The role of phenotypic heterogeneity in collective cell migration

Rebecca Crossley
Wolfson College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy
Hilary 2026

For the greatest Nan that ever could have existed.
You gave your everything so that we could have the very best something.
I hope we continue to do you proud.

Acknowledgements

Even after being awarded the studentships that funded my DPhil (thanks to the Engineering and Physical Sciences Research Council and Wolfson College, Oxford), I never quite believed I would be writing my acknowledgements—not because I doubted I would finish, but because I planned to avoid this bit altogether.

However, if my DPhil has taught me anything, it is the importance of acknowledging contributions. Despite my name being on the front of this thesis, I would never have completed it without the support, patience and encouragement of others.

Firstly, to my family. Thank you for believing in me and letting me chase my dreams, despite the lack of financial incentives. Thank you for reading my papers, cheering on my "pretty pictures", and celebrating my achievements louder and prouder than anyone else. Dad, I would not be doing maths without the hours you spent teaching me your tricks (many of which got me into trouble) and pushing me to do better than 99.9%. Mum, thank you for showing me that no one gets to stand in my way, and that we all deserve to achieve our dreams. Jingles (aren't you lucky that this nickname is now permanently archived?), thank you for never knowing what I was on about, but always knowing when it mattered—and for learning π purely to spite me. William, thank you for proposing and giving me something other than this thesis to focus on for the last few years. Thank you for always believing in me, picking me back up on every bad day and for moving to Oxford to support me. Finishing this journey as your wife is the greatest joy.

I am also deeply grateful to the Mathematical Institute in Oxford, whose staff, students and researchers welcomed me into an extraordinary community. The WCMB, along with Wolfson, Pembroke and St Hugh's Colleges gave me a home in Oxford. To my collaborators, friends and colleagues (too many to name!) thank you for the inspiration and joy that came from sharing this journey.

Finally, I owe my greatest thanks to my supervisors. Philip, thank you for believing in me, welcoming me into the WCMB, and for bringing perspective and lightness to many difficult days. Your patience, generosity, and breadth of knowledge have shaped both this thesis and my development as a researcher. Ruth, much like many of our meetings, I am not entirely sure where to begin. Being assigned to one another without ever having met could have gone very differently, yet your unwavering support, advocacy and honesty carried me through every stage of my DPhil. Knowing I had you on my side gave me the confidence to persist through even the hardest moments, and for this I will always be indebted. I am deeply grateful to you both for your guidance and mentorship, and I hope our collaborations continue in the future.

Statement of Originality

This thesis is submitted to the University of Oxford in fulfilment of the requirements of the Doctor of Philosophy. This thesis contains no material which has been previously submitted for a degree or diploma at this University or any other. The modelling, analysis and numerical results in Chapters 2, 4 and 5 of this thesis are the work of Rebecca M. Crossley under the supervision of Professor Ruth E. Baker and Professor Philip K. Maini. Some of the work in Chapter 1 was carried out by Rebecca Crossley in collaboration with Samuel Johnson, Erika Tsingos, Zoe Bell, Massimiliano Berardi, Margherita Botticelli, Quirine Braat, John Metzcar, Marco Ruscone, Yuan Yin and Robyn Shuttleworth. The work in Chapter 2 was carried out by Rebecca Crossley in collaboration with Tommaso Lorenzi, who contributed to the supervision of this work, alongside Ruth Baker and Philip Maini. The work in Chapter 3 was carried out by Rebecca Crossley in collaboration with Carles Falcó and Ruth Baker. The work in Chapter 4 was carried out by Rebecca Crossley in collaboration with Tommaso Lorenzi and Kevin Painter, who contributed to the supervision of this work, alongside Ruth Baker and Philip Maini. The work in Chapters 1, 2, 4 and 5 have been published in [75], [77], [79] and [76], respectively.

Abstract

Understanding collective cell migration requires robust mathematical models that effectively bridge the gap between individual cell behaviours and population-level dynamics. This thesis develops and analyses a suite of mathematical models for cell invasion into extracellular matrix (ECM), that are systematically derived from individual-based frameworks to ensure consistency in the inclusion of key biological mechanisms. Beginning with a single cell type, a volume-filling framework is introduced that accounts for spatial constraints imposed by both cells and ECM. These effects are incorporated coherently across all migration and proliferation processes to investigate their influence on invasion speed and wave structure by comparison to existing models in the literature. To further these insights, a variational principle approach is developed through the lens of optimal control theory, that characterises a new and improved lower bound on the speed of cell invasion in multi-species systems.

Building on the homogeneous foundation, the volume-filling model developed initially is then extended to incorporate phenotypic heterogeneity. Firstly, by introducing two distinct sub-populations: one specialising in movement and ECM degradation, and the other in proliferation. Secondly, by replacing discrete phenotypes with a continuum of states, allowing for a more flexible representation of phenotypic variation across diverse biological contexts, including tumour invasion and immune cell exhaustion. By exploring different environmentally-dependent phenotypic switching mechanisms, it is demonstrated that phenotypic switching fundamentally alters both the cell invasion dynamics and the structure of the travelling fronts. Overall, this work demonstrates the importance of systematically deriving population-level models from their individual-based counterparts and reveals that phenotypic heterogeneity plays a crucial role in shaping migration patterns. These insights highlight key features for consideration when modelling collective migration and help improve our understanding of cell invasion and its broader biological implications.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This introduction contains elements of a literature review completed in collaboration with S. Johnson, E. Tsingos, Z. Bell, M. Berardi, M. Botticelli, Q. J. S. Braat, J. Metzcar, M. Ruscone, Y. Yin and R. Shuttleworth regarding mathematical modelling of the extracellular matrix. It has been published in Frontiers in Cell and Developmental Biology [75]. The thesis author was first author of this review, held joint responsibility for all aspects of the research, and worked solely on the material presented in this chapter.

1.1 Overview

Across all levels of biology, we observe collective behaviour—the coordinated dynamics of many interacting individuals producing patterns at the population level [93, 113, 158, 227, 243]. Whether observed in ecological populations, microbial colonies, or the cells that form tissues, coordinated dynamics arise from local interactions between individuals and give rise to emergent patterns at the population scale [303, 338]. These systems are inherently symbiotic, comprising at least two interdependent components: a living species and the environment in which it resides. In multicellular systems, collective cell migration is one of the clearest examples of such behaviour. During embryonic development, wound repair, and tumour invasion, groups of cells migrate cohesively by responding to a complex array of both intrinsic and extrinsic mechanical and biochemical cues coming from other cells and their local environment [284]. These interactions, which are often nonlinear, dynamic, and bidirectional, allow cells to generate forces, remodel their micro-environment, and coordinate their movement, thereby producing large-scale invasion into surrounding tissue [5, 331].

Cell migration plays a central role in development, regeneration, and disease progression [64]. As such, understanding how microscopic, individual cell-level processes combine to generate macroscopic, tissue-level patterns of invasion is a central challenge in mathematical biology. Two main forms of collective cell migration are commonly distinguished [330]. Epithelial migration involves the coordinated motion of continuous sheets of tightly packed cells that deform and move coherently, as in wound closure [193, 284]. In this case, it is the intercellular junctions between cells that transmit forces and maintain tissue integrity, enabling collective movement [95]. Mesenchymal migration, by contrast, describes the motion of loosely connected cells moving in a more fluid and deformable manner, as seen during embryonic development or tumour invasion [5, 284]. This thesis focuses on mesenchymal cell migration, where groups of cells invade complex environments whilst interacting with neighbouring cells and the substrate they sit within.

Experimental studies have shown that invasive tumour cells can lose their migratory phenotype when placed in a non-invasive environment [161], and that leader–follower roles in neural crest collectives can dynamically interchange depending on cell positioning and local conditions [209, 211]. Such findings emphasise that feedback between cells and their environment is fundamental to understanding collective migration.

In this thesis, the primary micro-environment that we will consider cells having to encounter is the extracellular matrix (ECM). The ECM provides both a barrier to movement and a substrate on which cells can exert forces in order to be able to move [39, 75, 361]. Migrating cells extend protrusions, establish and release adhesions, proliferate, and secrete matrix-degrading enzymes, all while responding to the structural and biochemical cues encoded within the ECM. These coupled processes give rise to invasion fronts whose qualitative behaviour resembles that of ecological colonisation or classical reaction–diffusion waves [224].

Mathematical models provide a principled means of formalising this multi-scale behaviour. They allow us to translate biological hypotheses into quantitative predictions, to test how simple interaction rules generate emergent collective behaviour, and to bridge scales from individual cells to tissue-level organisation. Individual-based models (IBMs) [249, 268] have the ability to capture stochastic movement, volume exclusion, heterogeneity, and local environmental feedback directly [35, 216, 223], while continuum models, typically formulated as partial differential equations (PDEs), describe the evolution of population densities at larger scales [16, 145, 195, 259]. Establishing a systematic connection between these descriptions through coarse-graining

allows the resulting continuum equations to inherit mechanistic meaning from their microscopic origins [128, 310]. Such connections are particularly valuable when seeking analytical results, such as invasion speeds or front shapes, or when comparing models against classical phenomenological frameworks.

The central aim of this thesis is to construct and analyse a hierarchy of continuum models for collective cell invasion into ECM, each one derived systematically from an underlying IBM. Beginning with a homogeneous population, we incorporate random motility, proliferation, crowding effects, and ECM degradation in a consistent manner. These models are then extended to incorporate discrete and continuous forms of phenotypic heterogeneity. Travelling wave analysis is used throughout to link microscopic processes to macroscopic invasion dynamics. As such, this thesis also develops and presents a new optimal control approach to using the variational principle to find the travelling wave speed of invading populations.

In the remainder of this introductory chapter, the biological context of the processes modelled in this thesis is summarised, followed by a review of the mathematical structures on which the subsequent chapters build.

1.2 Biological background

Collective movement is a pervasive feature of living systems at all scales and is perhaps best understood when viewed through an ecological and evolutionary lens. One influential perspective on cancer was proposed by Nowell in 1976 which emphasised that selection within tumours is driven by micro-environmental conditions, yielding increasingly adapted subclones [240]. More broadly, Hanahan and Weinberg formalised the set of cellular capabilities required for tumour progression as the "hallmarks of cancer", that include sustaining proliferation, resisting cell death, the acquisition of invasive behaviour and the ability to metastasise [134, 135]. These ideas underscore two points central to this thesis. Firstly, biological collectives can be modelled as populations interacting with each other and their environment, and secondly, heterogeneity within the population and heterogeneity of the environment are fundamental drivers of emergent behaviour [155].

Cell migration is fundamental to the formation, maintenance, and disruption of tissue organisation [333]. Although cells may migrate individually, many contexts involve coordinated motion by groups of cells [284]. Solid tissues (and solid tumours, such as carcinomas, in particular) can be viewed as systems composed of at least two interdependent compartments: the cellular compartment (cancer, stromal and normal

cells) and the micro-environmental compartment (blood supply, pH, oxygen, ECM, soluble factors) [115]. Mesenchymal invasion, in particular, provides a representative setting in which cells move through complex, confining environments by exerting forces on their surroundings and degrading the ECM in their path [5]. The dynamics of such systems are typically governed by a small number of core processes: cell motility, directed movement in response to environmental gradients, cell proliferation and death, the creation of space through ECM degradation, and the spatial limitations imposed by crowding and volume exclusion.

When these processes act together, populations of cells can expand into new territory and form coherent invasion fronts. Even in simplified settings, the combined action of random motility and proliferation is sufficient to generate constant-speed travelling waves, as captured by the classical Fisher–KPP equation (described in more detail in Section 1.3) [109, 174, 224]. The addition of environmental feedback, such as local ECM degradation, introduces further layers of complexity and produces invasion behaviour that depends on the structure, composition and properties of the surrounding ECM [75, 216].

Cell motility. Cell motility lies at the heart of invasion. At the single-cell level, movement arises from stochastic protrusion and retraction of the actin cytoskeleton, the formation and turnover of adhesions, and fluctuations in intracellular signalling [8, 185]. Over larger scales, these microscopic fluctuations manifest as an effectively random component of motion. Although cells can exhibit more complex behaviours, such as directed migration in response to chemical or mechanical gradients [110], the focus of this thesis is on environments where stochastic motility is the dominant mechanism driving dispersal into unoccupied regions. Experimental evidence indicates that such random migration, when combined with the appropriate mechanisms, such as cell population growth, can be sufficient to produce persistent invasion fronts [64].

Cell proliferation. Cell proliferation is the second key driver of population expansion. Through repeated cell cycles involving growth and division, cell populations increase in density and exert pressure on surrounding tissue. Proliferation is regulated by a variety of biochemical and mechanical cues, including nutrient availability, physical confinement, and local cell density. In multicellular systems, these factors vary spatially, producing heterogeneous proliferation rates across tissues or tumours [117, 149, 188, 282]. Importantly, proliferation is also limited by available space: in

crowded environments, cells may experience contact inhibition of proliferation, reflecting the fact that division requires sufficient free volume for daughter cells. This space limitation motivates the volume-exclusion mechanisms incorporated into the models developed in this thesis, in which both movement and growth slow down as densities approach a maximal packing limit.

In many biological tissues, cell death acts alongside proliferation to regulate population size and composition [101, 148, 236, 365]. Although we often do not explicitly represent complex cell death pathways, cell loss arises naturally through environmental feedback: when ECM density is high or space is restricted, cells may be unable to divide or move into new regions, leading to stagnation or local decline. Furthermore, if the population exceeds the maximum capacity in a local neighbourhood, cell death may begin to occur. These effects appear implicitly through the space-limiting mechanisms described when considering the net proliferation of cells in isolation.

Cell-ECM interactions. A defining feature of mesenchymal invasion is the ability of cells to remodel their environment. Most relevant in this thesis is the ECM, which encompasses all biological components outside of cells. The ECM itself is a complex, yet highly organised network of more than 150 different ECM-associated proteins, collagen and other molecules that provides mechanical support and biochemical cues to the surrounding cells, and varies with its location in an organism [39, 62, 141, 159, 254, 301, 367, 373]. As such, it has a range of complicated properties, including its stiffness, porosity, elasticity and role during durotaxis [62, 74, 97, 136, 142, 179, 277, 349], that are hard to describe generally as they vary significantly between tissues. As can surely be imagined, the resulting cell-ECM interactions modulate a vast number of intracellular signalling pathways, which result in changes to migratory behaviour, proliferation, and adhesion during cell migration [22, 137, 159, 222, 291, 364].

The intricate hierarchical architecture structure of the ECM involves collagen fibres, cross-linking, and nonlinear mechanical behaviour (as described in Table 1.1 and the schematic in Figure 1.1 [75]), but this thesis aims to capture the high level impacts of the ECM, and thus treats the ECM as a continuous scalar field that impedes movement and proliferation, and that is degraded locally by cells [19]. This coarse description, despite its simplicity, captures a key biological mechanism: cells must create space by degrading the ECM in order to invade, and ECM density directly influences motility and growth [153, 230, 231, 346]. As such, these processes are the focus of this thesis, and further information regarding the biochemical and biomechanical properties of the ECM can be found elsewhere [62, 75, 111, 159, 222].

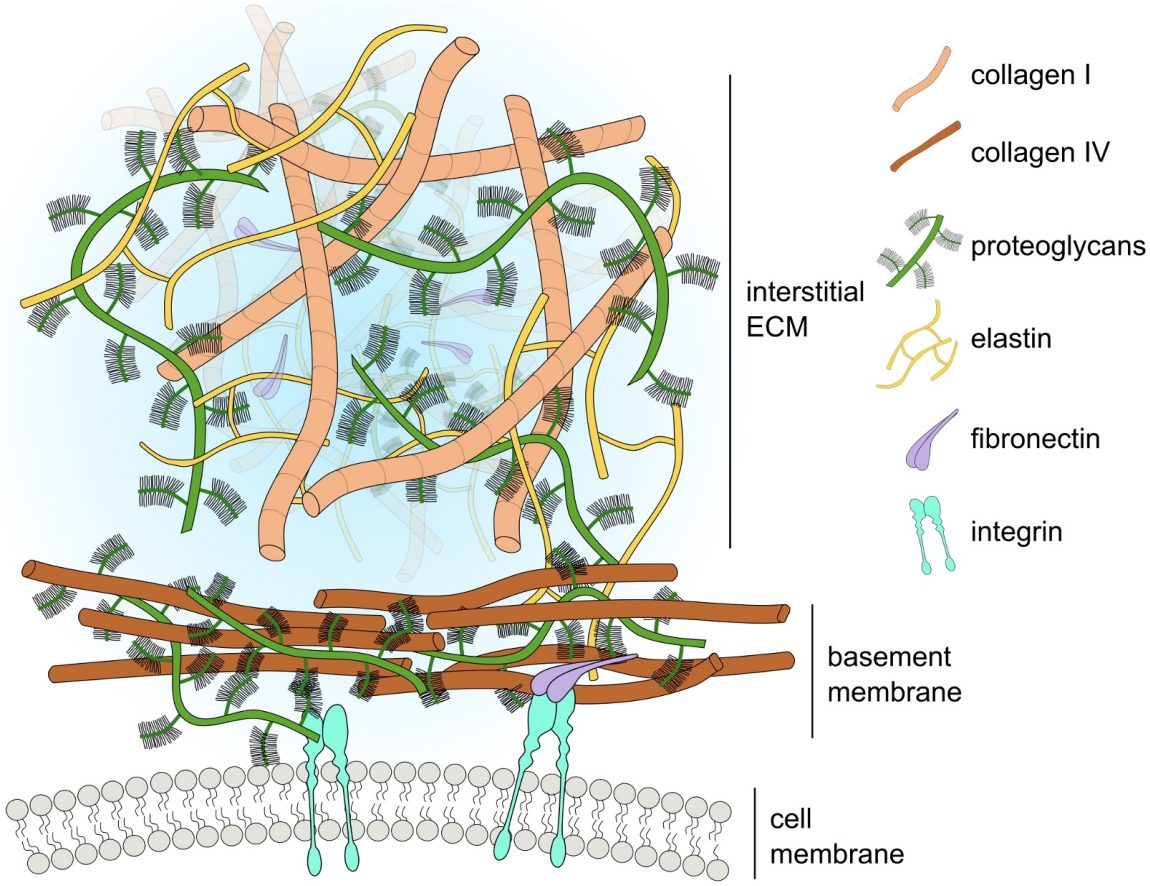

Figure 1.1: Schematic representation of the major ECM compartments and their components. Macromolecular components of ECM (size on the order of hundreds of nanometres) are smaller than cells (size on the order of tens of micrometers). However, many components, such as laminin and collagen, polymerize to form extended sheets or fibres (size order ranging from tens of micrometers to a few millimetres). Small molecules, such as cytokines, enzymes, or soluble factors, are not represented for visual clarity [75].

Biologically, there are two complementary strategies that are used by migrating cells to overcome the physical barriers of the ECM. In protease-dependent invasion, cells secrete or present matrix-degrading enzymes, including both soluble and membrane-tethered matrix metalloproteinases (MMPs), which locally cleave collagen and other matrix components to open paths for cell migration [49, 99, 346, 352]. Alternatively, protease-independent invasion exploits morphological and mechanical plasticity to allow cells to deform and squeeze through pre-existing pores in the ECM without extensive degradation. The ECM thus acts simultaneously as an impediment to migration and as a consumable substrate whose loss facilitates invasion.

Cells commonly switch between these strategies depending on the ECM archi-

tecture, confinement and protease availability. MMPs, particularly those that are membrane bound and have a very small diffusivity [356], have been identified as the key mediators of protease-dependent invasion in dense collagen matrices, indicating circumstances where protease-dependent invasion can dominate and is necessary for cancer invasion [152, 271, 288]. Given this, we focus solely on protease-dependent invasion by membrane-bound MMPs, which dominate over diffusible MMPs, due to their short-range action and fast degradation [261, 353], for the remainder of this thesis.

These four processes—motility, proliferation, death, and ECM degradation—operate within a finite spatial domain where physical crowding imposes strict constraints. Contact inhibition ensures that neither movement nor division can occur when local densities approach a maximum packing threshold. The resulting effects strongly influence invasion dynamics, particularly in systems where both cells and ECM occupy appreciable volume. This thesis therefore focuses on incorporating volume-filling constraints consistently across all mechanisms governing collective cell migration, reflecting the biological reality that crowding affects not only division but also the ability of cells to move and to degrade their surroundings.

Phenotypic heterogeneity. Finally, real tissues rarely consist of homogeneous populations. Empirical studies have repeatedly observed spatial structuring linked to different cell behaviours: for example, tumour peripheries often contain more proliferative and motile cells exhibiting glycolytic metabolism, whereas tumour cores contain relatively quiescent cells with lower motility [117, 149, 188, 282], analogous to niche-partitioning and invasion fronts observed in ecological systems [41]. The differences in cell behaviours can arise as a result of genetic variation, epigenetic modifications, or reversible phenotype switching. For example, throughout the human body, genetically identical cells often evolve into distinct cell types with very different functions, such as nerve cells for sensing, epithelial cells for lining, or stem cells for regeneration. These emergent differences between cell phenotypes often determine an individual cell's fitness (such as their net growth) in a given local environment [248], and may be either fixed or dynamic over time.

In a constant environment, a specialised phenotype may dominate, whereas variable environments often favour strategies that mitigate risk. Experiments have suggested that cells can transition between phenotypes depending on environmental conditions, such as nutrient availability, ECM density, or mechanical confinement. For example, cancer cells use the so-called 'invasion-metastasis cascade' to adapt their

phenotype in evolving micro-environments [156, 157]. In general, there are two broad strategies for cell phenotypic change that have been particularly well-studied.

Phenotypic plasticity allows individual cells to alter their phenotype in response to sensed cues, improving their behaviour to best suit the local environment [325, 366]. However, cells do not have infinite ability to change phenotypic state [91], and thus phenotypic plasticity is best suited to cells when cues are reliable and changes in the local environment are predictable [143]. Examples include transitions between cell types during development and epithelial–to–mesenchymal transitions (EMT) that change adhesion, motility and invasive potential [130]. In cancer, reversible switches between invasive/slow-proliferating and non-invasive/fast-proliferating states have been documented (*e.g.*, melanoma) and are associated with therapy resistance [3, 132, 166].

On the other hand, bet-hedging as a strategy is used to spread risk across a population of cells by producing diverse offspring phenotypes irrespective of immediate cues [238]. This may be advantageous when environmental fluctuations are irregular or unpredictable [65, 68, 126, 146, 265, 313]. Classic examples include bacterial persistence under antibiotic stress [21], stochastic switching to enhance survival in fluctuating nutrient conditions [1, 24, 44, 256, 295, 314, 339], and drug-resistant subpopulation formation during tumour development [180, 302].

Clearly, both phenotypic plasticity and bet-hedging strategies appear to play roles in maintaining diverse phenotypes within invasive populations, allowing cells to respond efficiently to fluctuating environments. Understanding how these adaptive strategies interact with spatial structure and with cell motility is essential for explaining collective migration. This thesis further investigates the role of phenotypic heterogeneity during cell migration by examining the impact of both discrete and continuous phenotype structures, enabling a systematic comparison of how variations in cell behaviour influence invasion speed, population composition, and wave front morphology.

1.3 Mathematical background

Having established the biological basis of cell migration and ECM structure, this section focuses on how these processes can be represented mathematically. Mathematical and computational models have long been used to describe biological phenomena and to test hypotheses that are difficult to probe experimentally. They provide a complementary means of investigation to laboratory studies, enabling exploration of

parameter regimes, spatial and temporal scales, and mechanistic hypotheses that are often inaccessible *in vitro* or *in vivo*.

Mathematical models of invasion span a wide spectrum of spatial and temporal scales. At the microscopic level, stochastic IBMs explicitly track the trajectories of single cells and describe migration through rules that govern cell interactions with one another and their environment [9, 72, 217, 336, 350, 357]. At larger scales, deterministic continuum models describe the collective dynamics of populations moving through tissue or extracellular environments containing chemical attractants, adhesive substrates, or other cell types that influence migration [11, 43, 59, 69, 76, 77, 116, 198, 202, 207, 212, 235, 261, 333]. Continuum models often take the form of systems of differential equations representing the dynamic evolution of cell densities, and can be advantageous due to their ease of analytical tractability and computational implementation.

In designing an experiment in a laboratory, careful selection of the appropriate equipment and methodology must be used. This same principle applies to mathematical modelling. There is no single "best" modelling framework: the most suitable approach depends on the scientific question, the biological scale of interest, and the level of detail required. Over the years, a variety of mathematical and computational models have been used to explore the role of the ECM during cell migration [75]. An overview of these is shown in Figure 1.2. The models developed in this thesis are continuum PDEs, but their formulation is grounded in detailed underlying IBMs. The remainder of this section outlines the key mathematical ideas and existing literature that underpin these approaches.

Discrete models. IBMs provide an explicit representation of individual cells as discrete entities. Each agent moves, proliferates, interacts with neighbours, and modifies the ECM according to prescribed stochastic rules. A key strength of discrete cell models is the capacity to incorporate heterogeneity naturally: each agent can carry its own phenotypic label, an internal state, mutation history or stochastic decision rule. This makes them indispensable for studying rare, stochastic events and for deriving mechanistic hypotheses about how single-cell processes aggregate into collective behaviour (*e.g.*, leader–follower patterns) [11, 345].

Two broad classes of IBMs exist: lattice-based approaches (cellular automata, lattice gas cellular automata, Cellular Potts Models) and off-lattice, force-based approaches (centre-based models, subcellular element models, boundary-tracking and vertex models) [90, 217, 326]. Lattice-based approaches are often computationally

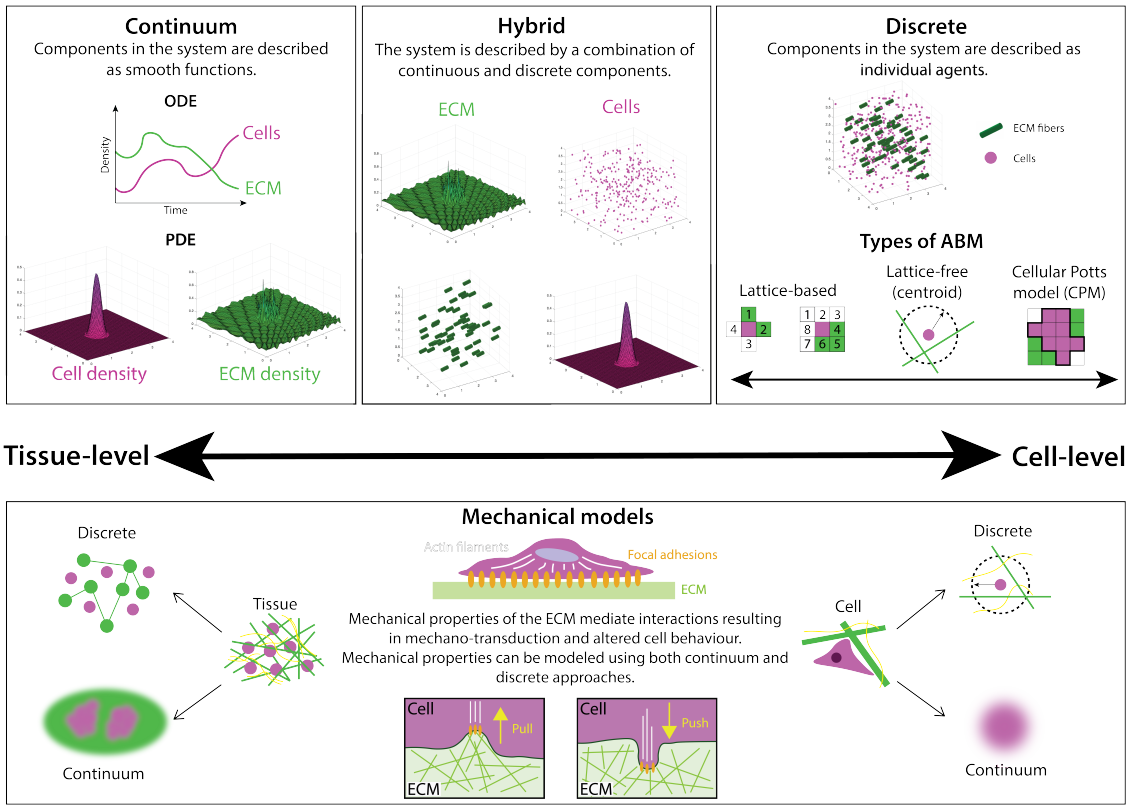

Figure 1.2: There are many different mathematical modelling approaches to studying ECM and cell–ECM interactions, ranging from continuum to discrete frameworks. Continuum models describe the densities of cells and ECM components with either ordinary differential equations (ODEs), typically used to describe total population size via a spatially-averaged density, or PDEs that often describe both the spatial distribution and temporal evolution of a population density. On the other hand, discrete frameworks, such as IBMs, represent biological components as separate, interacting agents. IBMs can be further categorised as either lattice-based, where positions are restricted to a finite set of points, or lattice-free, where positions lie in a continuous range. In some cases, multiple modelling approaches are combined in a hybrid fashion to balance the computational complexity of frameworks with many interacting agents with the accuracy of their outputs. Mechanical models are a subset of models focused on representing material properties such as stiffness and elasticity. Both continuum and discrete mechanical models are utilised to study the mechanical interactions between cells and ECM [75].

simple to implement and interpret. Cellular automata and lattice-gas models track occupancy and local state changes on a fixed grid and are therefore efficient for simulating large populations over long times [10, 11, 90, 139]. Off-lattice, force-based models represent cells in continuous space and update the positions of the cells accord-

ing to mechanical interactions (adhesion, repulsion, drag) and active motility forces. Although these have been extremely valuable in understanding the role of packing, jamming and mechanical feedback during cell invasion [150, 169, 217, 219, 250], and in particular in exploring the role of collagen fibres, cross-linking, stiffness and orientation in cell-ECM interactions [10, 11, 81, 129, 138, 251, 273, 280, 294, 297, 298], the focus on forces makes off-lattice models less well-suited for the mechanisms studied in this thesis. As such, we focus on on-lattice models.

Despite the choice of IBM, the principal limitations include the high computational cost, the need to calibrate many parameters, and limited analytic tractability. To overcome this when further analysis is required, IBMs have been used as the starting point for coarse-graining procedures that yield continuum descriptions of a larger population of cells.

Deriving continuum models from underlying discrete descriptions. Several methods exist for deriving continuum equations from individual-based descriptions. The approach used in this thesis follows a master-equation framework in which the population-level density is obtained by coarse-graining the stochastic dynamics of discrete agents [186, 257, 310, 311, 376]. In this setting, a mesoscopic balance law is first constructed for the expected occupancy of discrete lattice sites, and a continuum limit is obtained through Taylor expansion in the limit of small lattice spacing or large population size. This procedure yields PDEs whose terms correspond directly to the underlying microscopic rules: nonlinear diffusion emerges naturally from movement restrictions, while reaction terms reflect proliferation or ECM degradation mechanisms. The principal advantage of this approach is its mechanistic transparency. Each term in the resulting equations retains a clear and interpretable link to an individual-level process, in contrast to phenomenological models where nonlinearities are often introduced heuristically without reference to microscopic behaviour. Moreover, this class of derivation guarantees consistency between movement, growth, and environmental interaction terms, ensuring that space-limiting effects are not inadvertently incorporated into one part of the model but omitted from others.

Although coarse-graining a master equation provides a rigorous and intuitive route to finding a continuum limit, it is only one approach among many. An alternative and widely used framework is based on kinetic theory. In kinetic approaches, the microscopic state of the system is described by a probability density over position and velocity, typically governed by a Boltzmann, Fokker–Planck, or velocity-jump equation [213, 214, 234]. Diffusion approximations or parabolic scalings then lead

to macroscopic PDEs that resemble transport or reaction–diffusion equations. Such formalisms have been particularly successful in modelling chemotaxis and movement with velocity persistence, as in the classical derivation of the Keller–Segel model from velocity-jump processes [162, 163, 260, 304]. Kinetic approaches excel when directional persistence, turning frequencies, or velocity-dependent interactions are central to the dynamics. However, encoding strong short-range crowding or finite-volume effects within kinetic frameworks can be cumbersome, and closure approximations are often required to handle multi-particle interactions. Furthermore, kinetic equations are typically high-dimensional, and their reduction to tractable macroscopic models may depend sensitively on the assumed scaling regime.

Statistical mechanics approaches provide a third route to deriving continuum limits [2, 360]. These methods begin with numerous equations governing the joint distributions of increasing numbers of agents, before employing moment-closure approximations, such as mean-field or pair-correlation closures, to truncate the representations. Such techniques are powerful for deriving complex nonlinear diffusion and cross-diffusion systems, especially in contexts involving short-range interactions such as volume exclusion, adhesion, or mechanical repulsion. By retaining explicit information about pairwise correlations, these approaches can capture effects sometimes overlooked, such as clustering or density-dependent motility. Their main limitation lies in the necessity of choosing an appropriate closure; different closure schemes can lead to quantitatively or even qualitatively different PDEs. Moreover, higher-order closures quickly become analytically intractable, and their interpretation can be challenging when correlations are strong or spatial structure is complex.

A further class of methods relies on homogenisation and multiscale asymptotics [86]. In these approaches, the continuum limit is obtained by identifying separate spatial or temporal scales, for example, microscopic cell movement versus macroscopic invasion dynamics, and exploiting this scale separation to derive PDEs that average over the finer scale structure. Homogenisation has been used extensively in modelling transport in porous media, cell movement in heterogeneous environments, and diffusion through periodically structured matrices [286]. These methods are particularly well suited when the microscopic environment contains repeated structure, but their applicability depends on the existence of clear scale separation and may fail in situations where there is local, highly complex ECM architecture.

Finally, variational frameworks offer powerful tools for deriving macroscopic equations from energetic or probabilistic principles. In this setting, many nonlinear diffusion equations, including porous-medium and Fokker-Planck equations, arise as

gradient flows in Wasserstein space [98, 205]. Such formulations provide insights into stability, equilibrium structure, and dissipative properties. However, variational approaches are most naturally suited to problems that can be easily characterised as a free-energy minimisation and may not readily incorporate processes such as proliferation or degradation without further detailed information.

Each of these methods has distinct strengths and limitations. Kinetic and statistical mechanics approaches can capture directional persistence and pairwise correlations more effectively than master equation methods, while homogenisation excels in structured environments. Conversely, the master equation approach adopted in this thesis is selected as it is particularly well suited to invasion processes driven by local mechanisms including volume exclusion, proliferation, and ECM degradation. These are all naturally expressed as probabilistic transitions on a lattice and coarse-grain to produce continuum equations that preserve mechanistic clarity. Advantageously, the method is also relatively straightforward to implement in one-dimension and the resulting systems of equations are analytically tractable for travelling wave solutions.

Continuum models. When conducting studies at the tissue scale, both the cells and individual ECM components are of a small size relative to the characteristic length scale of tissue, and are often densely packed. Thus, in many cases, a continuum framework for modelling spatially averaged behaviour at the cell scale and its effects at the macroscopic scale of the tissue is appropriate. At tissue scales, fully continuum models describe cell populations as spatially continuous densities that evolve in time, typically via PDEs [48]. Often these equations will be accompanied by an equation describing the dynamics of an environmental feature (such as the ECM) over time, which may take the form of a coupled ODE or PDE, depending on whether it varies in space and/or time [151]. Continuum models are the natural choice when population-level questions, such as the existence of patterning, or trying to quantify the invasion speed, are central and when the population is sufficiently dense that mean-field approximations are reasonable [107].

Continuum cell and ECM models typically include random motility (diffusion), volume-filling and crowding, logistic proliferation, and ECM degradation [13, 60]. Some continuum models are exactly solvable [262], however, in more complex cases where exact solutions cannot be obtained, there exists a suite of numerical techniques to approximate solutions at a relatively low computational cost. In general, ODE and PDE models are amenable to mathematical analysis, for example using boundary layer techniques [106], or asymptotic methods [164], wherein the behaviour of a system

under certain limits (for example, long-time behaviour) may be studied, as well as standard travelling wave or stability analyses that yield macroscopic invasion laws and conditions for pattern formation [69, 77, 100, 116].

The archetypal continuum model used to describe collective cell migration is the Fisher–Kolmogorov–Pietrovskii–Piskunov (Fisher–KPP) equation. This was first proposed in the context of the spread of an advantageous gene [109, 174], but has since been well-studied and seen a broad spectrum of applications in the natural sciences: most notably in cell biology [121, 196] and ecology [178, 242, 244]. This seminal reaction–diffusion equation explicitly considers changes in cell density with respect to space and time, and combines linear diffusion (random motility) alongside logistic growth [109, 174, 196]. In one spatial dimension, the Fisher–KPP model takes the form

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} + ru\left(1 - \frac{u}{K}\right), \qquad x \in \mathbb{R}, \tag{1.1}$$

where $u(x,t)$ is the cell density, $D > 0$ is the (constant) diffusion coefficient, $r > 0$ is the (constant) rate of proliferation and $K > 0$ is the (constant) carrying capacity. When paired with compactly supported or exponentially decaying initial conditions and boundary conditions connecting the state $u = 1$ and the state $u = 0$, then this model exhibits constant profile, constant speed solutions, which can be analysed using a process known as *travelling wave analysis* that seeks solutions of the form $u(x,t) = U(x - ct)$, where c is the constant travelling wave speed [51, 226]. In doing this, an ODE is obtained describing the steady wave profile in a moving coordinate frame. This reduction allows for the determination of wave speed ($c \geq 2\sqrt{Dr}$ for Fisher–KPP), front structure, and parameter sensitivity [121], and has been extended to incorporate directed movement in response to gradients: chemotaxis (soluble cues), haptotaxis (ECM or adhesion gradients) or durotaxis (see, for example [58, 116, 118, 145, 252, 307]).

Many biological invasion models extend the PDE structure by incorporating nonlinear diffusion to account for crowding [310], environmental coupling to describe ECM dependence [43, 69], or multiple interacting species to represent phenotypic diversity [311, 321]. In the context of ECM modelling, continuum models have also been extended to study the role of fibre orientation [54, 84, 96, 108, 245, 246, 251, 372], MMPs for degrading fibre cross-links [11, 58, 237, 326, 356] in a variety of contexts, and in modelling wound healing in particular [83, 184, 247, 258, 305]. They have also been coupled to the invasion of cells to examine the role of haptotaxis [281], haptokinesis [112], durotaxis [140, 167, 168], ECM-mediated cell proliferation [275],

and ECM production and degradation [361]. However, as previously stated, these mechanisms are beyond the scope of this thesis and further information can be found elsewhere [58, 75, 118, 144, 304].

To capture phenotypic heterogeneity of cells, continuum models can be structured by introducing a phenotypic variable that can be discrete (as in the case of the classic Lotka-Volterra model [224, 263, 351]) [53, 63, 79, 105, 318] or continuous, whereby population members exhibit a variety of behaviours to different degrees [37, 194]. These phenotypes may truly represent a number of different cell features, such as age, motility, invasiveness or drug resistance, but may not always be easily measurable experimentally.

In discretely structured phenotype models, phenotype change can be included by introducing switching terms between the different sub-populations of cells [318]. Various models of this form have been developed as the continuum limit of an underlying IBM. In particular, the model presented in [310, 311] demonstrates the impacts of volume-filling effects on collective cell invasion of multiple species. Furthermore, an example of an extended proliferation-invasion model of three interacting sub-populations of glioblastoma cells is presented in [220], which includes cross-species density-dependent diffusion terms, and takes account of volume-filling effects.

Alternatively, structured PDEs account for continuous trait distributions and their dynamics under selection, mutation and environmentally dependent fitness [37, 189, 191, 194]. Such models link ecological and evolutionary dynamics and have been used to study bet-hedging, EMT transitions, and therapy-driven selection in various geometries and varying environmental conditions [189, 342].

The models analysed in this thesis exhibit both discrete and continuous phenotypic structuring across a range of biological scenarios, as well as novel terms describing volume-filling effects of the cells and the ECM. However, unlike some of the aforementioned models, the terms in the models developed in this thesis always emerge systematically from the coarse-grained IBM rather than being postulated in advance, providing a deeper insight into how microscopic mechanisms influence macroscopic invasion behaviour.

1.4 Thesis structure, scope and aims

As George Box famously stated, "all models are wrong, but some are useful" [40]. This statement captures the philosophy that underpins this thesis: mathematical models are simplified representations of reality whose value lies not in their precision but in

their capacity to provide insight into complex systems. By embracing simplifications deliberately, essential mechanisms that drive biological behaviour can be isolated and studied to gain meaningful insights that guide both theory and experiment.

The central aim of this thesis is to derive and analyse mathematical models that describe the collective migration of cells as they invade into ECM. In particular, we develop new continuum limits of IBMs that incorporate key biological processes—such as volume exclusion, phenotypic heterogeneity and environmental feedback—to better understand how collective invasion emerges from individual cell interactions. By systematically coarse-graining these discrete models, we obtain continuum formulations that retain mechanistic consistency with their microscopic origins, enabling direct comparison between individual- and population-level perspectives.

Throughout, the ECM is treated as an isotropic, scalar field whose density restricts cell movement and proliferation, and is degraded locally by cells. Mechanical forces, anisotropy, fibre alignment, and subcellular processes are not explicitly modelled, nor are detailed biochemical signalling pathways. These simplifications are motivated by the desire to retain a clear and mechanistically interpretable connection between the individual-based rules and the resulting continuum equations. They also allow for a rigorous analytical treatment using travelling wave methods.

The thesis begins in Chapter 2 with the derivation of a continuum model describing a single homogeneous cell population invading into ECM. This model is constructed by coarse-graining an IBM that consistently accounts for volume exclusion across all mechanisms influencing migration, including random motility, proliferation and degradation of the surrounding ECM. By analysing travelling wave solutions and their dependence on key parameters such as the initial ECM density and the degradation rate, we demonstrate how volume exclusion modifies invasion speed compared with classical Fisher–KPP type formulations and other models in the literature that do not explicitly consider the impact of volume exclusion [43, 69, 109]. Comparison of analytical and numerical predictions reveals that linearisation of nonlinear effects often underestimates the complexity of invasion dynamics, yet the new model still recovers the qualitative behaviours of simpler, volume-filling-free frameworks in most parameter regimes [69, 100, 224].

In Chapter 3, the wave speed of a travelling front is analysed using a variational approach. Beginning with single species models, the variational principle [26, 27, 28, 320] is recast as an optimal control problem [176, 344], before being used to improve the characterisation of the minimum bound for the travelling wave speed in a number of canonical models [290, 306]. Thereafter, this analysis is extended to models of two

interacting species which couple a parabolic PDE for one cell population with an ODE describing the dynamics of another quantity, usually either another cell population, chemical or local micro-environment [43, 69]. Using the optimal control framework, a novel variational principle is derived for two-species reaction–diffusion equations and a new, more accurate, lower bound on the travelling wave speed is found for models of cell invasion into ECM described in Chapter 2.

Chapter 4 extends the coarse-graining modelling framework to account for cellular heterogeneity by introducing two interacting sub-populations: one primarily motile and degrading, the other primarily proliferative. This extension allows investigation of the so-called 'go-or-grow' dichotomy, whereby cells alternate between invasive and proliferative states. While previous studies have explored this concept phenomenologically, most neglect the fact that cells can dynamically switch between phenotypic states in response to environmental conditions. Here, we derive a continuum model for two coupled cell populations with explicit phenotypic switching mechanisms, which may be constant, stochastic or environment-dependent, and study how these affect collective migration. By analysing the resulting travelling wave solutions, we identify distinct invasion speeds and front compositions corresponding to different switching dynamics, suggesting that macroscopic observables could, in principle, be used to infer underlying cellular strategies.

The following chapter, Chapter 5, generalises the approach used in Chapter 4 to a continuum of populations with many coexisting phenotypes. Starting from an IBM consisting of N discrete sub-populations, we take the continuum limit as $N \to \infty$ to obtain a structured PDE describing invasion into a general local environment. This formulation bridges the gap between discrete and continuous representations of phenotypic variability, allowing the study of populations that evolve smoothly across trait space. To illustrate the versatility of the framework, it is then applied to several exemplar biological scenarios, including T cell exhaustion, lineage differentiation during population expansion, and, obviously, collective cell migration into ECM. These examples demonstrate how the same mathematical structure can be adapted to diverse biological problems, highlighting the unifying nature of the methodology developed in this thesis.

Finally, this thesis concludes in Chapter 6 with a discussion of the main findings, as well as the limitations of the frameworks. Potential extensions and future work to mediate these are also outlined.

Together, the chapters present a coherent modelling framework that progresses from mechanistic individual-level rules, through systematically derived continuum

equations, to analytical characterisation and questions of model identifiability. All of the research presented here is theoretical in nature, encompassing model development, analysis and simulation. Although some work was completed in collaboration with researchers who are acknowledged at the outset of each chapter and in the associated publications [75, 76, 77, 79], my contribution lies in formulating consistent mathematical frameworks for multi-scale invasion, developing the methods required to study them (modelling and simulation), and interpreting their implications for collective cell behaviour through detailed mathematical analysis. The work as a whole emphasises that, while no model can capture biological reality in its entirety, careful simplification—guided by biological relevance and analytical rigour—can yield models that are not only useful but genuinely illuminating.

ECM component	Function	ECM Location
Proteins		
Fibril collagens, e.g., type I, II, III	Provides structural integrity and strength to the tissue.	Interstitial ECM
Network-forming collagen, e.g., type IV	Sheet-like ECM forms from fibres.	Basement membrane
Elastin	Provides tissue with elasticity and load-bearing capabilities.	Interstitial ECM
Glycoproteins		
Fibronectin	Cross-links with ECM proteins, influences cell adhesion and migration.	Interstitial ECM (also found concentrated near basement membrane)
Laminin	Forms a mesh-like network to aid in cell adhesion and migration.	Basement membrane
Nidogen	Promotes BM structural integrity by facilitating laminin and collagen type IV cross-linking.	Basement membrane
Proteoglycans	Regulate ECM structure, e.g., perlecan, performing a similar role to nidogen, and supports cellular functions by storing growth factors and other signaling molecules.	Interstitial ECM and Basement membrane
Enzymes		
Lysyl oxidase (LOX)	Mediates ECM structure, facilitating ECM protein cross-linking.	Interstitial ECM
Matrix metalloproteinases (MMPs)	Degrades ECM proteins such as collagens, fibronectin and laminin.	Interstitial ECM and Basement membrane

Table 1.1: Table describing some of the major ECM components alongside their functions and locations [75].

Chapter 2

A coarse-grained model for cell invasion into extracellular matrix

This work was completed in collaboration with T. Lorenzi, P. K. Maini and R. E. Baker, and is published in Studies in Applied Mathematics [77]. The author was responsible for all aspects of the research and writing presented in this chapter.

2.1 Introduction

Cell invasion is central to a variety of biological phenomena, playing a key role in morphogenesis, tumour growth and tissue engineering. In many instances of cell invasion, cells must migrate through ECM–a network of proteins and other molecules that can impact collective cell movement by reducing the space available for migration [85, 253, 261].

A wide range of mathematical approaches have been used to model cell invasion into ECM. These range from IBMs describing processes at the single-cell level to PDE models that describe population-level cell density dynamics [124]. However, existing PDE models often take inconsistent approaches to incorporating volume-filling effects—the fact that both cells and ECM occupy space, limiting movement and proliferation. Some models [100] include volume-filling effects of ECM in the proliferation term but not in the motility term, while others incorporate crowding effects into motility through phenomenological density-dependent diffusion [131, 202, 208, 300]. Crucially, these models are not always derived directly from explicit assumptions at the single-cell level, making it unclear how different modelling choices affect the resulting population-level dynamics.

This gap motivates the work in this chapter. The aim is to start from a cell-level description of invasion into ECM, explicitly incorporating volume-filling effects

in both cell motility and proliferation, and then to coarse-grain this to obtain the corresponding PDE model. By beginning at the individual-cell level, we can ensure a consistent treatment of volume-filling affects across both movement and growth mechanisms, and can directly examine how specific single-cell assumptions shape the emergent PDE.

In this chapter, we extend the methods in [310, 311] to develop an IBM for cell invasion into ECM, taking into account volume-filling effects, where both cell motility and proliferation are impacted by the presence of other cells or ECM components. We make the simplifying assumption that space is the only factor limiting cell invasion, whereas other models in the literature [12, 170] assume various other factors, such as nutrient-limited growth [327]. By coarse-graining this model, a limiting PDE description is formally derived and explored both analytically and numerically.

This framework allows us to carry out a systematic comparison between the population-level behaviours observed in this model and those predicted by existing simpler population-level models, including the classical Fisher–KPP model [109, 174] and with ECM invasion models [69, 100]. Each of these simpler models can be recovered from the model derived in this work by neglecting specific terms, such as those capturing ECM volume-filling effects. By comparing qualitative and quantitative predictions across these models, we investigate to what extent different assumptions at the cell-level influence the population-level description, and what this implies for the interpretation and application of ECM invasion models in biological contexts.

2.2 The mathematical model

In this section, we begin by developing a simple one-dimensional, on-lattice, IBM of cell invasion into ECM that incorporates cell motility and proliferation, degradation of ECM, and volume-filling effects. We then coarse-grain this model to formally derive a corresponding PDE model that comprises a system of coupled PDEs describing the evolution of the densities of cells and ECM [45, 310].

2.2.1 Individual-based model

In the simplified setting of this model, cells are represented as discrete agents that can proliferate and move on a one-dimensional uniform lattice, which constitutes the spatial domain, and can also degrade the surrounding ECM, which is regarded as being composed of discrete constitutive elements. The novel aspect of this model

is the introduction of volume-filling effects, similar to the model described in [220], which uses the methods described in [252], but extended to multiple populations [311].

Let the number of cells and ECM elements on lattice site $i \in \mathbb{Z}$ of width Δ at time $\tilde{t} \in \mathbb{R}^+$ of realisation $j = 1, 2, \ldots, J$ of the model be denoted, respectively, by $u_i^j(\tilde{t})$ and $m_i^j(\tilde{t})$. We assume that ECM elements have constant density and are chosen to occupy a volume equal to that of a cell, such that at most N cells or ECM elements occupy each lattice site.

The dynamics of the cells are governed by two mechanisms: proliferation, in which a cell places a daughter cell into the same lattice site it occupies; and motility, whereby cells can move to one of their two adjacent lattice sites. Moreover, ECM elements can be degraded by cells in the same lattice site as them. To incorporate volume-filling effects into the model, we prescribe that each lattice site has a maximum occupancy level N [329] and assume that:

(A1) if a cell attempts a move to a neighbouring lattice site, the probability that the move is successful decreases linearly with the occupancy level of the target site, such that the probability of a successful move to a target site with occupancy level N is zero;

(A2) if a cell attempts to proliferate, the probability of success decreases linearly with the occupancy level of the site where the cell is located, such that the probability of a successful proliferation event in a site with occupancy level N is zero.

As such, we can now mathematically define the functions describing the probabilities of movement and proliferation of cells.

Probability of cell movement. A cell attempts a movement in a time interval $[t, t+\tau)$ with probability $p_{\mathrm{m}} \in [0, 1]$, and the attempted movement from lattice site i to either of the neighbouring lattice sites $i \pm 1$ occurs with equal probability $1/2$. Using assumption (A1), we can define the probability of movement to the left, $T_{i-}^{\mathrm{m}^j}(\tilde{t})$, or right, $T_{i+}^{\mathrm{m}^j}(\tilde{t})$, during the time interval $[\tilde{t}, \tilde{t}+\tau)$ of realisation j, as

$$T_{i\pm}^{\mathrm{m}^j}(\tilde{t}) = \frac{p_{\mathrm{m}}}{2}\left(1 - \frac{u_{i\pm1}^j(\tilde{t}) + m_{i\pm1}^j(\tilde{t})}{N}\right). \tag{2.1}$$

Probability of cell proliferation. A cell in lattice site i attempts to proliferate in time interval $[t, t+\tau)$ with probability $p_{\mathrm{p}} \in [0,1]$. If proliferation occurs, then the cell places a daughter cell into the same lattice site as itself. Using assumption (A2), we can define the probability of proliferation, $T_i^{\mathrm{p}^j}(\tilde{t})$, during the time interval $[\tilde{t}, \tilde{t}+\tau)$ of realisation j, as

$$T_i^{\mathrm{p}^j}(\tilde{t}) = p_{\mathrm{p}} \left(1 - \frac{u_i^j(\tilde{t}) + m_i^j(\tilde{t})}{N}\right). \tag{2.2}$$

In order to ensure that the probabilities $T_{i\pm}^{\mathrm{m}^j}(\tilde{t})$, $T_i^{\mathrm{p}^j}(\tilde{t}) \geq 0$ are well-defined, the initial distributions of cells and ECM elements must be such that at most N cells or ECM elements occupy each lattice site to begin with. Under the assumption that the initial distributions of cells and ECM elements satisfy $0 \leq u_i^j(0) + m_i^j(0) \leq N$ for all $j = 1, 2, \ldots, J$ and $i \in \mathbb{Z}$, the definitions for the probabilities of cell movement and proliferation given by Equations (2.1) and (2.2) ensure that

$$0 \leq u_i^j(\tilde{t}) + m_i^j(\tilde{t}) \leq N \ \text{ for all } \ j = 1, 2, \ldots, J \ \text{ and } \ i \in \mathbb{Z} \ \text{ for any } \ \tilde{t} \in \mathbb{R}^+. \tag{2.3}$$

Furthermore, we assume that cells degrade surrounding ECM only at a constant rate. Hence, we may define the function describing the degradation of ECM.

Probability of ECM degradation. During the time interval $[\tilde{t}, \tilde{t}+\tau)$ of realisation j, an element of ECM in lattice site i is degraded by a cell on the same lattice site with probability $p_{\mathrm{d}} \in [0,1]$, such that the degradation per unit element of ECM, $T_i^{\mathrm{d}^j}(\tilde{t})$, is

$$T_i^{\mathrm{d}^j}(\tilde{t}) = p_{\mathrm{d}} u_i^j(\tilde{t}).$$

2.2.2 Corresponding coarse-grained model

In order to derive a coarse-grained description of the IBM, we introduce the average occupancy of lattice site i at time $\tilde{t}$ by cells and ECM elements over J realisations of the model, denoted, respectively, by

$$\langle u_i(\tilde{t}) \rangle = \frac{1}{J} \sum_{j=1}^{J} u_i^j(\tilde{t}) \quad \text{and} \quad \langle m_i(\tilde{t}) \rangle = \frac{1}{J} \sum_{j=1}^{J} m_i^j(\tilde{t}).$$

Coarse-grained model of cell dynamics. We proceed to derive a coarse-grained model by considering how the average occupancy in lattice site i changes during the time interval $[\tilde{t}, \tilde{t}+\tau)$. To do this, we use probabilistic approximations of the mean-field type which are frequently used for the coarse-graining of IBMs and involve

assuming independence of lattice site occupancies (see, for example, [257]). As such, we can write

$$\begin{aligned}\langle u_i(\tilde{t}+\tau)\rangle = \langle u_i(\tilde{t})\rangle & \\ &+ \frac{p_\mathrm{m}}{2}\langle u_{i+1}(\tilde{t})\rangle\left(1-\frac{\langle u_i(\tilde{t})\rangle+\langle m_i(\tilde{t})\rangle}{N}\right)\\ &+ \frac{p_\mathrm{m}}{2}\langle u_{i-1}(\tilde{t})\rangle\left(1-\frac{\langle u_i(\tilde{t})\rangle+\langle m_i(\tilde{t})\rangle}{N}\right)\\ &- \frac{p_\mathrm{m}}{2}\langle u_i(\tilde{t})\rangle\left(1-\frac{\langle u_{i+1}(\tilde{t})\rangle+\langle m_{i+1}(\tilde{t})\rangle}{N}\right)\\ &- \frac{p_\mathrm{m}}{2}\langle u_i(\tilde{t})\rangle\left(1-\frac{\langle u_{i-1}(\tilde{t})\rangle+\langle m_{i-1}(\tilde{t})\rangle}{N}\right)\\ &+ p_\mathrm{p}\langle u_i(\tilde{t})\rangle\left(1-\frac{\langle u_i(\tilde{t})\rangle+\langle m_i(\tilde{t})\rangle}{N}\right). \end{aligned} \tag{2.4}$$

Rearranging Equation (2.4) and dividing both sides by τ yields:

$$\begin{aligned}\frac{\langle u_i(\tilde{t}+\tau)\rangle-\langle u_i(\tilde{t})\rangle}{\tau} = \frac{p_\mathrm{m}\Delta^2}{2\tau}&\left[\frac{\langle u_{i-1}(\tilde{t})\rangle-2\langle u_i(\tilde{t})\rangle+\langle u_{i+1}(\tilde{t})\rangle}{\Delta^2}\right]\\ &+\frac{p_\mathrm{m}\Delta^2}{2\tau N}\left[\frac{\langle u_i(\tilde{t})\rangle(\langle m_{i-1}(\tilde{t})\rangle-2\langle m_i(\tilde{t})\rangle+\langle m_{i+1}(\tilde{t})\rangle)}{\Delta^2}\right]\\ &-\frac{p_\mathrm{m}\Delta^2}{2\tau N}\left[\frac{\langle m_i(\tilde{t})\rangle(\langle u_{i-1}(\tilde{t})\rangle-2\langle u_i(\tilde{t})\rangle+\langle u_{i+1}(\tilde{t})\rangle)}{\Delta^2}\right]\\ &+\frac{p_\mathrm{p}}{\tau}\langle u_i(\tilde{t})\rangle\left(1-\frac{\langle u_i(\tilde{t})\rangle+\langle m_i(\tilde{t})\rangle}{N}\right). \end{aligned} \tag{2.5}$$

We now divide both sides of Equation (2.5) by length scale Δ, perform a Taylor expansion and take limits as $\Delta,\tau\to 0$ to obtain a description of the cell density dynamics in terms of the variables $\tilde{u}(\tilde{x},\tilde{t})$ and $\tilde{m}(\tilde{x},\tilde{t})$, that are the continuum counterparts of $\langle u_i(\tilde{t})\rangle/\Delta$ and $\langle m_i(\tilde{t})\rangle/(\tilde{\mu}\Delta)$ that represent, respectively, the number density of cells and the density of ECM at position $\tilde{x}\in\mathbb{R}$ and time $\tilde{t}\in(0,\infty)$. The factor $\tilde{\mu}$ can be interpreted as the number of cells equivalent to a unit mass of ECM. It serves as a conversion factor between the density of ECM—defined as mass of ECM per unit volume—and the number density of ECM elements, which is given by $\tilde{\mu}\tilde{m}(\tilde{x},\tilde{t})$. Under the assumptions

$$\lim_{\Delta,\tau\to 0}\frac{p_\mathrm{m}\Delta^2}{2\tau}=\tilde{D},\qquad \lim_{\tau\to 0}\frac{p_\mathrm{p}}{\tau}=\tilde{r},\qquad \lim_{\Delta\to 0}\frac{N}{\Delta}=\tilde{K}, \tag{2.6}$$

we obtain the following PDE for the cell density $\tilde{u}(\tilde{x},\tilde{t})$:

$$\frac{\partial\tilde{u}}{\partial\tilde{t}}=\tilde{D}\frac{\partial}{\partial\tilde{x}}\left[\left(1-\frac{\tilde{u}+\tilde{\mu}\tilde{m}}{\tilde{K}}\right)\frac{\partial\tilde{u}}{\partial\tilde{x}}+\tilde{u}\frac{\partial}{\partial\tilde{x}}\left(\frac{\tilde{u}+\tilde{\mu}\tilde{m}}{\tilde{K}}\right)\right]+\tilde{r}\tilde{u}\left(1-\frac{\tilde{u}+\tilde{\mu}\tilde{m}}{\tilde{K}}\right), \tag{2.7}$$

where $\tilde{x} \in \mathbb{R}$ and $\tilde{t} \in (0, \infty)$. Note that the first term on the right-hand side of Equation (2.7) describes the movement of cells down gradients in cell density, with movement prevented by the presence of surrounding cells and ECM, as expected by the introduction of volume-filling effects. The second term models the motion of cells down the "total density gradient" of cells and ECM, $\tilde{u} + \tilde{\mu}\tilde{m}$. The third term captures cell proliferation, which is also impacted by volume-filling effects. From Equation (2.7) it is clear that the parameter $\tilde{D} \geq 0$, which is defined via Equation (2.6), can be regarded as the diffusion coefficient of the cells in the absence of ECM, while the parameters $\tilde{r} \geq 0$ and $\tilde{K} > 0$, which are also defined via Equation (2.6), are the intrinsic growth rate of the cell population, and the density corresponding to the maximum occupancy level (*i.e.*, the carrying capacity), respectively.

Coarse-grained model of ECM dynamics. Probabilistic approximations similar to those underlying Equation (2.4) give the following conservation equation for the evolution of ECM elements in lattice site i during the time interval $[\tilde{t}, \tilde{t} + \tau)$:

$$\langle m_i(\tilde{t} + \tau) \rangle = \langle m_i(\tilde{t}) \rangle - p_{\mathrm{d}} \langle u_i(\tilde{t}) \rangle \langle m_i(\tilde{t}) \rangle. \tag{2.8}$$

Rearranging Equation (2.8), dividing by Δ and τ and taking limits as $\Delta, \tau \to 0$, under the assumption

$$\lim_{\Delta, \tau \to 0} \frac{p_{\mathrm{d}} \Delta}{\tau} = \tilde{\lambda}, \tag{2.9}$$

we formally obtain the following differential equation for ECM density $\tilde{m}(\tilde{x}, \tilde{t})$:

$$\frac{\partial \tilde{m}}{\partial \tilde{t}} = -\tilde{\lambda} \tilde{m} \tilde{u}, \tag{2.10}$$

where $\tilde{x} \in \mathbb{R}$ and $\tilde{t} \in (0, \infty)$. Here, the parameter $\tilde{\lambda} \geq 0$ defined via Equation (2.9) is the per cell degradation rate of ECM.

2.2.3 Non-dimensional coarse-grained model

To reduce the number of parameters in the model and to facilitate the following analysis, we non-dimensionalise the mathematical model defined via Equations (2.7) and (2.10) by the introduction of the following non-dimensional variables:

$$u = \frac{\tilde{u}}{\tilde{K}}, \quad m = \frac{\tilde{\mu}\tilde{m}}{\tilde{K}}, \quad t = \tilde{t}\tilde{r}, \quad x = \sqrt{\frac{\tilde{r}}{\tilde{D}}}\tilde{x}.$$

We obtain the following system of equations:

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left[(1-m)\frac{\partial u}{\partial x} + u\frac{\partial m}{\partial x}\right] + u(1-u-m), \tag{2.11}$$

$$\frac{\partial m}{\partial t} = -\lambda m u, \tag{2.12}$$

where $x \in \mathbb{R}$ and $t \in (0, \infty)$. Here, the only remaining parameter is $\lambda = \tilde{\lambda}\tilde{K}/\tilde{r} \geq 0$, which is interpreted as the rescaled ECM degradation rate. We complement the model defined via Equations (2.11) and (2.12) with no flux boundary conditions for Equation (2.11):

$$(1-m)\frac{\partial u}{\partial x} + u\frac{\partial m}{\partial x} = 0, \tag{2.13}$$

at the ends of the domain $x \in \mathbb{R}$ and $u \to 0$ as $x \to \infty$. We also implement the following initial conditions:

$$u(x,0) = u_0(x) \geq 0, \quad m(x,0) = m_0(x) \geq 0, \quad 0 \leq u_0(x) + m_0(x) \leq 1 \quad \forall\, x \in \mathbb{R}. \tag{2.14}$$

We note that by assuming at the single-cell level that both the presence of cells and ECM elements impair the movement and proliferation of the cells, the resulting population-level description for cell density evolution in Equation (2.11) exhibits a number of differences to similar models without volume-filling effects that currently exist in the literature and are summarised in Table 2.1.

Model	Volume-filling in movement		Diffusion term	Volume-filling in proliferation		Reaction term
	by cells	by ECM		by cells	by ECM	
Colson [69]	-	+	$\frac{\partial}{\partial x}\left[(1-m)\frac{\partial u}{\partial x}\right]$	+	-	$u(1-u)$
Browning [43, 100]	-	+	$\frac{\partial}{\partial x}\left[(1-m)\frac{\partial u}{\partial x}\right]$	+	+	$u(1-u-m)$
Equations (2.11) and (2.12)	+	+	$\frac{\partial}{\partial x}\left[(1-m)\frac{\partial u}{\partial x} + u\frac{\partial m}{\partial x}\right]$	+	+	$u(1-u-m)$

Table 2.1: Description of the volume-filling effects of cells and ECM considered by the models compared in this chapter.

For example, the model developed by Browning et al. [43] and studied by El Hachem et al. in [100] does not consider volume-filling of cells to impair cell movement, and therefore contains one less flux term, namely that accounting for movement of cells down the "total density gradient". This model can be recovered from Equation (2.11) by employing different underlying assumptions such that the probability of movement depends on the average available space (where space is only filled by cells)

between the target lattice site and the lattice site the cell occupies at time $\tilde{t}$. The results of the models summarised in Table 2.1 are compared in detail in Section 2.6.

Model simplification when there is no ECM degradation. We observe that when there is no ECM degradation (*i.e.*, if $\lambda = 0$) and ECM is uniformly distributed at $t = 0$ (*i.e.*, if $m(x, 0) \equiv m^0$ where $m^0 \in \mathbb{R}^+$ with $0 \leq m^0 \leq 1$), the mathematical model defined via Equations (2.11) and (2.12) simplifies to the following Fisher–KPP model of cell dynamics [109]:

$$\frac{\partial u}{\partial t} = (1 - m^0)\frac{\partial^2 u}{\partial x^2} + u\big(1 - m^0 - u\big). \tag{2.15}$$

2.3 Numerical methods

In later sections, we will show that the PDE model defined in Equations (2.11) and (2.12) gives rise to travelling wave behaviours which can be analytically studied using the corresponding ODE framework. First however, we detail the numerical methods used to simulate the solutions to the before mentioned ODE and PDE equations studied in this chapter.

2.3.1 The ODE model

We numerically solve the system of ODEs (2.29)–(2.31) subject to the boundary conditions (2.21)–(2.23) using Python's built-in stiff solver `scipy.integrate.ODE` with tolerance 10^{-15} and order 5. Convergence checks were completed by considering a range of tolerances and time steps, ensuring that the parameters used for the simulations produced solutions within known numerical error. In Figure 2.1, we numerically solve Equations (2.29)–(2.31) with the initial condition $(U, V, M) = (0.9, -0.01, 0.01)$ for $c = \{1,\, 2(1 - m_0),\, 3\}$, time step $\tau = 0.01$, and final time $t = 100$.

2.3.2 The PDE model

Although the analytical travelling wave solution is defined on the real line, it is necessary that numerical simulations are performed on a finite domain, due to computational restrictions. As such, we simulate Equations (2.11) and (2.12) on a one-dimensional spatial domain $x \in [0, L]$, where $L > 0$ is chosen to be sufficiently large to ensure that the travelling wave remains well separated from the domain boundaries over the duration of the simulation. Under this assumption, boundary effects are negligible and the finite-domain problem provides an accurate approximation of

the infinite-domain model. Accordingly, Equations (2.11) and (2.12) are solved numerically subject to zero flux boundary conditions (2.13) in u at $x = 0$ and $x = L$ using the method of lines. The spatial domain is uniformly discretised with spacing $\Delta = 0.1$ between each of the $q = 1, \ldots, Q$ spatial points, and the following explicit central difference scheme is used to discretise and approximate the nonlinear diffusion terms [321]:

$$\frac{\partial}{\partial x}\left[D\frac{\partial a}{\partial x}\right]_q \approx \frac{1}{2\Delta^2}\Big[(D_{q-1}+D_q)a_{q-1} - (D_{q-1}+2D_q+D_{q+1})a_q + (D_q+D_{q+1})a_{q+1}\Big], \tag{2.16}$$

where a_q and D_q represent the value of a and D evaluated at the spatial point q, respectively. Simulating the model described by Equations (2.11) and (2.12) requires using the discretisation scheme in Equation (2.16) twice. For the first term with a derivative in x, we use $D = (1 - m)$ and $a = u$, and for the second term in the flux $D = u$ and $a = m$ is employed.

This discretisation of Equations (2.11) and (2.12) results in a system of $2I$ time-dependent ODEs, given by:

$$\frac{\mathrm{d}u_q}{\mathrm{d}t} = \frac{1}{\Delta^2}\Big[u_{q-1}(1-m_q) + u_q(m_{q+1}+m_{q-1}-2) + u_{q+1}(1-m_q)\Big] + u_q(1-u_q-m_q), \tag{2.17}$$

$$\frac{\mathrm{d}m_q}{\mathrm{d}t} = -\lambda m_q u_q, \tag{2.18}$$

for $1 \leq q \leq Q - 1$. To implement Neumann boundary conditions, we introduce the ghost points x_{-1} and x_{Q+1} [221] and enforce zero flux across the boundary. We solve the system of Equations (2.17) and (2.18) with zero flux boundary conditions using the built-in Python solver `scipy.integrate.solve_ivp` with the explicit Runge-Kutta integration method of order 5 and time step $\tau = 1$. Convergence checks were completed by considering a range of tolerances, time and spatial steps.

For the simulations of the PDE systems in this work, we consider compactly supported initial conditions (2.38) and (2.39) (see Section 2.4.2) with $\psi = 1$. In Section 2.4.2.1 we use $\xi = 10^{-7}$, $\gamma = 0.1$ and $\epsilon = 1$ when considering compactly supported initial conditions (2.40) and (2.41), and $a = 0.1$, $\gamma = 0.1$ and $\beta = 10$ for non-compactly supported initial conditions (2.40) and (2.42).

2.4 Travelling wave analysis

We seek constant speed, constant profile travelling wave solutions of Equations (2.11) and (2.12) by adopting the usual travelling wave co-ordinate $z = x - ct$. As such,

we introduce the new variables $u(x,t) = U(z)$ and $m(x,t) = M(z)$ and require that $c > 0$ is a constant so that the wave invades from left to right in the spatial domain. By substituting the ansatz $u(x,t) = U(z)$ and $m(x,t) = M(z)$ in Equations (2.11) and (2.12), we deduce that, if a travelling wave solution exists, then it must satisfy

$$\frac{\mathrm{d}}{\mathrm{d}z}\left[(1-M)\frac{\mathrm{d}U}{\mathrm{d}z} + U\frac{\mathrm{d}M}{\mathrm{d}z}\right] + c\frac{\mathrm{d}U}{\mathrm{d}z} + U\big(1-U-M\big) = 0, \tag{2.19}$$

$$c\frac{\mathrm{d}M}{\mathrm{d}z} - \lambda MU = 0, \tag{2.20}$$

for $-\infty < z < \infty$. We also choose to implement the following boundary conditions, that simulate biologically realistic scenarios where cells start at one end of the domain exclusively and there is only ECM ahead of them:

$$U(z) \to 1 \quad \text{as} \quad z \to -\infty, \tag{2.21}$$

$$U(z) \to 0 \quad \text{as} \quad z \to \infty, \tag{2.22}$$

$$M(z) \to m_0 \quad \text{as} \quad z \to \infty. \tag{2.23}$$

The ODE system (2.19) and (2.20) can be simplified further by expanding Equation (2.19) and using Equation (2.20) to substitute in $\mathrm{d}^2M/\mathrm{d}z^2$. We find that

$$(1-M)\,\frac{\mathrm{d}^2U}{\mathrm{d}z^2} + c\frac{\mathrm{d}U}{\mathrm{d}z} + U\big[\,(1-M)-U\big] = -M\,\left[U\left(\frac{\lambda\,U}{c}\right)^2 + \frac{\mathrm{d}U}{\mathrm{d}z}\left(\frac{\lambda\,U}{c}\right)\right]. \tag{2.24}$$

Furthermore, it is notable that Equation (2.20), subject to the boundary condition (2.23), has a semi-explicit solution. That is, if $U(z)$ is known, then we can evaluate $M(z)$ as

$$M(z) = m_0\,\exp\left\{-\frac{\lambda}{c}\int_z^\infty U(s)\mathrm{d}s\right\}, \tag{2.25}$$

which gives

$$M(z) \to 0 \quad \text{as} \quad z \to -\infty, \tag{2.26}$$

and $M \leq m_0$ for all $z \in \mathbb{R}$.

Under the boundary condition $U(z) \to 0$ as $z \to \infty$ at the leading edge of the travelling front (*i.e.*, for $z \in (\ell, \infty)$ with $-\infty \ll \ell < \infty$ sufficiently large), we assume that there is an exponential tail and, as such, we use the ansatz

$$U(z) \approx \exp\{-\alpha z\}, \tag{2.27}$$

with scaling constant $0 < \alpha < \infty$ for $z \in (\ell, \infty)$. Inserting Equation (2.27) into Equation (2.25) we find

$$M(z) \approx m_0 \exp\left\{-\frac{1}{\alpha}\,\left(\frac{\lambda\,U(z)}{c}\right)\right\}, \tag{2.28}$$

for $z \in (\ell, \infty)$, which will become useful in analysing the travelling waves in asymptotic regimes (see Sections 2.5.1 and 2.5.2).

Moreover, writing $\mathrm{d}U/\mathrm{d}z = V$, we can rewrite Equations (2.19) and (2.20) as a system of three first-order ODEs

$$\frac{\mathrm{d}U}{\mathrm{d}z} = V, \tag{2.29}$$

$$\frac{\mathrm{d}V}{\mathrm{d}z} = \frac{1}{(1-M)}\bigg[- cV - \frac{\lambda}{c}MUV - \frac{\lambda^2}{c^2}MU^3 - U(1-U-M)\bigg], \tag{2.30}$$

$$\frac{\mathrm{d}M}{\mathrm{d}z} = \frac{\lambda}{c}MU, \tag{2.31}$$

with boundary conditions given by

$$U(z) \to 1, \quad V(z) \to 0 \quad \text{and} \quad M(z) \to 0 \quad \text{as} \quad z \to -\infty, \tag{2.32}$$

$$U(z) \to 0, \quad V(z) \to 0 \quad \text{and} \quad M(z) \to m_0 \quad \text{as} \quad z \to \infty. \tag{2.33}$$

2.4.1 Linear stability analysis

The system (2.29)–(2.31) with boundary conditions (2.32) and (2.33) has two equilibrium points $\mathcal{S}_1 = (1, 0, 0)$ and $\mathcal{S}_2 = (0, 0, m_0)$. We seek travelling wave solutions, which are trajectories in the phase space that connect the spatially homogeneous steady state $\mathcal{S}_1$ at $z = -\infty$ to $\mathcal{S}_2$ at $z = \infty$ [82, 183, 225].

We first find the Jacobian of the linearised system (2.29)–(2.31). To do this, we introduce the following combinations for simplicity

$$\nu = \frac{\lambda}{c}, \qquad T = \frac{1}{1-M}, \qquad W = MU,$$

so that the Jacobian is given by

$$\mathbf{J} = \begin{pmatrix} 0 & 1 & 0 \\ T(M - 1 + 2U - 3\nu^2 WU - \nu W) & -T\big(c + \nu W\big) & Y \\ \nu M & 0 & \nu U \end{pmatrix},$$

where $Y = UT(1 - \nu^2 U^2 - \nu V + T(U + M - 1 - \nu^2 WU)) - T^2(cV + \nu)$. Then the Jacobian at $\mathcal{S}_2 = (0, 0, m_0)$ is given by

$$\mathbf{J}_{(0,0,m_0)} = \begin{pmatrix} 0 & 1 & 0 \\ \dfrac{m_0 - 1}{1 - m_0} & \dfrac{-c}{1 - m_0} & 0 \\ \dfrac{\lambda}{c} m_0 & 0 & 0 \end{pmatrix},$$

and the Jacobian at $\mathcal{S}_1 = (1, 0, 0)$ is

$$\mathbf{J}_{(1,0,0)} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -c & 1 - \left(\frac{\lambda}{c}\right)^2 \\ 0 & 0 & \frac{\lambda}{c} \end{pmatrix}.$$

For $(1, 0, 0)$, the eigenvalues are:

$$\sigma_1 = \lambda/c, \quad \sigma_{2,3} = (-c \pm \sqrt{c^2 + 4})/2,$$

which have associated eigenvectors

$$\mathbf{v_1} = \left(\frac{c^2 - \lambda^2}{c^2(\lambda - 1) + \lambda^2}, \quad \frac{\lambda(c^2 - \lambda^2)}{c(c^2(\lambda - 1) + \lambda^2)}, \quad 1 \right)^T, \tag{2.34}$$

$$\mathbf{v_{2,3}} = \left(\frac{c \pm \sqrt{c^2 + 4}}{2}, \quad 0, \quad 1 \right)^T. \tag{2.35}$$

These indicate that $(1, 0, 0)$ is a three-dimensional, hyperbolic, unstable saddle point since it has one negative and two positive eigenvalues [14]. However, at $(0, 0, m_0)$ the eigenvalues of the system are

$$\sigma_1 = 0, \quad \sigma_{2,3} = \frac{-c \pm \sqrt{c^2 - 4(1 - m_0)^2}}{2(1 - m_0)},$$

showing that $(0, 0, m_0)$ is a non-hyperbolic, stable steady state, since one of these eigenvalues has zero real part [359]. The corresponding eigenvectors at $\mathcal{S}_2$ are

$$\mathbf{w_1} = \begin{pmatrix} 0, & 0, & 1 \end{pmatrix}^T, \tag{2.36}$$

$$\mathbf{w_{2,3}} = \left(\frac{c(c \pm \sqrt{c^2 - 4(1 - m_0)^2})}{2\lambda m_0(m_0 - 1)}, \quad \frac{c(c^2 \pm c\sqrt{c^2 - 4(1 - m_0)^2} - 2(1 - m_0)^2)}{2\lambda m_0(1 - m_0)^2}, \quad 1 \right)^T. \tag{2.37}$$

In all cases, we use the index 2 to refer to the positive of the two choices, and 3 for the negative. When $c^2 - 4(1 - m_0)^2 < 0$, the steady state $(0, 0, m_0)$ is a stable spiral as the eigenvalues have non-zero imaginary parts; however, when $c^2 - 4(1 - m_0)^2 > 0$, the steady state is a stable node. Since these equilibria's stability can only be determined for $m_0 \neq 1$, we continue our analysis assuming $m_0 \in [0, 1)$. In the case that the state $(0, 0, m_0)$ is a stable spiral, U oscillates around this point on its approach and can therefore take negative values, see Figure 2.1. However, when $(0, 0, m_0)$ is a stable node, trajectories may exist between $(1, 0, 0)$ and $(0, 0, m_0)$ that are contained entirely

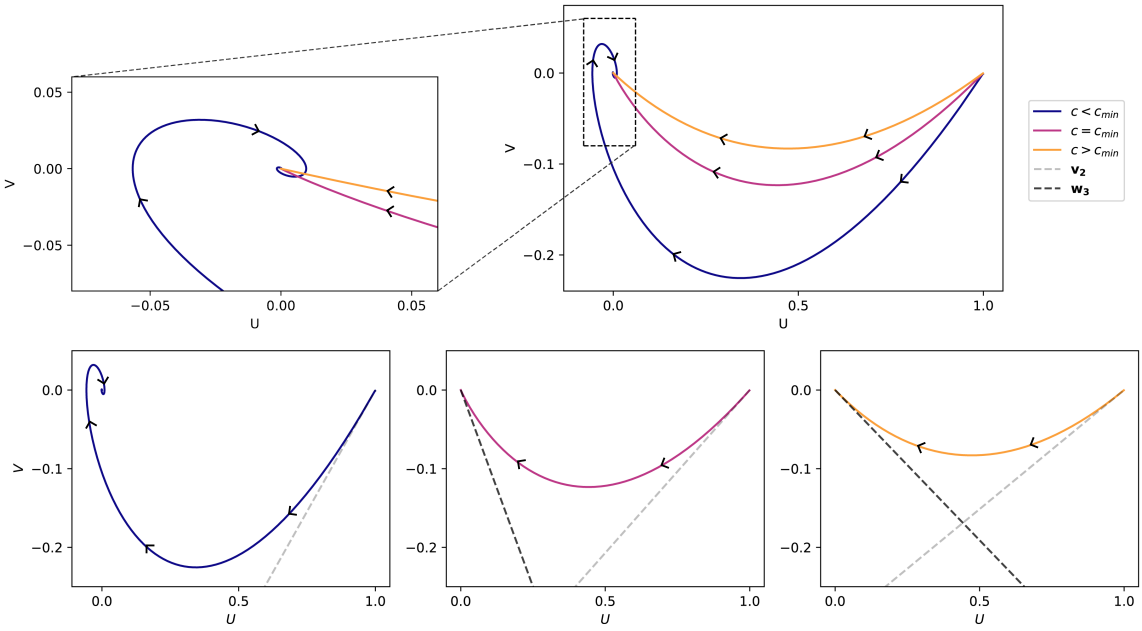

Figure 2.1: Phase plane plot of the ODE system (2.29)–(2.31), for different travelling wave speeds, c, demonstrating the change from a stable spiral to a stable node as the travelling wave speed exceeds $c_{\min}$. The corresponding unstable eigenvector (light grey dashed) given by Equation (2.35) and stable eigenvector (dark grey dashed) given by Equation (2.37) are overlaid in the lower plots [82]. Further specifics of the parameter values and the numerical methods used can be found in Section 2.3.

in the region of phase space defined by $U \geq 0$, $V \leq 0$ and $M \geq 0$. This ensures non-negativity of U and M, consistent with the PDE (which preserves non-negativity) and the biological grounding of the model. This demonstrates the existence of a minimum wave speed, $c_{\min} = 2(1 - m_0)$, such that the dependent variables, U and M, remain non-negative for all time.

2.4.2 Travelling wave solutions

From numerical simulations, we assume that the system (2.11) and (2.12) permits travelling wave solutions connecting $\mathcal{S}_1 = (1, 0, 0)$ and $\mathcal{S}_2 = (0, 0, m_0)$ for $m_0 \in [0, 1)$. To explore this, we simulate Equations (2.11) and (2.12) subject to the following initial conditions:

$$u(x,0) = \begin{cases} 1, & \text{if} \quad x < \psi, \\ 0 & \text{if} \quad x \geq \psi, \end{cases} \tag{2.38}$$

$$m(x,0) = \begin{cases} 0, & \text{if} \quad x < \psi, \\ m_0 & \text{if} \quad x \geq \psi, \end{cases} \tag{2.39}$$

where $0 < \psi \ll L$ represents the width of the initially invaded region at $t = 0$ and $m_0 \in [0, 1]$ corresponds to the uninvaded density of ECM ahead of the cells.

Figure 2.2 shows that solutions of Equations (2.11) and (2.12) subject to the initial conditions (2.38) and (2.39) converge to travelling wave profiles with $c > 0$ whereby the cell density, u, decreases monotonically from one to zero and the ECM density, m, increases monotonically from zero to m_0. The numerical results in Figure 2.2 also indicate that the speed of the travelling waves changes as the values of the parameters λ and m_0 are changed. This is illustrated in more detail in Figure 2.3, that shows that when $m_0 \in (0, 1)$: if $\lambda \to 0^+$ then the speed of the travelling waves converges to $c = 2\,(1 - m_0)$; whereas if $\lambda \to \infty$ then the speed of the travelling waves converges to $c = 2$.

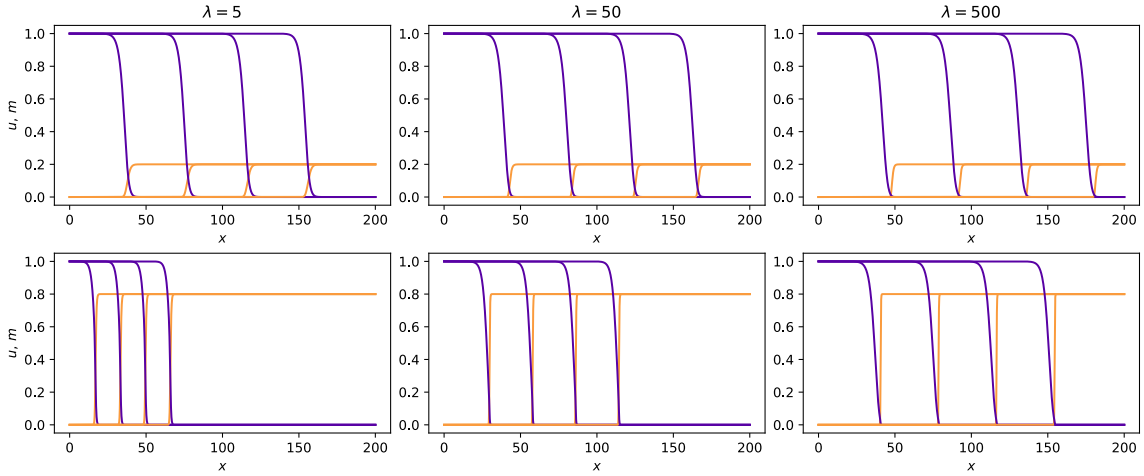

Figure 2.2: Numerical solutions of Equations (2.11) and (2.12) subject to the initial conditions (2.38) and (2.39), for $m_0 = 0.2$ in the top row and $m_0 = 0.8$ in the bottom row, and for rescaled ECM degradation rates $\lambda = 5$, 50, 500. Cell densities are shown in purple and ECM densities in orange at times $t = 25$, 50, 75, 100 from left to right. Further specifics of the parameter values and the numerical methods used can be found in Section 2.3.

The numerical results summarised by Figure 2.3 for $m_0 \in [0, 1)$ show similar behaviours to that in [100], where volume-filling effects of cells are not included in the cell movement term. However, a marked difference is observed for the case $m_0 = 1$, as discussed in the following section.

2.4.2.1 Numerical simulations for $m_0 = 1$

By examining Figure 2.3 it is clear that when $m_0 = 1$, the system (2.11) and (2.12) subject to the initial conditions (2.38) and (2.39) does not permit travelling wave

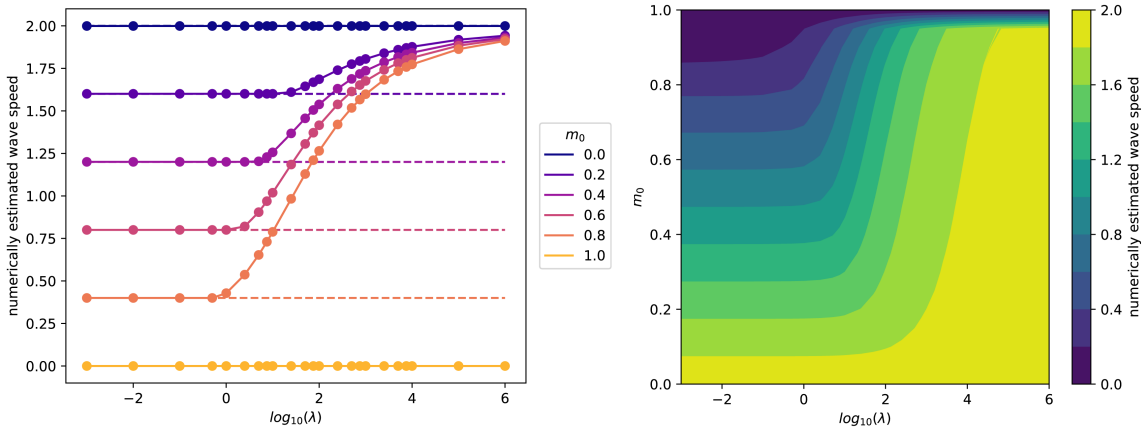

Figure 2.3: The relationship between the numerically estimated speed (solid lines) of travelling wave solutions of Equations (2.11) and (2.12) subject to the initial conditions (2.38) and (2.39) for a range of values of λ. The dashed lines in the plot on the left highlight the value of $2(1-m_0)$. The numerically estimated travelling wave speed is obtained by tracing the point $X(t)$ such that $u(X(t),t) = 0.1$. Further specifics of the parameter values and the numerical methods used can be found in Section 2.3.

solutions. To investigate this further, we simulate the system (2.11) and (2.12) subject to different initial conditions. In every case, we consider the initial condition for the ECM density, m, given by

$$m(x,0) = \begin{cases} m_0 - u(x,0), & \text{if} \quad m_0 > 1-\gamma, \\ m_0, & \text{if} \quad m_0 \leq 1-\gamma, \end{cases} \tag{2.40}$$

with $0 \leq \gamma \leq 1$. When $\gamma = 0$, there are no cells initially, so we do not consider this case going forward.

To explore the behaviours observed at $m_0 = 1$, we consider two different options for $u(x,0)$. First, we consider the compactly supported initial condition

$$u(x,0) = \begin{cases} \gamma\big(1-\tanh(\frac{x}{\epsilon})\big), & \text{if} \quad \gamma\big(1-\tanh(\frac{x}{\epsilon})\big) \geq \xi, \\ 0, & \text{if} \quad \gamma\big(1-\tanh(\frac{x}{\epsilon})\big) < \xi, \end{cases} \tag{2.41}$$

which models an initial population of cells at the far left of the domain that decays to an entirely unpopulated region of cells towards the right of the domain. Here, $\gamma \in [0,1]$ represents the maximum cell density at $t = 0$ and $m_0 \in [0,1]$ corresponds to the uninvaded density of ECM. Moreover, in the definition given by Equation (2.41), the parameter $\xi \in (0,1]$ is used to control the tolerance below which the cell density can be assumed, on a first approximation, to be zero, and $\epsilon > 0$ represents the initial width of the cell density profile. This is similar in structure to the initial conditions considered earlier (Equation (2.38)), but now instead of a fully occupied region of cells

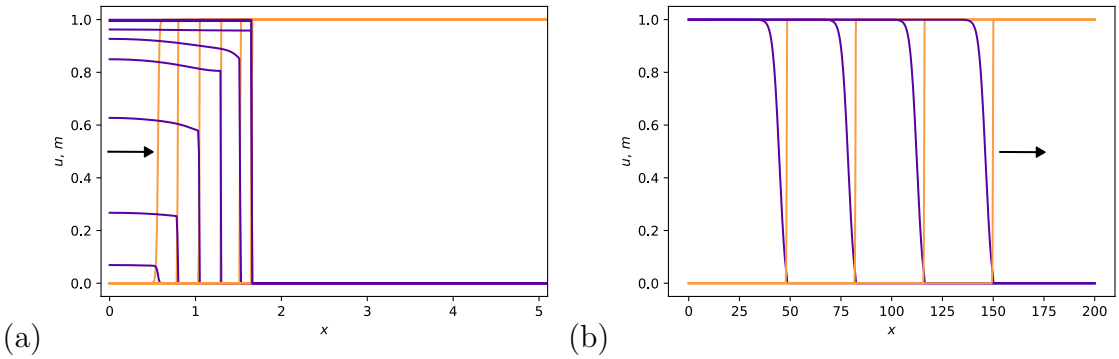

Figure 2.4: Numerical solutions to the system of Equations (2.11) and (2.12) subject to the initial conditions (2.40) and (2.41) (panel (a)) or (2.40) and (2.42) (panel (b)), for $m_0 = 1$ and $\lambda = 250$. Cell densities are shown in purple and ECM densities in orange at times $t = 2, 4, 6, 8, 10, 12, 14, 16$ (from left to right) in panel (a) and times $t = 25, 50, 75, 100$ (from left to right) in panel (b). Note that the axis in the plot in panel (a) is zoomed in on $x \in [0, 5]$ to display the initial behaviour in the transient region before invasion stops. Further specifics of the parameter values and the numerical methods used for simulation can be found in Section 2.3.

in the back, the cell density smoothly decays to zero, leaving some available space in the domain initially, allowing room for cell growth. As an alternative, we also consider the following non-compactly supported initial condition for $u(x, 0)$ (along with $m(x, 0)$ defined by Equation (2.40)), as used in [100],

$$u(x,0) = \begin{cases} \gamma, & x < \beta, \\ \gamma \exp\{-a(x-\beta)\}, & x \geq \beta, \end{cases} \tag{2.42}$$

where $\beta \in \mathbb{R}$ is used to define a region where the cell density is initially constant and equal to $\gamma \in [0, 1]$, while the parameter $a > 0$ is used to prescribe the length scale over which the cell density profile decays. We note that, since γ, $m_0 \in [0, 1]$, the initial conditions (2.40)–(2.42) are such that the total density of cells and ECM at $t = 0$ does not locally exceed the extreme value 1, which corresponds to complete local saturation, *i.e.*, $u(x, 0) + m(x, 0) \leq 1$ for all $x \in [0, L]$. We also note that when $m_0 = 0$ the initial condition (2.40) reduces to the trivial initial condition $m(x, 0) \equiv 0$, and that if $\gamma = 0$, then there are no cells to consider invading.

The numerical results in Figure 2.4(a), which complement the results summarised in Figure 2.3, show that when $m_0 = 1$ the system (2.11) and (2.12) subject to the initial conditions with compactly supported cell density (2.41) cannot sustain travelling wave solutions. In fact, invasion is entirely prevented beyond a point x^*, that is the smallest x such that $u(x, 0) = 0$ for all $x \geq x^*$. On the other hand, the numerical results in Figure 2.4(b) demonstrate that travelling wave solutions can be

sustained in the case where non-compactly supported initial conditions (2.42) in u are considered.

This result is a direct consequence of the volume-filling effects incorporated in the model. By considering an initial condition where $m(x, 0) = 1$ ahead of the invading population, the invading cells cannot penetrate the region where $u = 0, m = 1$. This is consistent with the underlying individual-based description, since cells can only degrade ECM in the same lattice site as themselves. Indeed, by design, the model (2.11) and (2.12) does not permit travelling waves under initial conditions with compactly supported cell density and $m_0 = 1$, as cells require available space ahead of the wave for invasion. Thus, in regions initially devoid of cells, the ECM cannot be degraded, preventing cell invasion. We acknowledge that this is a limitation of the model proposed in this work, as we would expect that, at the very least, cells could degrade ECM within their contact range, and this limit will be discussed in more detail in Section 2.7. However, for now, we proceed by considering $m_0 \in [0, 1)$.

2.4.2.2 Travelling waves when no ECM is present

When we consider the case of no ECM ahead of the invading population of cells, which is described by $m_0 = 0$, the solutions to Equation (2.12) subject to the initial condition (2.39) are such that $m(x, t) \equiv 0$ for all $t \geq 0$. As such, the model (2.11) and (2.12) simplifies to the Fisher–KPP model (1.1) with $D = r = K = 1$, that is

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + u\,(1 - u)\,. \tag{2.43}$$

Consistent with this, numerical simulations indicate that when $m_0 = 0$, the cell density u converges to a travelling wave that decreases monotonically from one to zero (results not shown for brevity), and travels with speed $c = 2$ (*i.e.*, the minimal speed of travelling wave solutions to the Fisher–KPP model (2.43)), see Figure 2.3.

2.4.2.3 Travelling wave profiles for $\lambda \to 0^+$ and $\lambda \to \infty$

In order to better understand the behaviour of the travelling waves we now consider the analytically tractable limits. For $\lambda \to 0^+$ and $\lambda \to \infty$, we began by simulating solutions to Equations (2.11) and (2.12) subject to the initial conditions (2.38) and (2.39) for very small and very large values of the rescaled ECM degradation rate, λ.

Figure 2.5 shows a similar travelling wave profile for all three rescaled ECM degradation rates considered, which is much steeper than those displayed in Figure 2.2.

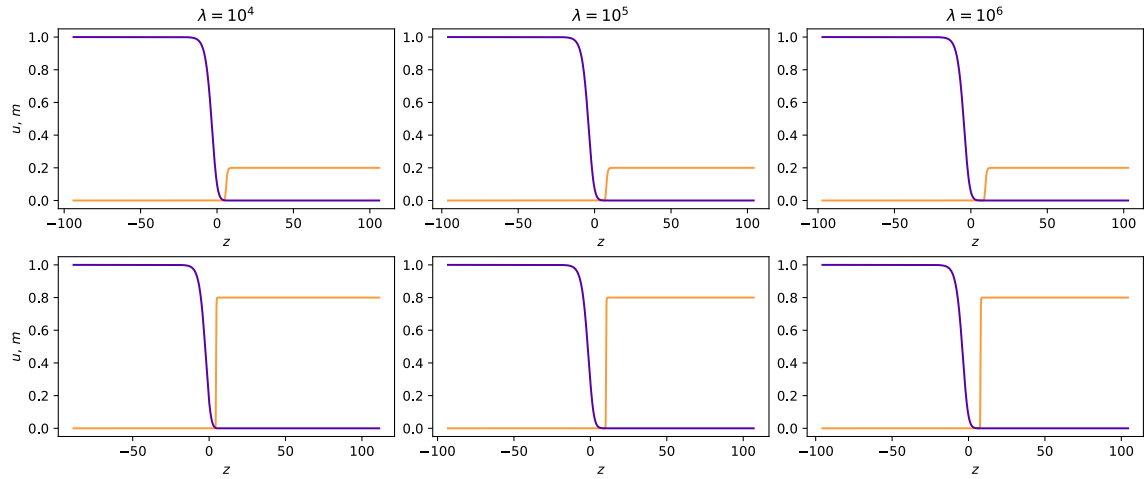

Figure 2.5: Travelling wave solutions of Equations (2.11) and (2.12) subject to the initial conditions (2.38) and (2.39), for $m_0 = 0.2$ in the top row and $m_0 = 0.8$ in the bottom row, and for rescaled ECM degradation rates $\lambda = 10^4$, 10^5, 10^6. Cell densities are shown in purple and ECM densities in orange. Further specifics of the parameter values and the numerical methods used can be found in Section 2.3.

We can also observe that for high λ values, there is no observable overlap between the non-zero regions of the travelling wave in the cells and the ECM. This separation is a result of extremely fast degradation rates allowing even a very small cell density to quickly degrade any ECM in the same spatial location, and thus making room for the cell density profile to evolve. It is also noticeable that, as before, the travelling wave profile of the ECM is sharper when the initial ECM density, m_0, is larger.

Alternatively, by considering very small rescaled ECM degradation rates, we can see in Figure 2.6 that the travelling wave profile becomes much wider, especially for $\lambda = 10^{-3}$, where we observe a distinctive 'corner' in the solution where the cell density decreases rapidly to zero. It also takes a much longer time for the full travelling wave profile to evolve. By examining Figure 2.7, it becomes clear that as λ decreases, the degradation of ECM becomes very slow, thus only allowing room for a small portion of the total available space at the front of the wave to be filled by cells. Behind this wave front, a very shallow travelling wave profile begins to appear.

2.5 Asymptotic analysis of the travelling wave speed

After observing disparities in the analytically predicted and numerically simulated travelling wave speeds of the model given by Equations (2.11) and (2.12) (as shown in Figure 2.3), we are particularly interested in investigating the dependence of travelling

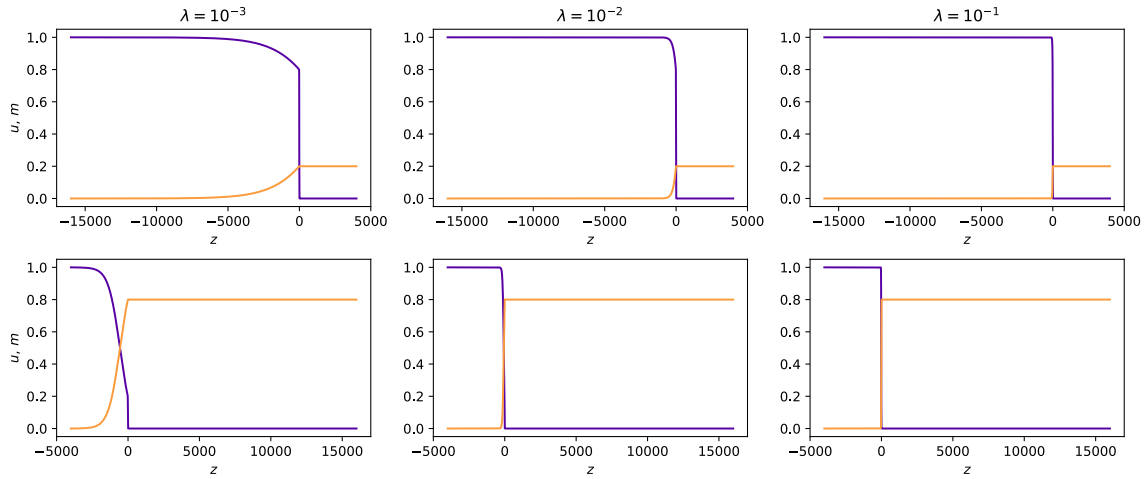

Figure 2.6: Travelling wave solutions of Equations (2.11) and (2.12) subject to the initial conditions (2.38) and (2.39), for $m_0 = 0.2$ in the top row and $m_0 = 0.8$ in the bottom row, and for rescaled ECM degradation rates $\lambda = 10^{-3}, 10^{-2}, 10^{-1}$. Cell densities are shown in purple and ECM densities in orange. Further specifics of the parameter values and the numerical methods used can be found in Section 2.3.

wave solutions on the parameters λ, the rescaled ECM degradation rate, and m_0, the density of ECM far ahead of the wave. Having now clearly determined that the minimum travelling wave speed decreases linearly as m_0 increases, we now aim to explore the relationship between the numerically estimated travelling wave speed and the rescaled ECM degradation rate, λ.

Standard perturbation techniques were not successfully amenable when applied to the travelling wave equations (2.19) and (2.20) in order to unveil how the travelling wave speed depends on λ. As a result, we examine Figure 2.3 for clues as to how to proceed. We immediately see that for sufficiently small λ it appears that the numerically estimated travelling wave speed is independent of λ and matches the speed predicted by standard travelling wave analysis. It can also be seen from the contour plot in Figure 2.3 that for large values of λ, the speed converges to 2 for all values of $m_0 \in [0, 1)$. Therefore, we proceed to investigate the limits corresponding to slow and fast rescaled ECM degradation rates, $\lambda \to 0^+$ and $\lambda \to \infty$, respectively, using asymptotic analysis.

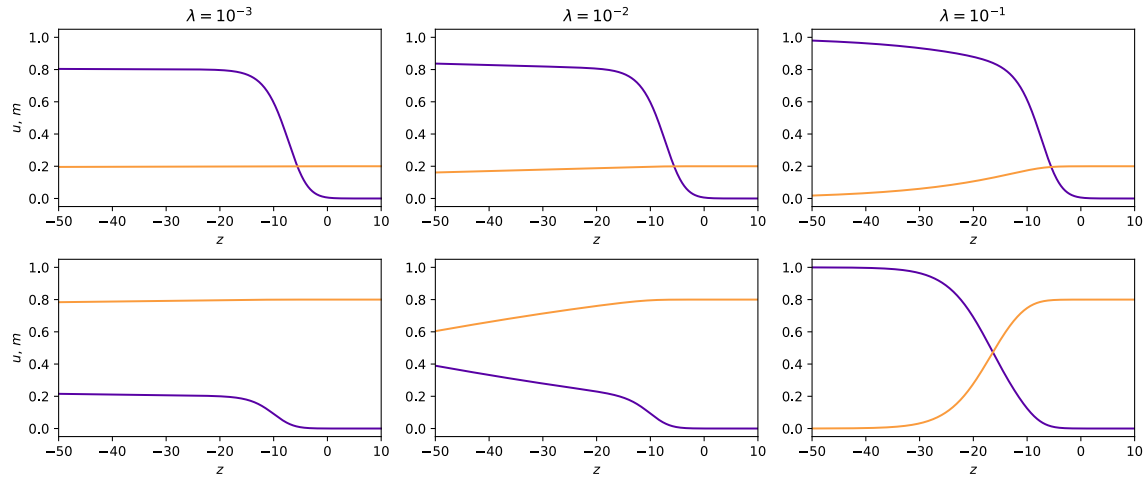

Figure 2.7: Travelling wave solutions of Equations (2.11) and (2.12) subject to the initial conditions (2.38) and (2.39), for $m_0 = 0.2$ in the top row and $m_0 = 0.8$ in the bottom row, and for rescaled ECM degradation rates $\lambda = 10^{-3}, 10^{-2}, 10^{-1}$. Cell densities are shown in purple and ECM densities in orange, zoomed in on the evolved travelling wave front, as shown in Figure 2.6. Further specifics of the parameter values and the numerical methods used can be found in Section 2.3.

2.5.1 Formal asymptotic analysis for $\lambda \to \infty$

In the case of very large rates of ECM degradation, by considering the semi-explicit solution for M in terms of U given by Equation (2.28), we see that

$$M(z) \approx m_0 \exp\left\{-\frac{1}{\alpha}\left(\frac{\lambda U(z)}{c}\right)\right\} \to 0 \quad \text{as} \quad \lambda \to \infty, \tag{2.44}$$

for $z \in (\ell, \infty)$ (see Figure 2.5 for the travelling wave profiles). In the asymptotic regime $\lambda \to \infty$, substituting Equation (2.44) into Equation (2.24) and using the fact that, since $0 \leq U(z) < 1$ for $z \in (\ell, \infty)$ and $\mathrm{d}U(z)/\mathrm{d}z \approx -\alpha\, U(z)$ for $z \in (\ell, \infty)$ (cf. the ansatz given by Equation (2.27)), the following asymptotic relation holds

$$m_0 \exp\left\{-\frac{1}{\alpha}\left(\frac{\lambda U(z)}{c}\right)\right\}\left[U(z)\left(\frac{\lambda U(z)}{c}\right)^2 + \frac{\mathrm{d}U(z)}{\mathrm{d}z}\left(\frac{\lambda U(z)}{c}\right)\right] \to 0 \quad \text{as} \quad \lambda \to \infty, \tag{2.45}$$

for $z \in (\ell, \infty)$. Therefore

$$\frac{\mathrm{d}^2U(z)}{\mathrm{d}z^2} + c\frac{\mathrm{d}U(z)}{\mathrm{d}z} + U(z)\big(1 - U(z)\big) \approx 0, \tag{2.46}$$

for $z \in (\ell, \infty)$. Hence, when $\lambda \to \infty$ we expect $U(z)$ at the leading edge of the travelling front to behave, to a first approximation, as the solution to the Fisher–KPP equation (Equation (2.43)) in travelling wave co-ordinates, for which $c_{\min} = 2$.

This result can also be observed numerically in the plot on the right of Figure 2.8. The same behaviour is observed in similar models without volume-filling effects [69, 100], demonstrating that the model in Equations (2.11) and (2.12) can be approximated, to an extent, with any of these simpler models in the parameter regime $\lambda \to \infty$, as growth and diffusion are unrestricted by the ECM within a neighbourhood of the travelling wave front.

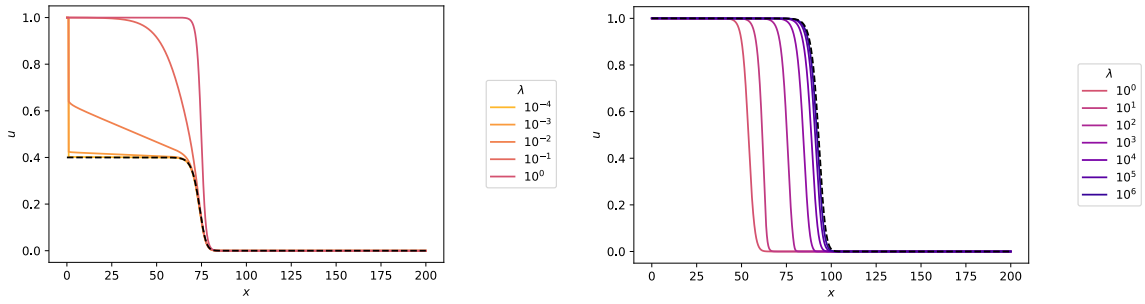

Figure 2.8: Left: plot of the cell density, u, obtained through numerical simulations of Equations (2.11) and (2.12) subject to the initial conditions (2.38) and (2.39) (solid lines) for small values of λ, and numerical simulations of the Fisher–KPP model (using Equation (2.15)) with rescaled coefficients (dashed black line) with $t = 100$ and $m_0 = 0.6$. Right: plot of the cell density, u, obtained through numerical simulations of Equations (2.11) and (2.12) subject to the initial conditions (2.38) and (2.39) (solid lines) for large values of λ, and numerical simulations of the Fisher–KPP model (Equation (2.43)) (dashed black line) in the plot on the right for $t = 50$ and $m_0 = 0.4$. Qualitatively, the same behaviour is observed for all $m_0 \in [0, 1)$. Further specifics of the parameter values and the numerical methods used for the simulations can be found in Section 2.3.

2.5.2 Formal asymptotic analysis for $\lambda \to 0^+$

Using Equation (2.28) it is clear that

$$M(z) \approx m_0 \exp\left\{-\frac{1}{\alpha}\left(\frac{\lambda U(z)}{c}\right)\right\} \to m_0 \quad \text{as} \quad \lambda \to 0^+, \tag{2.47}$$

for $z \in (\ell, \infty)$ (see Figures 2.6 and 2.7 for the travelling wave profiles).

In the asymptotic regime $\lambda \to 0^+$, substituting Equation (2.47) into Equation (2.24) and using the fact that, since $0 \leq U(z) < 1$ for $z \in (\ell, \infty)$ and $\mathrm{d}U(z)/\mathrm{d}z \approx -\alpha\, U(z)$ for $z \in (\ell, \infty)$ (cf. the ansatz given by Equation (2.27)), the following asymptotic relation holds

$$m_0 \exp\left\{-\frac{1}{\alpha}\left(\frac{\lambda U(z)}{c}\right)\right\}\left[U(z)\left(\frac{\lambda U(z)}{c}\right)^2 + \frac{\mathrm{d}U(z)}{\mathrm{d}z}\left(\frac{\lambda U(z)}{c}\right)\right] \to 0 \quad \text{as} \quad \lambda \to 0^+, \tag{2.48}$$

for $z \in (\ell, \infty)$. Considering again Equation (2.24) in the limit $\lambda \to 0^+$, then we find

$$(1 - m_0)\frac{\mathrm{d}^2 U(z)}{\mathrm{d}z^2} + c\frac{\mathrm{d}U(z)}{\mathrm{d}z} + U(z)\big[\,(1 - m_0) - U(z)\big] \approx 0, \tag{2.49}$$

for $z \in (\ell, \infty)$. Equation (2.49) is equivalent to the Fisher–KPP model in travelling wave co-ordinates

$$\hat{D}\frac{\mathrm{d}^2 \hat{U}}{\mathrm{d}z} + \hat{c}\frac{\mathrm{d}\hat{U}}{\mathrm{d}z} + \hat{r}\hat{U}\left(1 - \frac{\hat{U}}{\hat{K}}\right) = 0, \tag{2.50}$$

with $\hat{D} = \hat{r} = \hat{K} = 1 - m_0$, with $\hat{c}_{\min} = 2(1 - m_0)$, as predicted earlier in Equation (2.15). An excellent match around the leading edge of the travelling wave front between the travelling wave solution to the Fisher–KPP model (Equation (2.50) in travelling wave co-ordinates) and Equations (2.11) and (2.12) for low values of the rescaled ECM degradation rate can be seen in the plot on the left in Figure 2.8.

We now consider the region $z \in (-\infty, \ell)$ by rescaling Equations (2.19) and (2.20) using the new variable $\epsilon = z\lambda$ for $\epsilon \in (-\infty, \ell\lambda]$, and rewriting $U(z) = U(\epsilon)$ and $M(z) = M(\epsilon)$. The system of Equations (2.19) and (2.20) becomes

$$-c\lambda\frac{\mathrm{d}U}{\mathrm{d}\epsilon} = \lambda^2\frac{\mathrm{d}}{\mathrm{d}\epsilon}\left[(1 - M)\frac{\mathrm{d}U}{\mathrm{d}\epsilon} + U\frac{\mathrm{d}M}{\mathrm{d}\epsilon}\right] + U(1 - U - M), \tag{2.51}$$

$$\lambda\frac{\mathrm{d}M}{\mathrm{d}\epsilon} = \frac{\lambda}{c}MU. \tag{2.52}$$

For $\lambda \to 0^+$, we find from Equation (2.51) that $U(\epsilon)(1 - U(\epsilon) - M(\epsilon)) = 0$, so that for $\epsilon \in (-\infty, \ell\lambda]$ we have $U(\epsilon) = 1 - M(\epsilon)$ since $U(\epsilon) \to 1$ as $\epsilon \to -\infty$. By substitution into Equation (2.52), we find

$$\frac{\mathrm{d}M}{\mathrm{d}\epsilon} = \frac{M(1 - M)}{c},$$

which, using the matching condition that $M(\epsilon = \ell\lambda) = m_0$, gives

$$M(\epsilon) = \frac{m_0 \exp\{-(\lambda\ell - \epsilon)/c\}}{1 - m_0 + m_0 \exp\{-(\lambda\ell - \epsilon)/c\}}. \tag{2.53}$$

Recalling that $U(\epsilon) = 1 - M(\epsilon)$, we obtain

$$U(\epsilon) = \frac{1 - m_0}{1 - m_0 + m_0 \exp\{-(\lambda\ell - \epsilon)/c\}}, \tag{2.54}$$

which tends to 1 as $\epsilon \to -\infty$ and to $1 - m_0$ as $\epsilon \to \ell\lambda$. In the travelling wave co-ordinate, z, Equation (2.54) can be written as

$$U(z) = \frac{1 - m_0}{1 - m_0 + m_0 \exp\{-\lambda(\ell - z)/c\}}, \tag{2.55}$$

for $z \in (-\infty, \ell]$ and the solution to the Fisher–KPP model, as given by Equation (2.49), for $z \in (\ell, \infty)$. In the travelling wave co-ordinate, z, the solution for the wave profile of the ECM, given by Equation (2.53), is

$$M(z) = \frac{m_0 \exp\{-\lambda(\ell - z)/c\}}{1 - m_0 + m_0 \exp\{-\lambda(\ell - z)/c\}}, \tag{2.56}$$

for $z \in (-\infty, \ell]$, and $M(z) = m_0$ for $z \in (\ell, \infty)$, as given by Equation (2.47). An excellent agreement between these analytical solutions and the numerical results can be observed in Figure 2.9.

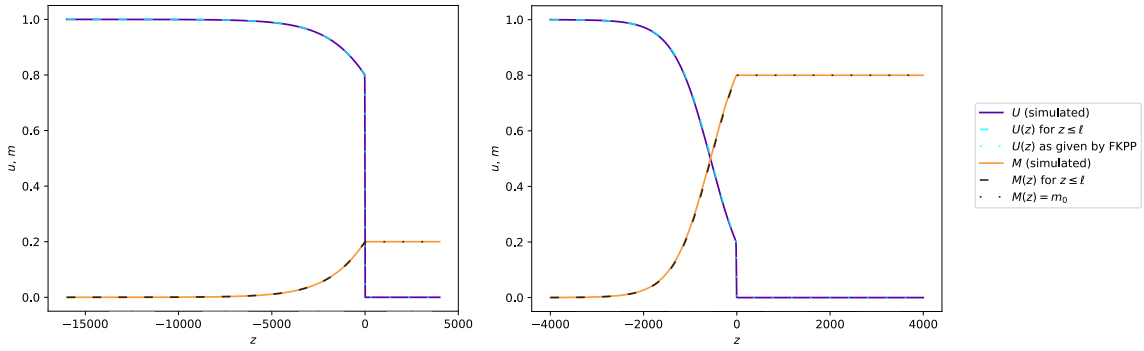

Figure 2.9: Numerical solutions of Equations (2.11) and (2.12) subject to the initial conditions (2.38) and (2.39), for $m_0 = 0.2$ on the left and $m_0 = 0.8$ on the right with rescaled ECM degradation rate $\lambda = 10^{-3}$ translated into the travelling wave co-ordinate, z. Solid lines represent the cell and ECM densities from numerical simulations in purple and orange, respectively. The Fisher–KPP solution (2.49) in travelling wave co-ordinates is plotted as a dotted blue line. The solution $M(z) = m_0$ is plotted in dotted black, and the analytical solutions given by Equations (2.55) and (2.56) are plotted in dashed blue and black lines, respectively. Further specifics of the parameter values and the numerical methods used can be found in Section 2.3.

Similar models, such as those described at the end of Section 2.4 that do not have volume-filling effects taken into account, demonstrate qualitatively similar behaviour. In all of these models, at very low rescaled ECM degradation rates we observe convergence of the solutions to those of the Fisher–KPP model with rescaled parameters. For models with the same cell proliferation term as in Equation (2.11), the rescaled parameters are the same and the convergence has qualitatively similar behaviour, as displayed in the plot on the left in Figure 2.8. As a result, in the limit of very small rescaled ECM degradation rates, $\lambda \to 0^+$, the model (2.11) and (2.12) can be simplified to that presented in [100], which neglects the volume-filling effects of cells upon cell movement. This model can, in turn, be well approximated by the Fisher–KPP model with rescaled parameters $\hat{D} = \hat{r} = \hat{K} = 1 - m_0$. This result is consistent with

predictions from standard travelling wave analysis. However, for the model presented in [69], the parameters of the rescaled Fisher–KPP model to which the model converges are, instead, $\hat{D} = 1 - m_0$ and $\hat{r} = \hat{K} = 1$, which entails a higher cell carrying capacity density since proliferation is not impacted by the surrounding ECM. See Section 2.6 for a more detailed comparison. As such, the model (2.11) and (2.12) is poorly approximated using the before-mentioned model simplifications, such as that in [69], which have different underlying assumptions for cell proliferation. These differences highlight the importance of fully laying out all of the model assumptions at the single-cell level and then deriving the PDE model, so that the population-level model fully captures behaviours associated with the underlying cell-level assumptions, in all parameter regimes.

2.6 Comparing the impact of volume-filling effects in models for cell invasion into ECM

This section systematically compares the PDE model derived in this work, which incorporates volume-filling effects into both diffusion and proliferation of cells, with simpler models in the literature (see Table 2.1 for a summary) where these effects are only partially included or neglected entirely. By completing this comparison, we can identify the consequences of consistently modelling volume exclusion at the population level, and understand how model simplifications affect predictions of invasion speed.

The following model, proposed as a minimal model for tumour growth into ECM in [69]:

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left[(1-m)\frac{\partial u}{\partial x}\right] + u(1-u), \tag{2.57}$$

$$\frac{\partial m}{\partial t} = -\lambda m u, \tag{2.58}$$

assumes cell motility to be impacted by the presence of surrounding ECM only and that cell proliferation is impacted only by other cells, that is, the resource limiting cell proliferation is not space. Another similar model is presented in [43] to describe melanoma growth into skin and it is subsequently analysed in [100]. The model can be interpreted to assume that cell motility is decreased by ECM, and that cell proliferation is impacted by both other cells and ECM:

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left[(1-m)\frac{\partial u}{\partial x}\right] + u(1-u-m), \tag{2.59}$$

$$\frac{\partial m}{\partial t} = -\lambda m u. \tag{2.60}$$

The model variables and parameters in both of the aforementioned models are interpreted in the same way as in the model presented in this chapter (Equations (2.11) and (2.12)). They share the same equation governing ECM evolution (Equation (2.12)), assuming that ECM degradation occurs at rate λ. However, the main difference between the two is whether ECM also limits proliferation, and neither incorporate volume exclusion effects of the total cell and ECM density in the motility term.

2.6.1 Travelling wave analysis

Comparing the travelling wave analysis in Section 2.4 to the travelling wave analysis for Equations (2.59) and (2.60) shows that the additional motility-limiting term present in the full model, Equations (2.11) and (2.12), is lost during linearisation, yielding the same minimum wave speed as in [100]: $c_{\min} = 2(1 - m_0)$. Following the same methods, the Colson model, Equations (2.57) and (2.58), yields $c_{\min} = 2\sqrt{1 - m_0}$ [69].

As such, all volume-filling effects encoded solely in the nonlinear flux terms (within this suite of models) have no effect on the predicted minimum travelling wave speed, as they are all identical after linearisation. However, alterations to the net proliferation terms do significantly impact the minimum travelling wave speeds predicted by standard travelling wave analysis. In either case, the minimum wave speed recovered using linear stability analysis does not (for all parameter values) accurately predict the minimum travelling wave speed observed when the full nonlinear PDE system is simulated—a point explored further in later sections.

2.6.2 Numerical comparison

We are particularly interested in comparing the population-level behaviours of the PDE model for cell invasion into ECM presented in this work, which incorporates volume-filling effects into both the diffusion and proliferation of cells, to the simpler models without these volume-filling effects, presented in the literature. By looking at Figure 2.10, we can draw the following conclusions: all three models produce travelling wave solutions with a speed $c \geq c_{\min}$, where $c_{\min}$ is the minimum speed predicted by standard travelling wave analysis. In fact, all of these speeds are dependent on both the initial density of ECM ahead of the wave, m_0, and the rescaled ECM degradation rate, λ. The two models with the same reaction (growth) terms, depending on both cell and ECM preventing growth (Equations (2.11) and (2.12) and the Browning model [43, 100]), predict the same travelling wave speed $c_{\min} = 2(1 - m_0)$, that

is achieved numerically for $\lambda \to 0^+$. However, the model in Equations (2.57) and (2.58) presented in [69] predicts a speed $c_{\min} = 2\sqrt{1 - m_0}$, that is also revealed for $\lambda \to 0^+$. As a result, the behaviours observed in these models for small rescaled ECM degradation rates λ can be reproduced by studying a Fisher–KPP model with the appropriate parameters. In the same manner, by looking at Figure 2.10, it is clear that all three models produce travelling waves with speed $c \to 2^-$ as $\lambda \to \infty$. The behaviours observed here can be studied by considering the standard Fisher–KPP model (2.43) with all parameters equal to unity. The Fisher–KPP model (2.43) is also a suitable model simplification for all three systems when $m_0 = 0$.

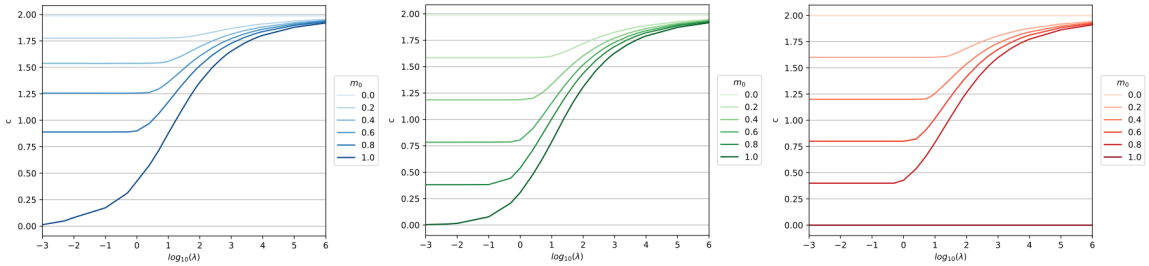

Figure 2.10: The relationship between the numerically estimated speed of travelling wave solutions to the system (2.57) and (2.58) on the left (blue), (2.59) and (2.60) in the middle (green) and (2.11) and (2.12) on the right (red), subject to the initial conditions (2.38) and (2.39). The numerically estimated travelling wave speed is obtained by tracing the point $X(t)$ such that $u(X(t), t) = 0.1$. Further specifics of the parameter values and the numerical methods used can be found in Section 2.3.

The transition between the two asymptotic regions is yet to be fully characterised for any of these models, but it is clear that c is a monotonic, increasing function of λ and m_0 for all of the models. The threshold value of λ above which the travelling wave speed begins to diverge from the analytically predicted minimum value is similar between models, depends on m_0, and increases as more volume-filling effects are included. Consistent with intuition, invasion is slower when volume exclusion acts on more aspects of cell behaviour (left to right in Figure 2.10).

The most marked qualitative difference is that the model in Equations (2.11) and (2.12) does not permit travelling waves from compactly supported initial u when $m_0 = 1$, because volume exclusion prevents cell invasion into regions already fully occupied by ECM. Other models [69, 100] do not bound the total density of cells and ECM above and therefore, for the parameters explored in this work, the cells can invade into a region where $u = 0$, $m = 1$, thus exhibiting travelling wave solutions even with compactly supported initial conditions.

For non-compactly supported initial data, results match those of the model described by Equations (2.59) and (2.60) (Section 2.4.2.1). Thus, the Browning model provides a useful qualitative simplification of the model presented in this chapter when $m_0 \in [0, 1)$, while the Fisher–KPP model accurately captures the asymptotic behaviour for $\lambda \to 0^+$, $\lambda \to \infty$, and $m_0 = 0$.

2.7 Discussion

In this chapter we studied a model for cell invasion into the surrounding ECM, focusing primarily on its travelling wave solutions. A key contribution of this work is the systematic derivation of the continuum PDE model directly from an underlying IBM that captures individual-level cell behaviours. In contrast to many previous approaches, which often adapt or assume pre-existing PDE forms to fit specific biological contexts, the derivation presented here ensures that volume-filling effects are consistently incorporated across all mechanisms—namely cell proliferation, diffusion, and ECM degradation. Existing models in the literature frequently treat volume exclusion only partially or heuristically, often neglecting its potential impact on one or more of these processes. By contrast, the approach presented in this chapter provides a unified framework grounded in the underlying microscopic dynamics.

By considering numerical solutions of the PDE model (2.11) and (2.12), we have been able to demonstrate a complex relationship between the travelling wave speed, c, the density of ECM far ahead of the wave of cells, m_0, and the rescaled ECM degradation rate, λ. Partial relationships between these parameters in asymptotic regimes of interest have been established, including that $c \to 2(1 - m_0)$ as $\lambda \to 0^+$, and that $c \to 2^-$ as $\lambda \to \infty$. A good agreement with the Fisher–KPP model has been demonstrated in the case where $\lambda \to \infty$, and we showed that the impacts of introducing volume-filling effects of cells to reduce cell movement (in comparison to the model in [100]) are minimal. As such, the Fisher–KPP model provides a suitable model simplification to reproduce the qualitative behaviours of the system in the case of a large ECM degradation rate, $\tilde{\lambda}$, compared to the proliferation rate, $\tilde{r}$. Since $\lambda = \tilde{\lambda}\tilde{K}/\tilde{r}$, the results equivalently suggest that as $\tilde{K} \to \infty$ the system can be well described by the Fisher–KPP model. In this model, the speed of the invasion front is given by $c_{\min} = 2$. For $\lambda \to 0^+$, which is representative of very large rescaled proliferation rates compared to the rescaled ECM degradation rates, or extremely small carrying capacities, the system can be studied by considering the simplification to a rescaled Fisher–KPP model (2.49). In this case, travelling waves are observed for $m \in [0, 1)$,

but the speed of the invasion front is now given by $c_{\min} = 2(1 - m_0)$. Converting back to dimensional variables, as with the Fisher–KPP model, the analytically predicted travelling wave speed increases with the cell proliferation rate, but with a more complicated relationship for the regions of parameter space corresponding to where the relationship between the travelling wave speed and rescaled ECM degradation rate is not yet well established.

It is also clear that qualitatively similar results are observed between this new model with volume-filling, and previously studied models outside this framework, as described by Table 2.1, in all cases where $m_0 \in [0, 1)$. Therefore, it could be said that the model originally proposed in [43] provides a good model simplification for any case where $m_0 \in [0, 1)$. In the case where $m_0 = 1$, the region that is initially uninvaded by cells is full with ECM, such that proliferation and movement of cells into this region is entirely prevented. Thus volume-filling effects prevent diffusion, so cells require space ahead of the wave front in order to invade the domain. This result provides the starkest difference between the model developed in this work and those previously studied elsewhere [69, 100]. It is observed that in the case of compactly-supported initial cell density, cell invasion cannot occur into the region where $m(x, 0) = 1$ and $u(x, 0) = 0$, and thus pinning occurs and travelling waves cannot form [348]. It is biologically reasonable to assume that an invading cell population might have zero density far ahead of the invading front. However it is important to note that the model considered here is a very simplistic model for cell invasion into ECM, and if further biological complications, such as the secretion of diffusible MMPs by cells to degrade and remodel ECM, were introduced then these phenomenological results would no longer be observed [261].

It can therefore be concluded that there exist simpler models for cell invasion into ECM, such as [69, 100, 109], that are defined by similar guiding principles and can be used to reproduce the qualitative behaviours of the travelling waves observed in the model presented in this work. Analysis of these systems confirms that the qualitative model predictions are conserved, and therefore the simpler models can be used in future studies to reduce computational complexity and render the resulting PDE model more analytically tractable. The disadvantage, however, is that in order to use these models to infer parameter values from data, extra steps would be required to validate whether the correct model has been selected. For example, analysis of cell trajectories can help infer the cell-cell interactions underlying the motility mechanism, and distinguish between the suite of models with qualitatively similar behaviours [38, 285, 312]. Our results reveal that the reaction term significantly impacts the

travelling wave speed for small and intermediate values of λ and thus it could be used to inform model development.

Chapter 3

An optimal control approach to nonlinear wave speed selection

This work was completed in collaboration with C. Falcó and R. E. Baker. C. Falcó and the author developed the optimal control approach and application to Fisher–Stefan models in tandem, however all other aspects of the research and writing presented in this chapter were completed by the author.

3.1 Introduction

As observed throughout this thesis, reaction–diffusion equations have long served as a powerful and widely-used mathematical framework for modelling the collective movement of populations. These models typically take the form of PDEs with density-dependent reaction and diffusion terms, enabling a detailed description of both the spatial and temporal evolution of a population across a variety of biological contexts.

A hallmark feature of such models is the emergence of travelling wave solutions that maintain a fixed profile while propagating at a constant speed. These travelling waves have become central to our theoretical understanding of diverse biological processes, from wound healing to tumour invasion, and species migration. In many of these settings, the wave connects an unstable steady state (*e.g.*, uncolonised space) to a stable steady state (*e.g.*, an established population), providing a natural mathematical description of an invasion front.

One of the seminal examples is the classical Fisher–KPP model (1.1). Along with its many extensions, the Fisher–KPP model illustrates how adapting the reaction and diffusion terms can substantially alter the speed, shape and structure of propagating waves [77, 109, 174, 224]. Considerable analytical effort has focused on understanding the influence of the various nonlinearities in these equations, with recent work yielding

increasingly refined existence and predicted wave speed results [36, 69, 87, 100, 121, 133, 147, 196, 198, 290, 318, 374].

Several complementary perspectives have emerged to explain how travelling wave speeds are selected, each with its own strengths and limitations. A foundational viewpoint uses the linear marginal stability criterion, which identifies the travelling wave with the steepest permissible exponential decay into the unstable state [337]. While this criterion provides an elegant and computable prediction, it relies on linearisation around the leading edge and therefore cannot account for cases where nonlinear interactions in the bulk of the wave influence the selected speed. To address this limitation, Benguria and Depassier introduced a variational method for wave speed selection, initially for nonlinear reaction terms [27, 28] and later extended to incorporate nonlinear diffusion [30]. Their approach generates rigorous upper and lower bounds on the admissible wave speeds and, crucially, provides a selection criterion identifying when the minimal admissible speed is indeed the realised one [26]. Hadeler and Rothe [133] clarified that, under suitable conditions, the asymptotic propagation speed can coincide with the linearly predicted value, even when the variational upper and lower bounds do not match. These results highlight cases where the linear and nonlinear speed selection criteria actually coincide. Though broadly undeveloped thereafter, more recent studies, such as that of Stokes et al. [320], have since extended these ideas by systematically analysing how density-dependent diffusion shapes both travelling wave existence and speed, thereby revealing parameter regimes where nonlinear diffusion either reinforces or invalidates predictions from earlier linear or variational theories.

Taken together, these developments establish a rich set of tools for analysing travelling waves in single-species systems, yet they also expose a notable gap: most classical speed-selection principles—including both the marginal stability criterion and the Benguria–Depassier variational framework—are formulated for scalar equations. However, many biological invasion processes, including those developed and studied in Chapter 2, fundamentally involve interactions between multiple populations (for example, cells and ECM). Extending our theory to determine travelling wave speed selection in these multi-species systems is therefore essential for connecting mechanistic models of collective migration with the analytical insights that have proved so valuable in the single-species setting.

This chapter addresses precisely this gap. Establishing a tractable lower bound for the wave speed in multi-species systems is particularly important for biological applications where numerical simulations suggest nonlinear wave speed selection but

analytic tools are lacking. A generalisation of the variational method therefore provides both conceptual insight and practical predictive power.

In this chapter, we extend the variational approach to systems of reaction–diffusion equations, aiming to understand how wave speed selection manifests in multi-species models of collective migration. We begin in Section 3.2 by revisiting the classical variational principle for a single species, before reconsidering it through the lens of optimal control theory. We then apply the resulting perspective to a number of examples in the literature before developing an optimal control approach for finding the travelling wave speed in models involving two interacting species. Finally, we use this new framework to obtain an improved lower bound for the travelling wave speed in systems describing cells invading into ECM, including some of those seen in Chapter 2.

3.2 A variational principle for one-species models

In this section, the goal is to introduce and justify the variational principle used to derive lower bounds on travelling wave speeds in one-species reaction–diffusion models. This scalar setting serves two purposes: first, it establishes the analytical framework that we later generalise to systems of equations; and second, it provides a benchmark against which we can assess how speed-selection mechanisms change when additional biological interactions are incorporated.

Following the approach by Benguria and Deppasier [27] and later developments such as those in [320], we begin by formulating a variational principle for the speed of one-species invading fronts. These are particular solutions to general reaction–diffusion equations of the form

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x}\left(D(\rho)\frac{\partial \rho}{\partial x}\right) + f(\rho)\,, \tag{3.1}$$

where $\rho = \rho(x,t)$ describes the density of the invading species, $D(\rho) > 0$ describes a density-dependent diffusivity, and $f(\rho)$ is a reaction term, often modelling net proliferation. The domain is $x \in \mathbb{R}$ and $t > 0$. We assume $f(0) = f(1) = 0$ to ensure we have biologically meaningful steady states. Equation (3.1) is complemented by the boundary conditions

$$\begin{cases} \rho(x,t) \to 1 & \text{as } x \to -\infty, \\ \rho(x,t) \to 0 & \text{as } x \to +\infty, \end{cases} \qquad \text{for all } t \geq 0.$$

The aim in what follows is threefold: to derive an inequality that produces a lower bound on the admissible wave speeds; to show how appropriate choices of test functions recover and improve classical predictions such as the linear speed; and to establish that this bound is sharp by demonstrating the existence of a test function that attains equality, thereby confirming that the formulation is indeed a variational principle. This structure mirrors the approach we will later extend to multi-species models.

To prepare for the variational characterisation, we first rewrite the PDE in travelling wave co-ordinates, which reduces the problem to a system of ODEs. To do this, we look for a monotonically decreasing function of $z = x - ct$ that joins the state $u = 1$ to the state $u = 0$ with constant travelling wave speed $c \geq 0$. Hence, we look for a solution of the form $\rho(x, t) = u(x - ct)$, with $u(-\infty) = 1$ and $u(+\infty) = 0$. In travelling wave co-ordinates, and denoting $\mathrm{d}u/\mathrm{d}z = -v$, Equation (3.1) becomes the following system of ODEs

$$\begin{aligned} \frac{\mathrm{d}u}{\mathrm{d}z} &= -v\,, \\ \frac{\mathrm{d}}{\mathrm{d}z}\left(D(u)v\right) &= -cv + f(u)\,, \end{aligned} \tag{3.2}$$

with $u(z) \to 1$ as $z \to -\infty$, $u(z) \to 0$ as $z \to \infty$, $v \geq 0$, and $v(z) \to 0$ as $z \to \pm\infty$. While, in general, monotonicity of travelling wave profiles is not guaranteed for all choices of $D(u)$ and $f(u)$, it has been rigorously established for a range of biologically relevant models, including the classical Fisher–KPP equation and the Porous–Fisher model [18, 171, 290], upon which this thesis focuses, and is a crucial assumption for the success of the variational principle in selecting the travelling wave speed. This reformulation makes explicit the relationship between the profile, $u(z)$, and the flux, $v(z)$, which is essential for the subsequent optimisation argument. Monotonicity is thereafter not merely a technical assumption: it enables us to treat $v(z)$ as a function of $u(z)$ and thus to convert the ODE system back into a scalar equation that is amenable to the variational approach. Without this structure, the derivation that follows would not be valid.

As such, it is therefore essential to determine, *a priori*, whether there are travelling wave solutions, and if they are monotonic. To do this, the system can be solved numerically to test if the resultant solutions satisfy these hypotheses, but it is essential to note that this is by no means rigorous, and intended only to indicate whether use of this method might be justified.

Assuming that monotonic travelling waves exist in all of the systems considered hereafter, we focus solely on the case where the travelling wave solutions $u(z)$ are monotonically decreasing with z (*i.e.*, $\mathrm{d}u/\mathrm{d}z < 0$), such that $v(z) = -\mathrm{d}u/\mathrm{d}z > 0$ can be rewritten as a function of $u(z)$, noting that $v(u = 0) = v(u = 1) = 0$. Using the chain rule to expand the left-hand side of Equation (3.2), we obtain

$$v(u)\frac{\mathrm{d}}{\mathrm{d}u}\left(D(u)v(u)\right) = cv(u) - f(u)\,. \tag{3.3}$$

The next step introduces a positive test function $\varphi(u)$ for $u \in [0, 1]$, following the classical variational framework. This manipulation converts the differential equation given in Equation (3.3) into an integral identity, from which a lower bound on c can be extracted. Multiplying Equation (3.3) by $D(u)\varphi(u)$, where $\varphi(u)$ is the smooth test function with $\phi(u) = -\mathrm{d}\varphi/\mathrm{d}u > 0$, and integrating by parts with respect to u, we find

$$\int_0^1 \left[-\frac{1}{2}\phi(u)D(u)^2v(u)^2 + cv(u)D(u)\varphi(u)\right]\mathrm{d}u = \int_0^1 D(u)f(u)\varphi(u)\,\mathrm{d}u\,. \tag{3.4}$$

Interpreting the left-hand side of the integrand in Equation (3.4) as a quadratic polynomial in v reveals a natural inequality:

$$-\frac{1}{2}\phi(u)D(u)^2v(u)^2 + cv(u)D(u)\varphi(u) \leq \frac{c^2\varphi(u)^2}{2\phi(u)}. \tag{3.5}$$

This step transforms the integral identity into a bound that holds for all admissible test functions. Equality in Equation (3.5) is only attained when

$$v(u) = v^*(u) := \frac{c\varphi(u)}{D(u)\phi(u)}\,. \tag{3.6}$$

Inserting the bound in Equation (3.5) into Equation (3.4) yields a family of lower bounds on the wave speed, indexed by the choice of test function:

$$c^2 \geq \frac{2\int_0^1 D(u)f(u)\varphi(u)\,\mathrm{d}u}{\int_0^1 \varphi(u)^2/\phi(u)\,\mathrm{d}u}. \tag{3.7}$$

This bound on the travelling wave speed (Equation (3.7)) holds for all smooth test functions $\varphi(u)$ with $\varphi'(u) = -\phi(u)$. The remainder of this section shows how to optimise these bounds.

To make the bound explicit, we now select a convenient one-parameter family of test functions, $\varphi_\beta(u) = ((1-u)/u)^\beta$, for $\beta \in [0, 2)$. This choice, previously used in

[26, 320], balances analytical tractability with the ability to capture the asymptotic behaviour near the steady states. Using this test function, we find

$$\frac{1}{2}c^2 \geq \sup_{\beta\in[0,2)} F(\beta), \quad \text{with } F(\beta) = \beta\,\frac{\int_0^1 D(u)f(u)u^{-\beta}(1-u)^{\beta}\,\mathrm{d}u}{\mathrm{B}(2-\beta,2+\beta)}\,. \tag{3.8}$$

Equation (3.8) uses $\mathrm{B}(x,y)$—the Beta function—which is defined by the integral

$$\mathrm{B}(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\mathrm{d}t, \tag{3.9}$$

where x and y are complex numbers with positive real parts. Equation (3.9) defines a symmetric function, such that $\mathrm{B}(x,y) = \mathrm{B}(y,x)$, and is closely related to the Gamma function, $\Gamma(z)$, through the relationship

$$\mathrm{B}(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}, \tag{3.10}$$

where the Gamma function is defined as

$$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}\mathrm{d}t,$$

where z is a complex number with positive real part.

To confirm that the inequality in Equation (3.8) defines a genuine variational principle, we must also show that there exists a test function, $\varphi = \hat{\varphi}$ (and correspondingly, $\phi = \hat{\phi}$) for which the bound is attained. This requires analysing the asymptotic behaviour of the travelling wave profile near $u = 0$ and $u = 1$ and verifying that the corresponding integrals converge. Equality in Equation (3.7) is attained (as per Equation (3.6)) when

$$v = \frac{c\hat{\varphi}(u)}{D(u)\hat{\phi}(u)} = -\frac{c\hat{\varphi}(u)}{D(u)\hat{\varphi}'(u)}, \tag{3.11}$$

where $' = \mathrm{d}/\mathrm{d}u$. Equation (3.11) can then be integrated to obtain

$$\hat{\varphi}(u) = \exp\left\{-c\int_{\bar{u}}^{u}\frac{\mathrm{d}q}{D(q)v(q)}\right\}, \tag{3.12}$$

for some $\bar{u} \in (0,1)$. Clearly $\hat{\varphi} > 0$ and is monotonically decreasing. Recalling Equation (3.3), which can be rewritten as

$$\frac{\mathrm{d}}{\mathrm{d}u}\Big(D(u)v(u)\Big) - c + \frac{f(u)}{v(u)} = 0,$$

we can employ the change of variables $p(u) = D(u)v(u)$, inspired by the denominator in the exponent in Equation (3.12), to obtain

$$p(u)\frac{\mathrm{d}}{\mathrm{d}u}p(u) - cp(u) + D(u)f(u) = 0. \tag{3.13}$$

In order to ensure convergence of $\hat{\varphi}(u)$ across all of $u \in [0,1]$, we must then analyse the asymptotic behaviour of Equation (3.13) as $u \to 0^+$ and $u \to 1^-$ [29].

To establish convergence, we begin by considering a linear analysis around the endpoint $u = 0$. Near $u = 0$, we can use linear expansions to write $f(u) \sim f'(0)u$ and $D(u) \sim D(0)$, since $f(0) = 0$ and $D(0) \neq 0$. Given that we are interested in smooth fronted travelling wave solutions, we employ the ansatz $p(u) = \alpha u$, where α is the positive root of $m^2 - cm + D(0)f'(0) = 0$. We find that

$$c \geq 2\sqrt{D(0)f'(0)} = c_L,$$

in order for α to be real. Hereon in, c_L is the travelling wave speed predicted by linear stability analysis. If we now substitute $p(u) = \alpha u$ into Equation (3.12), we find

$$\hat{\varphi}(u) = \exp\left\{ -\frac{c}{\alpha}\int_{\bar{u}}^{u} \frac{\mathrm{d}q}{q} \right\},$$

which when solved gives $\hat{\varphi} = (u/\bar{u})^{-c/\alpha}$. We note here that we most often choose $\bar{u} = u(z \to -\infty) = 1$, giving $\hat{\varphi} = u^{-c/\alpha}$. This result therefore ensures that all integrals in Equation (3.7) converge as long as $c/\alpha < 2$.

Now consider the linear analysis around the endpoint $u = 1$. Define the new variable, s, as $s = 1 - u$. Using Taylor series expansions, $f(u) \sim f'(1)(u-1) = -f'(1)s$ and $D(u) \sim D(1)$. Allowing $p(u) = r(1-u) = rs$ and employing this ansatz in Equation (3.13) then, following the same methods as before, we find that for $\bar{u} = 1$, $\hat{\varphi} = (1-u)^{c/r}$ with r as the positive root of the auxiliary equation $r^2 + cr + f'(1)D(1) = 0$, giving

$$r = \frac{-c + \sqrt{c^2 - 4D(1)f'(1)}}{2}.$$

As such, the integrals in the functional (3.8) must converge at $u = 1$ because $c/r > 0$ and does not restrict the speed. This is sufficient to show convergence of all of the integrals across the entire domain $[0,1]$ for an attainable test function $\hat{\varphi}$, such that Equation (3.8), does in fact define a variational principle.

3.2.1 An optimal control interpretation

The variational principle above, which was originally derived by Benguria and Depassier [26, 27, 28], can alternatively be viewed through the lens of an optimal control problem. In this reinterpretation, the travelling wave variable $u \in [0, 1]$ serves as the independent variable, whilst the local flux, given by $v(u) = -\mathrm{d}u/\mathrm{d}z > 0$, is the control variable. The idea is to recast the inequality leading to the speed bound as an optimisation problem in which one seeks the control $v(u)$ that makes this inequality as tight as possible.

In the optimal-control viewpoint, instead of bounding the left-hand side of Equation (3.4), we treat it as a functional to be minimised:

$$\mathcal{I}_c[v] = \int_0^1 \left[-\frac{1}{2}\phi(u)D(u)^2 v^2(u) + cD(u)\varphi(u)v(u) \right] \mathrm{d}u. \tag{3.14}$$

The rationale is as follows. For any admissible profile $v(u)$, the integral identity forces $\mathcal{I}_c[v]$ to take the fixed value $\int_0^1 D(u)f(u)\varphi(u)\,\mathrm{d}u$. However, among all controls that satisfy the ODE in Equation (3.3), there is one—denoted by v^* in the derivation in Section 3.2—for which the quadratic expression is tightest, *i.e.*, when the inequality used in deriving the speed bound becomes an equality. Identifying this optimal flux is therefore equivalent to solving an optimisation problem: finding the control that minimises $\mathcal{I}_c[v]$ subject to the dynamical equation. Therefore, the functional, $\mathcal{I}_c[v]$, appears naturally as the quantity whose minimisation reproduces the variational principle, and the speed bound follows by rearranging the resulting inequality.

Analogous to the co-state variable in Pontryagin's minimum principle, we can introduce the adjoint variable $\omega(u)$ that allows us to enforce the constraint in Equation (3.3) in a weak sense. Defining the Hamiltonian as

$$\begin{aligned} H(u, v, \omega) = &-\frac{\phi(u)}{2}D^2(u)v^2(u) + cv(u)D(u)\varphi(u) \\ &+ \omega\left[cv(u) - f(u) - v(u)\frac{\mathrm{d}}{\mathrm{d}u}\left(D(u)v(u)\right) \right], \end{aligned} \tag{3.15}$$

Pontryagin's minimum principle [176] states that for an optimal trajectory (v^*, ω), the Hamiltonian must satisfy

$$\frac{\partial H}{\partial v} = 0, \qquad \frac{\mathrm{d}\omega}{\mathrm{d}u} = -\frac{\partial H}{\partial u},$$

and that v^* minimises H pointwise in v.

Enforcing the constraint given in Equation (3.3) in Equation (3.15) gives

$$H(u, v, \omega) = -\frac{\phi(u)}{2}D^2(u)v^2(u) + cv(u)D(u)\varphi(u),$$

which, when computing the derivative with respect to v, gives

$$\frac{\partial H}{\partial v} = -\phi(u)D^2(u)v(u) + cD(u)\varphi(u) = 0. \tag{3.16}$$

The solution of Equation (3.16) is then

$$v(u) = v^* = \frac{c\varphi(u)}{D(u)\phi(u)},$$

which agrees with the condition given in Equation (3.6) derived in the variational argument above. As such, we can conclude that the pointwise minimiser of the functional (3.14) satisfies the same optimality condition as the control problem, and hence the one-species variational principle is indeed an optimality system. This equivalent optimal control interpretation aids us in developing a new lower bound on the travelling wave speed in coupled systems of equations with cross-dependence.

3.2.2 A model of epidermal wound healing

We now apply this variational approach to find a lower bound on the travelling wave speed in a number of biological examples from the literature. First, we consider a generalised, non-dimensional equation which can be used to study wound healing and other applications [306]. The governing reaction–diffusion equation takes the form

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x}\left(\rho^m \frac{\partial \rho}{\partial x}\right) + \rho(1-\rho^n), \tag{3.17}$$

for positive integers m, n. This is an example of the well-studied Porous–Fisher equation, with Richards' growth terms [87, 362]. We want to apply the variational principle to estimate the wave speed, using $D(u) = u^m$, and $f(u) = u(1-u^n)$. Recalling Equation (3.8), $F(\beta)$ can be calculated explicitly, using the relationship between the Beta and Gamma functions as described in Equation (3.10), as

$$\begin{aligned}
F(\beta) &= \frac{\beta}{B(2-\beta, 2+\beta)} \int_0^1 u^{m-\beta+1}(1-u)^\beta(1-u^n)\mathrm{d}u \\
&= \frac{\beta}{B(2-\beta, 2+\beta)} \int_0^1 \left[u^{m-\beta+1}(1-u)^\beta - u^{m+n-\beta+1}(1-u)^\beta\right]\mathrm{d}u \\
&= \frac{\beta}{\mathrm{B}(2-\beta, 2+\beta)} \left[\mathrm{B}(m-\beta+2, \beta+1) - \mathrm{B}(m+n-\beta+1, \beta+1)\right] \\
&= \frac{6\beta}{(1+\beta)\Gamma(2-\beta)} \left[\frac{\Gamma(m-\beta+2)}{\Gamma(m+3)} - \frac{\Gamma(m+n-\beta+2)}{\Gamma(m+n+3)}\right].
\end{aligned} \tag{3.18}$$

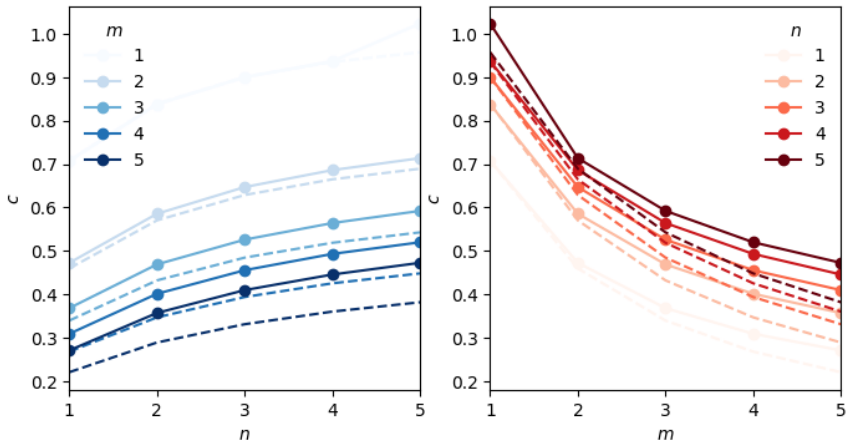

Figure 3.1: Plot of the numerically estimated minimum travelling wave speed (solid lines) and the minimum travelling wave speed of Equation (3.17) predicted using the variational principle with $F(\beta)$ found in Equation (3.18) (dashed lines) against n (left-hand side) and m (right-hand side). The numerically estimated travelling wave speed is obtained by tracing the point $X(t)$ such that $u(X(t), t) = 0.1$.

Taking the expression for $F(\beta)$ from Equation (3.18) and evaluating the supremum in Equation (3.8) gives a new lower bound on the travelling wave speed, c, which cannot be evaluated explicitly, but is plotted numerically in Figure 3.1.

Several key features are apparent in Figure 3.1. First, across all parameter regimes shown, the variational principle provides a strict lower bound, with the dashed curves always lying below the numerically observed minimum travelling wave speeds. The bound is tightest when the diffusion is closest to linear, *i.e.*, for small values of m, and when the reaction term is closest to logistic growth, *i.e.*, for small values of n. In these regimes, the wave profile behaves more similarly to the Fisher–KPP equation, for which the variational bound is known to be exact. Furthermore, as m increases, the diffusion exhibits a stronger degeneracy, leading to a sharper wave front. In this limit, the variational bound underestimates the travelling wave speed more strongly as the impact of the diffusion term increases. A similar trend occurs when n increases: the reaction term becomes more strongly nonlinear, moving the system away from the Fisher–KPP regime in which the variational bound is sharp. In both cases, the gap between the dashed and solid curves reflects precisely the departure from the classical marginal-stability scenario that underpins the variational method.

These trends motivate a closer examination of the special case $n = 1$ (*i.e.*, the Porous–Fisher equation), for which the variational expression simplifies substantially and exact analytical results are available for several integer values of m. When $n = 1$,

we employ Equation (3.8) to find

$$F(\beta) = \beta \frac{\mathrm{B}(m+2-\beta, 2+\beta)}{\mathrm{B}(2-\beta, 2+\beta)} = \beta \frac{\Gamma(4)}{\Gamma(4+m)} \frac{\Gamma(2-\beta+m)}{\Gamma(2-\beta)}, \tag{3.19}$$

which can be solved explicitly for small integer values of m, as $F(\beta)$ is a low-order polynomial in β. For instance, we recover the known values of the travelling wave speed when $m = 0$ and $m = 1$. For $m = 0$, Equation (3.19) reduces to $F(\beta) = \beta$, and thus we find $c \geq 2$ by Equation (3.8) (a.k.a. the Fisher–KPP travelling wave speed). For $m = 1$, using the fact that $\Gamma(z+1) = z\Gamma(z)$, we see

$$F(\beta) = \beta \frac{\Gamma(4)}{\Gamma(5)} \frac{\Gamma(3-\beta)}{\Gamma(2-\beta)} = \frac{\beta(2-\beta)}{4},$$

such that $c \geq 1/\sqrt{2}$. The case $m = 2$ also admits an explicit solution that exploits the identity

$$F(\beta) = \beta \frac{\Gamma(4)}{\Gamma(6)} \frac{\Gamma(4-\beta)}{\Gamma(2-\beta)} = \frac{\beta(3-\beta)(2-\beta)}{20},$$

which has a maximum at $\beta = -(\sqrt{7}-5)/3$, that when substituted into Equation (3.8) gives $c \geq \sqrt{(10+7\sqrt{7})/270}$. The next example illustrates a contrasting scenario: a pushed-wave regime where the variational principle becomes exact.

3.2.3 A model with degenerate density-dependent diffusion and the Allee effect

In this example we consider a degenerate, parabolic reaction–diffusion equation, adapted from that studied in [290] to demonstrate how the variational approach can be used to estimate the travelling wave speed when the Allee effect is present [317]. The governing equation takes the form

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x}\left(\left(\alpha\rho + \rho^2\right)\frac{\partial \rho}{\partial x}\right) + \rho(1-\rho)(\rho - a), \tag{3.20}$$

such that we are considering an equation of the form of Equation (3.1) with $D(\rho) = \alpha\rho + \rho^2$ and $f(\rho) = \rho(1-\rho)(\rho-a)$ where $\alpha > 0$ and $a \in [0, 0.5]$. Employing the

variational principle in Equation (3.8), we find

$$\begin{aligned}
F(\beta) &= \frac{\beta}{B(2-\beta,2+\beta)} \int_0^1 (\alpha u + u^2)(u-a)u^{1-\beta}(1-u)^{\beta+1}\,\mathrm{d}u \\
&= \frac{\beta}{B(2-\beta,2+\beta)} \int_0^1 \Big[\alpha u^{3-\beta}(1-u)^{\beta+1} + u^{4-\beta}(1-u)^{\beta+1} \\
&\qquad\qquad - a\alpha u^{2-\beta}(1-u)^{\beta+1} - a u^{3-\beta}(1-u)^{\beta+1}\Big]\mathrm{d}u \\
&= \frac{\beta}{B(2-\beta,2+\beta)} \Big[(\alpha-a)B(4-\beta,\beta+2) + B(5-\beta,\beta+2) \\
&\qquad\qquad - \alpha a B(3-\beta,\beta+2)\Big] \\
&= \frac{\beta}{120\Gamma(2-\beta)}[6(\alpha-a)\Gamma(4-\beta) + \Gamma(5-\beta) - 30\alpha a\Gamma(3-\beta)] \\
&= \frac{\beta(2-\beta)}{120}\left[(3-\beta)(4-\beta+6(\alpha-a)) - 30\alpha a\right], \qquad (3.21)
\end{aligned}$$

where the recursion formula $\Gamma(n) = (n-1)!$ is used to simplify the Gamma functions. Employing Equation (3.21) in the expression for the minimum travelling wave speed given in Equation (3.8), we numerically solve for the supremum and find a very good prediction of the numerically observed travelling wave speed, which can be observed in Figure 3.2.

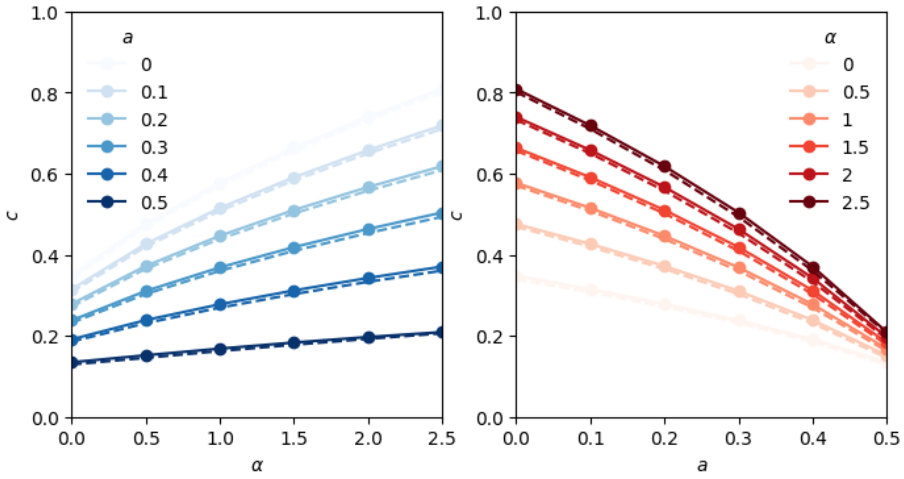

Figure 3.2: Plot of the numerically estimated minimum travelling wave speed (solid lines) and the minimum travelling wave speed of Equation (3.20) derived using the variational principle with $F(\beta)$ found in Equation (3.21) (dashed lines) as a function of α (left-hand side) and a (right-hand side). The numerically estimated travelling wave speed is obtained by tracing the point $X(t)$ such that $u(X(t),t) = 0.1$.

It is notable that the accuracy of the variational bound in this example is markedly higher than in the previous Porous–Fisher type case, and indeed becomes exact

throughout parameter space. This improvement is not accidental: it reflects key structural differences introduced by the Allee effect. When $f(u) = \rho(1-\rho)(\rho-a)$, the reaction term no longer exhibits logistic behaviour and the classical marginal-stability mechanism—where the wave speed is determined solely by the linearisation near $u = 0$—is replaced by a pushed-wave regime in which the entire interior of the wave profile contributes to speed selection. The variational approach presented in this chapter, although applicable to all types of waves, often has the best performance for pushed waves, where other methods like perturbation analysis are less successful due to the nonlinearities inherent in the system. In the case of pushed waves, the optimal test function reproduces the correct wave profile exactly, and the variational inequality becomes an equality.

In biological terms, the Allee effect suppresses growth at low densities, shifting the leading-edge dynamics away from the small u asymptotics that the variational principle relies on, and thus, in turn, reduces the accuracy of the variational bound. As a consequence, the variational method captures the dominant nonlinear contributions, producing a near-exact estimate across the full parameter ranges of α and a, as shown in Figure 3.2.

3.2.4 The Fisher–Stefan model

In this example, the variational principle is adapted to study a well-known moving boundary problem, the Fisher–Stefan model, which is described by

$$\frac{\partial \rho}{\partial t} = \frac{\partial^2 \rho}{\partial x^2} + \rho(1-\rho), \quad \text{for} \quad x < s(t), \tag{3.22}$$

and

$$\rho = 0, \quad \frac{\mathrm{d}s}{\mathrm{d}t} = -\kappa \frac{\partial \rho}{\partial x}, \quad \text{at} \quad x = s(t)\,, \tag{3.23}$$

where $\kappa \geq 0$ is a parameter that controls the speed of the moving boundary.

By fixing the moving boundary at $z = 0$, we can use the travelling wave ansatz, $z = x - s(t) = x - ct$ with $\rho(x,t) = u(z)$, to obtain

$$\begin{aligned} \frac{\mathrm{d}^2 u}{\mathrm{d}z^2} + c\frac{\mathrm{d}u}{\mathrm{d}z} + u(1-u) &= 0, \\ u(0) = 0, \quad \frac{\mathrm{d}}{\mathrm{d}z}u(0) &= -\frac{c}{\kappa}\,. \end{aligned} \tag{3.24}$$

As in the previous sections, expressing $v(u) = -\mathrm{d}u/\mathrm{d}z$ as a function of u, and integrating Equation (3.24) against a test function, $\varphi(u)$, we obtain

$$\int_0^1 \left(-\frac{1}{2}\phi(u)v^2(u) + cv(u)\varphi(u) \right) \mathrm{d}u + \frac{c^2\varphi(0)}{2\kappa^2} = \int_0^1 u(1-u)\varphi(u)\,\mathrm{d}u\,,$$

where we used that $v(0) = c/\kappa$. Using the same argument as previously shown on the second order polynomial in v, we obtain the bound

$$\frac{1}{2}c^2 \geq \frac{\int_0^1 u(1-u)\varphi(u)\,\mathrm{d}u}{\varphi(0)/\kappa^2 + \int_0^1 \varphi^2(u)/\phi(u)\,\mathrm{d}u},$$

with equality when

$$v(u) = \frac{c\varphi(u)}{\phi(u)}.$$

Here

$$\varphi(u) \propto \exp\left\{-\int_0^u \frac{c\,\mathrm{d}q}{v(q)}\right\},$$

which gives the condition

$$\varphi(0) = \phi(0)/\kappa\,,$$

in order to satisfy the moving boundary condition as described in Equation (3.24).

If we now assume that $\phi(u) = -\varphi'(u)$ and $\kappa\varphi(0) = \phi(0)$, then

$$\frac{1}{2}c^2 = \sup_{\varphi(u)} \frac{\int_0^1 u(1-u)\varphi(u)\,\mathrm{d}u}{\varphi(0)/\kappa^2 + \int_0^1 \varphi^2(u)/\phi(u)\,\mathrm{d}u}.$$

Since we can assume without loss of generality that $\varphi(0) = \tilde{\varphi}(0) = 1$, the simplest possible test function we could employ that satisfies these bounds and ensures that $\kappa\varphi(0) = \phi(0)$ is $\varphi(u) = e^{-\kappa u}$. Using this, we find

$$\frac{1}{2}c^2 \geq \frac{\int_0^1 u(1-u)e^{-\kappa u}\,\mathrm{d}u}{1/\kappa^2 + \int_0^1 e^{-\kappa u}/\kappa\,\mathrm{d}u} = \frac{\kappa - 2 + e^{-\kappa}(2+\kappa)}{\kappa(2-e^{-\kappa})}. \tag{3.25}$$

Equation (3.25) therefore provides a lower bound for the speed that is valid for all $\kappa > 0$; however, the quality of this bound degenerates as $\kappa \to +\infty$ (see Figure 3.3, where it is notable that the bound tends to one in the limit $\kappa \to \infty$ while the speed should actually tend to two [206]). Other test functions were employed, however no improvement on the prediction could be found (results not shown).

The degeneration of the bound in the limit $\kappa \to \infty$ is not just a shortcoming of the chosen test function, but reflects a fundamental limitation of the variational formulation in this regime. As κ increases, the influence of the Stefan boundary condition weakens and the Fisher–Stefan model approaches the classical Fisher–KPP equation, for which the travelling wave speed is selected by linear marginal stability at the leading edge.

The variational principle employed here encodes the moving boundary condition through constraints on admissible test functions at $u = 0$. In the limit $\kappa \to \infty$, these

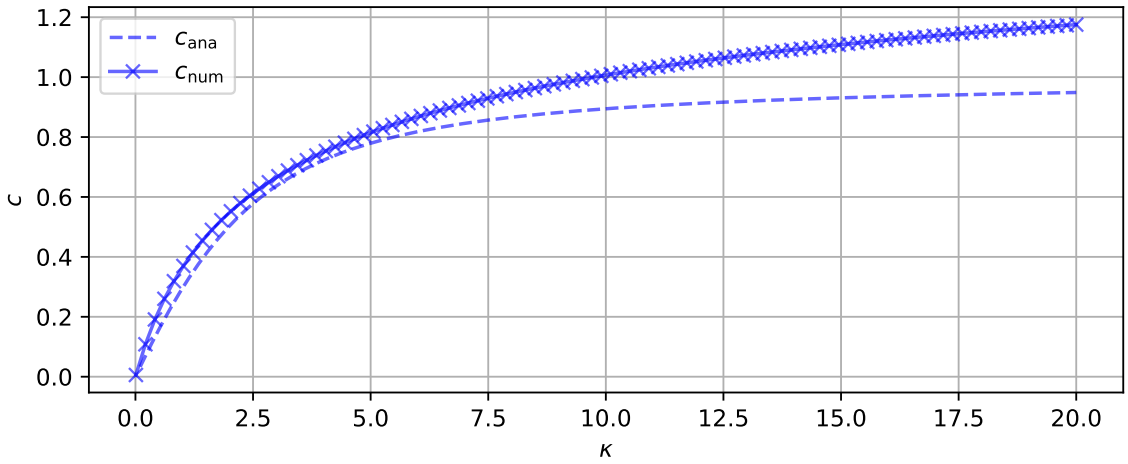

Figure 3.3: Plot of numerically simulated minimum travelling wave speed of the Fisher–Stefan model given by Equations (3.22) and (3.23) (solid lines) and analytically calculated travelling wave speed using the variational principle in Equation (3.25) (dashed lines) as a function of κ. The numerically estimated travelling wave speed is obtained by tracing the point $X(t)$ such that $u(X(t), t) = 0.1$.

constraints degenerate, and the class of admissible test functions no longer captures the linearised dynamics governing speed selection. As a consequence, the resulting bound saturates at a value strictly below the true wave speed $c = 2$.

This observation highlights that, while the variational approach yields a rigorous and informative lower bound for all $\kappa > 0$, it is most effective when wave speed selection is governed by nonlinear or boundary effects. Recovering the Fisher–KPP speed in the large κ limit using these techniques would require a reformulation of the variational principle that explicitly incorporates the linear leading edge behaviour in this asymptotic regime, which lies beyond the scope of the present work.

3.3 An optimal control approach to two-species models

In general, for models of two or more species, it is difficult to obtain an explicit variational principle due to the complex interactions present between two-species. As such, in this section, we instead demonstrate how to exploit the optimal control approach to uncover a lower bound for the minimum travelling wave speed, which improves the lower bound on the travelling wave speed found using linear stability analysis (as in Chapter 2) and is particularly useful in some specific parameter regimes

where a more accurate bound on the travelling wave speed makes further analytical insights possible.

One may ask how this approach provides additional information beyond direct numerical simulation of the original PDE system. While simulations can estimate travelling wave speeds for fixed parameter values, they do not yield rigorous bounds, nor do they clarify how specific mechanisms, such as cross-diffusion or species coupling, contribute to speed selection. The optimal control formulation developed here instead produces a computable lower bound on the wave speed that holds independently of numerical discretisation and initial data. Moreover, by expressing the problem in terms of a control variable and an associated adjoint system, this framework exposes how the second species influences the effective propagation dynamics of the first. This is particularly valuable in parameter regimes where linear stability analysis is inaccurate and where direct simulation becomes computationally demanding or opaque.

To begin, we describe this process for a general model of the form

$$\frac{\partial \rho_1}{\partial t} = \frac{\partial}{\partial x}\left(D(\rho_1, \rho_2)\frac{\partial \rho_1}{\partial x}\right) + f(\rho_1, \rho_2)\,, \tag{3.26}$$

$$\frac{\partial \rho_2}{\partial t} = g(\rho_1, \rho_2)\,, \tag{3.27}$$

where there are cross-species density-dependent diffusion and reaction terms. The boundary conditions are given by

$$\begin{cases} \rho_1(x,t) \to 1, & \rho_2(x,t) \to 0, \quad \text{as } x \to -\infty, \\ \rho_1(x,t) \to 0, & \rho_2(x,t) \to \nu, \quad \text{as } x \to +\infty, \end{cases}$$

and, as before, we have that $D(\rho_1, \rho_2) \geq 0$ is the density-dependent diffusivity, and $f(\rho_1, \rho_2)$ and $g(\rho_1, \rho_2)$ are the reaction terms, with $f(1, \rho_2) = f(0, \rho_2) = 0$ and $g(\rho_1, 0) = 0$.

As in the one-species setting, the variational approach relies on reducing the PDE to an ODE system along a travelling wave profile. This reduction requires that the primary population $u_1(z)$ is monotone. For general two-species systems, monotonicity is not guaranteed *a priori*; however, in the classes of models considered later in this chapter (*e.g.*, ECM degradation systems), numerical simulations in Chapter 2 consistently show monotone wave profiles, justifying this assumption. In more complex systems where oscillatory or non-monotone waves may arise, the reduction would fail and the variational principle would no longer apply.

Now, to make further progress, we look to rewrite ρ_1 as a monotonically decreasing function of $z = x - ct$, which defines $\rho_1(x,t) = u_1(x-ct)$ and $\rho_2(x,t) = u_2(x-ct)$ for constant $c \geq 0$. This gives the boundary conditions

$$\begin{cases} u_1(z) \to 1, & u_2(z) \to 0, \quad \text{as } z \to -\infty, \\ u_1(z) \to 0, & u_2(z) \to \nu, \quad \text{as } z \to \infty. \end{cases}$$

In travelling wave co-ordinates where we denote $\mathrm{d}u_1/\mathrm{d}z = -v$, we have the system

$$\begin{aligned} \frac{\mathrm{d}u_1}{\mathrm{d}z} &= -v, \\ \frac{\mathrm{d}}{\mathrm{d}z}(D(u_1,u_2)v) &= f(u_1,u_2) - cv, \\ \frac{\mathrm{d}u_2}{\mathrm{d}z} &= -\frac{1}{c}g(u_1,u_2). \end{aligned}$$

As before, if we can assume that u_1 and u_2 are monotonic, then we can rewrite $v(z) = v(u_1(z)) = v(u_1)$ and $u_2(z) = u_2(u_1(z)) = u_2(u_1)$ to remove the z dependence, which gives

$$v\frac{\mathrm{d}}{\mathrm{d}u_1}(D(u_1,u_2)v) = cv - f(u_1,u_2), \tag{3.28}$$

$$v\frac{\mathrm{d}u_2}{\mathrm{d}u_1} = \frac{1}{c}g(u_1,u_2). \tag{3.29}$$

Multiplying Equation (3.28) by $D(u_1,u_2)$ and a test function $\varphi(u_1)$ that satisfies $\varphi(0) = \varphi(1) = 0$ and $\varphi'(u_1) = -\phi(u_1)$, we can integrate by parts to obtain

$$\begin{aligned} \int_0^1 \left[-\frac{1}{2}\phi(u_1)D(u_1,u_2)^2 v(u_1)^2 + cv(u_1)D(u_1,u_2)\varphi(u_1) \right] \mathrm{d}u_1 \\ = \int_0^1 D(u_1,u_2)f(u_1,u_2)\varphi(u_1)\,\mathrm{d}u_1\,, \end{aligned} \tag{3.30}$$

where u_2 satisfies Equation (3.29).

To make progress, we now reformulate the problem within the framework of optimal control theory. To do this, we consider u_1 as the independent variable, $v(u_1) > 0$ to be the control, and the profile of the second species to be the state that we are trying to minimise. In this sense, introducing an adjoint variable and applying Pontryagin's minimum principle yields a simpler set of first-order optimality conditions. In the single-species case studied in Section 3.2.1, the constraint on the system was algebraic (in that it could be explicitly solved), whereas for two-species, the constraint now introduces an ODE for $u_2(u_1)$, requiring the use of the full Pontryagin system [176].

To proceed, we first denote the solution to Equation (3.29) by $\mathcal{U}_2[v](u_1)$ for a particular $v(u_1)$. Thereafter, we define $\mathcal{D}[v](u_1) = D(u_1,\mathcal{U}_2[v](u_1))$, $\mathcal{F}[v](u_1) = f(u_1,\mathcal{U}_2[v](u_1))$ and $\mathcal{G}[v](u_1) = g(u_1,\mathcal{U}_2[v](u_1))$.

Recalling Equation (3.30), the following integral identity

$$\int_0^1 \left\{ -\mathcal{D}[v](u_1)\mathcal{F}[v](u_1)\varphi(u_1) + c\mathcal{D}[v](u_1)\varphi(u_1)v(u_1) - \frac{1}{2}\mathcal{D}[v]^2(u_1)v^2(u_1)\phi(u_1) \right\} \mathrm{d}u_1 = 0\,, \tag{3.31}$$

serves as a starting point to construct the lower bound on the travelling wave speed. In particular, we define

$$\mathcal{I}_c[v] = \int_0^1 \left[-\frac{1}{2}\phi(u_1)\mathcal{D}[v]^2(u_1)v^2(u_1) + cv(u_1)\mathcal{D}[v](u_1)\varphi \right] \mathrm{d}u_1\,, \tag{3.32}$$

such that, if we look for v^* such that

$$v^* = \arg\min_v \mathcal{I}_c[v]\,,$$

subject to

$$\frac{\mathrm{d}\mathcal{U}_2}{\mathrm{d}u_1}[v](u_1) = \frac{\mathcal{G}[v](u_1)}{cv(u_1)}\,, \qquad \mathcal{U}_2[v](0) = 0, \qquad \text{and} \qquad \mathcal{U}_2[v](1) = \nu,$$

then all possible positive solutions v^* satisfying these constraints are admissible solutions.

To address this constrained minimisation problem, we introduce the adjoint variable, $\omega(u_1)$, and define the Hamiltonian

$$H(u_1,\mathcal{U}_2[v](u_1),v,\omega) = -\frac{\phi(u_1)}{2}\mathcal{D}[v]^2(u_1)v^2(u_1) + cv(u_1)\mathcal{D}[v](u_1)\varphi(u_1) + \frac{\omega(u_1)\mathcal{G}[v](u_1)}{cv(u_1)}\,.$$

Pontryagin's minimum principle implies that v^* minimises the Hamiltonian pointwise, *i.e.*, $\partial H/\partial v = 0$, together with the adjoint equation $\mathrm{d}\omega/\mathrm{d}u_1 = -\partial H/\partial\mathcal{U}_2$ and the state ODE [176]. The adjoint equation satisfies

$$\frac{\mathrm{d}\omega}{\mathrm{d}u_1} = v(u_1)\frac{\partial\mathcal{D}}{\partial\mathcal{U}_2}[v](u_1)\big[\phi(u_1)v(u_1)\mathcal{D}[v](u_1) - c\varphi(u_1)\big] - \frac{\omega(u_1)}{cv(u_1)}\frac{\partial\mathcal{G}}{\partial\mathcal{U}_2}[v](u_1),$$

with homogeneous endpoint conditions for ω, because both endpoints of $\mathcal{U}_2[v](u_1)$ are fixed. The stationarity condition $\partial H/\partial v = 0$ gives

$$-\phi(u_1)\mathcal{D}^2[v](u_1)v^3(u_1) + c\,\mathcal{D}[v](u_1)\varphi(u_1)v^2(u_1) - \frac{\omega(u_1)\mathcal{G}[v](u_1)}{c} = 0, \tag{3.33}$$

which implicitly determines v^* as the positive real root of this cubic equation. Although this expression (Equation (3.33)) cannot generally be solved in closed form, it can be evaluated numerically or approximated in particular parameter regimes. Substituting v^* and the corresponding $\mathcal{U}_2[v](u_1)$ into Equation (3.31), yields a condition that provides a computable lower bound for the travelling wave speed, c.

In order to improve the lower bound for the travelling wave speed that is found in Chapter 2, we now restrict our attention to a particularly relevant class of models where ρ_2 denotes a density field which gets degraded by the first species, such that $g(\rho_1, \rho_2) \propto -\rho_1\rho_2$. For example, when $g(\rho_1, \rho_2) = -\lambda\rho_1\rho_2$, we have the explicit solution

$$\mathcal{U}_2[v](u_1) = \nu \exp\left\{-\frac{\lambda}{c}\int_0^u \frac{q\,\mathrm{d}q}{v(q)}\right\}, \tag{3.34}$$

and bounds on the travelling wave speed, c, can be obtained numerically.

In contrast to numerical simulation, which must be repeated for each parameter set, the optimal control framework provides a systematic way to explore how changes in coupling strength or degradation rates alter the admissible range of travelling wave speeds. As such, the approach complements numerical experiments by providing both analytical structure and quantitative guarantees.

3.3.1 From the optimal control formulation to a variational approximation

The full optimality system described in Section 3.3 couples the control variable, $v(u_1)$, the secondary state, $\mathcal{U}_2[v](u_1)$, and the adjoint, $\omega(u_1)$, through a set of nonlinear first-order equations. In general, this system may only be solved numerically. However, in many parameter regimes the coupling term involving the adjoint, and the feedback of $\mathcal{U}_2[v](u_1)$ onto the diffusivity and reaction terms enter only as higher-order corrections during linearisation. This occurs because the feedback of $\mathcal{U}_2[v](u_1)$ onto the wave front is mediated only through smooth, multiplicative factors in $\mathcal{D}[v](u_1)$ and $\mathcal{F}[v](u_1)$ during Taylor series expansions. When the coupling parameter (*e.g.*, λ in $\mathcal{G}[v](u_1)$) is small, for example, the induced variations in $\mathcal{U}_2[v](u_1)$ remain $\mathcal{O}(\lambda)$ over the entire front since $\mathcal{U}_2[v](u_1) \in [0, 1]$. As a result, the adjoint equation introduces only $\mathcal{O}(\lambda)$ corrections to the optimality condition. In particular, the terms involving the adjoint $\omega(u_1)$, which account for the sensitivity of $\mathcal{U}_2[v](u_1)$ to variations in the control, contribute only higher-order corrections and are therefore neglected at leading order. In this limit, the leading-order stationarity condition reduces to the same pointwise minimisation that underlies the one-species variational principle.

Crucially, even in the two-species setting, $\mathcal{I}_c[v]$ retains the same concave quadratic dependence on the control v. Thus, the pointwise bound used in the one-species derivation (completing the square) still holds, regardless of the presence of the dynamical constraint relating u_1 and $\mathcal{U}_2[v](u_1)$. The resulting bound is

$$\mathcal{I}_c[v] \leq \frac{c^2}{2} \int_0^1 \frac{\varphi^2(u_1)}{\phi(u_1)} \, \mathrm{d}u_1.$$

Furthermore, recalling the integral identity in Equation (3.30), this bound on $\mathcal{I}_c[v]$ subsequently gives

$$\frac{1}{2}c^2 \geq \frac{\int_0^1 \mathcal{D}[v](u_1)\mathcal{F}[v](u_1)\varphi(u_1) \, \mathrm{d}u_1}{\int_0^1 \varphi(u_1)^2/\phi(u_1) \, \mathrm{d}u_1}, \tag{3.35}$$

with $\phi(u) = -\varphi'(u)$, and where equality is only achieved when

$$v^* = \frac{c\varphi(u_1)}{\mathcal{D}[v](u_1)\phi(u_1)} \, . \tag{3.36}$$

We now have a potentially analytically tractable lower bound on the travelling wave speed, representing the leading-order variational approximation of the full optimal control problem, obtained by neglecting the adjoint-dependent terms in the stationarity condition that arise from the non-local dependence of $\mathcal{U}_2[v](u_1)$ on the control. This approximation is valid when coupling effects enter only as higher-order corrections, for example when cross-species interactions are weak or vary smoothly across the wave front.

To show this is an equivalent variational principle, all that remains is to demonstrate the existence of a test function $\varphi(u_1) = \hat{\varphi}(u_1)$ for which equality holds. The same condition for equality holds in the two-species case (Equation (3.36)) as it did for one-species (see Equation (3.11)), which integrates to give

$$\hat{\varphi}(u_1) = \exp\left\{ -c \int_0^{u_1} \frac{\mathrm{d}q}{\mathcal{D}[v](q)v(q)} \right\}. \tag{3.37}$$

If $\mathcal{U}_2[v](u_1)$ is single-valued and smooth, and $\mathcal{D}[v](u_1)$ and $v(u_1)$ are continuous and bounded away from zero for $u \in [0,1]$, then provided that $\phi(u_1) = -\varphi'(u_1)$ does not vanish on $(0,1)$ (equivalently $\varphi(u_1)$ is strictly monotone), then the integrand in Equation (3.37) is continuous on $(0,1)$ and integrable on $[0,1]$. As such, Equation (3.35) is indeed a variational principle.

Again, as in the one-species case, we may choose any admissible one-parameter family of test functions and find the supremum with respect to that parameter to obtain a lower bound on the travelling wave speed. However, unlike the one-species

setting, this optimisation problem must now be carried out subject to a dynamic constraint. Each test function induces an ODE for $\mathcal{U}_2[v](u_1)$ (Equation (3.39)), and the solution of this ODE depends on the wave speed c. As a result, the variational bound is no longer a maximisation problem that can be solved analytically, but instead is a coupled problem in which both $\mathcal{U}_2[v](u_1)$ and the wave speed, c, must be determined numerically.

As before, we proceed by using the family of test functions given by $\varphi_\beta(u_1) = ((1-u_1)/u_1)^\beta$. This family is a natural choice for several reasons: (i) it has the correct monotonic behaviour and asymptotic structure near the endpoints; (ii) for $\beta \in [0,2)$, these functions guarantee the convergence of all integrals appearing in the variational bound, avoiding singularities introduced by the degeneracy of the diffusion or the reaction terms; and (iii) the dependence on a single parameter β allows the optimisation problem to remain computationally tractable, even though $\mathcal{U}_2[v](u_1)$ must now be determined dynamically. Employing this family of test functions, we obtain the bound

$$\frac{1}{2}c^2 \geq \sup_{\beta\in[0,2)} G_c(\beta), \tag{3.38}$$

with

$$G_c(\beta) = \beta\,\frac{\int_0^1 \mathcal{D}[v](u_1)\mathcal{F}[v](u_1)\,u_1^{-\beta}(1-u_1)^\beta\,\mathrm{d}u_1}{\mathrm{B}(2-\beta,2+\beta)}\,,$$

where $\mathcal{U}_2$ is subject to

$$\frac{\mathrm{d}\mathcal{U}_2}{\mathrm{d}u_1} = \frac{\beta}{c^2}\frac{\mathcal{G}[v](u_1)\mathcal{D}[v](u_1)}{u_1(1-u_1)}\,. \tag{3.39}$$

In order to simplify notation in the following applications, we define

$$N(\beta) = \int_0^1 \mathcal{D}[v](u_1)\mathcal{F}[v](u_1)\,u_1^{-\beta}(1-u_1)^\beta\,\mathrm{d}u_1\,, \tag{3.40}$$

such that

$$G_c(\beta) = \beta\,\frac{N(\beta)}{\mathrm{B}(2-\beta,2+\beta)}\,.$$

3.3.2 A criterion for nonlinear wave speed selection

Now that we understand how to better characterise the travelling wave speed for nonlinear coupled systems of equations, we now want to extend the results of [27, 320] to provide a criterion to evaluate, in advance, whether the speed is selected nonlinearly, *i.e.*, whether $c > 2\sqrt{f'(0)D(0)}$ in the case of a single-species model.

Since the minimal admissible wave speed is determined by $c^2 = 2\sup_{0\leq\beta<2} G_c(\beta)$, and since the linearly predicted (pulled) speed corresponds to the boundary value

$\beta = 2$, nonlinear (pushed) selection occurs precisely when the supremum of G_c is attained at some interior point $\beta < 2$. A sufficient and necessary local condition for this is that $G_c(\beta)$ increases as β decreases from two, *i.e.*,

$$\frac{\mathrm{d}}{\mathrm{d}\beta} G_c(\beta)\bigg|_{\beta=2^-} < 0.$$

Our goal is therefore to evaluate this derivative explicitly and obtain a criterion expressed directly in terms of the nonlinear diffusion and reaction terms.

We first focus on $N(\beta)$ as given in Equation (3.40). If we define $\mathcal{R}[v](u_1) = \mathcal{F}[v](u_1)/u_1$ and $\mathcal{P}[v](u_1) = \mathcal{D}[v](u_1)\mathcal{R}[v](u_1)$, we can rewrite $N(\beta)$ as

$$\begin{aligned}
N(\beta) &= \int_0^1 \mathcal{P}[v](u_1) u_1^{(1-\beta)} (1-u_1)^\beta \mathrm{d}u_1 \\
&= \mathcal{P}[v](0) \int_0^1 u_1^{(1-\beta)} (1-u_1)^\beta \mathrm{d}u_1 + \int_0^1 \Big(\mathcal{P}[v](u_1) - \mathcal{P}[v](0)\Big) u_1^{(1-\beta)} (1-u_1)^\beta \mathrm{d}u_1 \\
&= \mathcal{P}[v](0)\mathrm{B}(2-\beta, 1+\beta) + \int_0^1 \Big(\mathcal{P}[v](u_1) - \mathcal{P}[v](0)\Big) u_1^{(1-\beta)} (1-u_1)^\beta \mathrm{d}u_1.
\end{aligned}$$

If we then introduce $\epsilon = 2 - \beta$ and consider the limit $\epsilon \to 0^+$ (*i.e.*, $\beta \to 2$), then

$$\begin{aligned}
N(2-\epsilon) &= \mathcal{P}[v](0)\mathrm{B}(\epsilon, 3-\epsilon) + \int_0^1 \Big(\mathcal{P}[v](u_1) - \mathcal{P}[v](0)\Big) u_1^{(\epsilon-1)} (1-u_1)^{(2-\epsilon)} \mathrm{d}u_1 \\
&= \mathcal{P}[v](0)\Big[\frac{1}{\epsilon} - \frac{3}{2} + \mathcal{O}(\epsilon)\Big] \\
&\qquad + \int_0^1 \Big(\mathcal{P}[v](u_1) - \mathcal{P}[v](0)\Big) u_1^{-1} (1-u_1)^2 e^{\epsilon \log(u_1) - \epsilon \log(1-u_1)} \mathrm{d}u_1,
\end{aligned}$$

using expansions of the beta function around small ϵ. Expanding the exponential similarly, we find to $\mathcal{O}(1)$:

$$N(2-\epsilon) = \mathcal{P}[v](0)\Big[\frac{1}{\epsilon} - \frac{3}{2}\Big] + \int_0^1 \Big(\mathcal{P}[v](u_1) - \mathcal{P}[v](0)\Big) u_1^{-1} (1-u_1)^2 \mathrm{d}u_1.$$

Substituting this into $G_c(\beta)$, we obtain to first order

$$\begin{aligned}
G_c(2-\epsilon) &= (2-\epsilon) \frac{\mathcal{P}[v](0)\Big[\frac{1}{\epsilon} - \frac{3}{2}\Big] + \int_0^1 \Big(\mathcal{P}[v](u_1) - \mathcal{P}[v](0)\Big) u_1^{-1} (1-u_1)^2 \mathrm{d}u_1}{\mathrm{B}(\epsilon, 4-\epsilon)} \\
&= (2-\epsilon) \frac{\mathcal{P}[v](0)\Big[\frac{1}{\epsilon} - \frac{3}{2}\Big] + \int_0^1 \Big(\mathcal{P}[v](u_1) - \mathcal{P}[v](0)\Big) u_1^{-1} (1-u_1)^2 \mathrm{d}u_1}{\frac{1}{\epsilon} - \frac{11}{6}} \\
&= (2-\epsilon) \frac{\mathcal{P}[v](0) + N_0 \epsilon}{1 - \frac{11}{6}\epsilon},
\end{aligned}$$

where

$$N_0 = \int_0^1 \Big(\mathcal{P}[v](u_1) - \mathcal{P}[v](0)\Big) u_1^{-1}(1-u_1)^2 \mathrm{d}u_1 - \frac{3}{2}\mathcal{P}[v](0). \tag{3.41}$$

Then, expanding the denominator, we find

$$G_c(2-\epsilon) = 2\mathcal{P}[v](0) + \left(2N_0 + \frac{8}{3}\mathcal{P}[v](0)\right)\epsilon + \mathcal{O}(\epsilon^2). \tag{3.42}$$

Since $\epsilon = 2 - \beta > 0$, the expansion in Equation (3.42) shows that $G_c(\beta)$ increases as β decreases from two if and only if $2N_0 + 8/3\mathcal{P}[v](0) > 0$. Equivalently,

$$N_0 > -\frac{4}{3}\mathcal{P}[v](0).$$

Recalling the definition of N_0 in Equation (3.41), this condition can be written explicitly as

$$\begin{aligned}\int_0^1 \Big(\mathcal{P}[v](u_1) - \mathcal{P}[v](0)\Big) u_1^{-1}(1-u_1)^2 \mathrm{d}u_1 &\geq \frac{1}{6}\mathcal{P}[v](0) \\ &= \frac{1}{6}\mathcal{D}[v](0)\frac{\partial}{\partial u_1}\mathcal{F}[v](u_1)|_{u_1=0}. \end{aligned}\tag{3.43}$$

Condition (3.43) provides a sufficient criterion for nonlinear (pushed) wave speed selection in both two-species and single-species reaction–diffusion systems. When it is satisfied, the supremum of $G_c(\beta)$ is attained at some $\beta < 2$, and hence the selected wave speed strictly exceeds the linearly predicted value.

Consider $D(u) = u + \delta$ and $f(u) = u(1-u)$, as studied in [320]. In this case, $P(u) = (u+\delta)(1-u)$ and $P(0) = \delta$. Substituting into the criterion (3.43) gives

$$\int_0^1 u(1-u-\delta)(1-u)^2 \,\mathrm{d}u \;\geq\; \frac{1}{6}\delta,$$

which holds if and only if $\delta < 1/2$. Hence the criterion correctly predicts nonlinear wave speed selection for $\delta < 1/2$, in agreement with the results of [320].

3.3.3 Application to models of cell invasion into ECM

We now employ the derived framework to improve the estimated travelling wave speed in the class of models studied in Section 2.6. The models of focus (Equations (2.57) and (2.58), and Equations (2.59) and (2.60)) describe collective cell migration into ECM governed by the following dynamical system

$$\frac{\partial \rho_1}{\partial t} = \frac{\partial}{\partial x}\left((1-\rho_2)\frac{\partial \rho_1}{\partial x}\right) + f(\rho_1, \rho_2)\,, \tag{3.44}$$

$$\frac{\partial \rho_2}{\partial t} = -\lambda \rho_1 \rho_2\,, \tag{3.45}$$

where $f(0,\rho_2)=0$ and $f(1,\rho_2)=0$. The boundary conditions are given by

$$\begin{cases} \rho_1(x,t)\to 1, & \rho_2(x,t)\to 0, \quad \text{as } x\to-\infty,\\ \rho_1(x,t)\to 0, & \rho_2(x,t)\to \nu, \quad \text{as } x\to+\infty, \end{cases}$$

for $0\le\nu\le 1$.

We consider two different reaction terms. Firstly, $f_C(\rho_1,\rho_2)=f_C(\rho_1)=\rho_1(1-\rho_1)$ as described in [69]. Secondly, we consider $f_B(\rho_1,\rho_2)=\rho_1(1-\rho_1-\rho_2)$ as detailed in [43, 100]. For both models, numerical simulations and the analysis in Section 2.6 shows that the system exhibits travelling waves that attain the linearly predicted wave speed ($2\sqrt{1-\nu}$ when $f=f_C$, and $2(1-\nu)$ when $f=f_B$) for small values of the ECM degradation rate, $\lambda\le\lambda^*(\nu)$, only, where $\lambda^*(\nu)$ is the critical ECM degradation rate such that the linearly predicted and numerically observed travelling wave speeds coincide for $\lambda\le\lambda^*(\nu)$. For $\lambda>\lambda^*(\nu)$, the numerically estimated wave speed shows significant deviations from the linearly selected travelling wave speed and the critical value of the degradation rate, $\lambda^*(\nu)$, seems to follow a monotonically increasing relationship with ν as $\nu\to 1$ (see Figure 2.3).

Recalling the travelling wave analysis, we can translate the system into the monotone wave co-ordinate $z=x-ct$, which gives $\rho_1(x,t)=u_1(x-ct)=u_1(z)$ and $\rho_2(x,t)=u_2(x-ct)=u_2(z)$ that are described by

$$\frac{\mathrm{d}}{\mathrm{d}z}\left((1-u_2)\frac{\mathrm{d}u_1}{\mathrm{d}z}\right)+c\frac{\mathrm{d}u_1}{\mathrm{d}z}+f(u_1)=0, \tag{3.46}$$

$$c\frac{\mathrm{d}u_2}{\mathrm{d}z}-\lambda u_2u_1=0. \tag{3.47}$$

The boundary conditions of this system now read

$$\begin{cases} u_1(z)\to 1, & u_2(z)\to 0, \quad \text{as } z\to-\infty,\\ u_1(z)\to 0, & u_2(z)\to \nu, \quad \text{as } z\to+\infty. \end{cases}$$

Following the optimal control approach described in Section 3.3, we introduce the variable $v(u_1)=-\mathrm{d}u_1/\mathrm{d}z$, that gives the integral identity

$$\int_0^1(1-u_2)\Big(-u_1(1-u_1)\varphi-\frac{1}{2}\phi(1-u_2)v^2+cv\varphi\Big)\,\mathrm{d}u_1=0, \tag{3.48}$$

and define the cost for this system (Equations (3.46) and (3.47) subject to $f_C(u_1)=u_1(1-u_1)$) as

$$\mathcal{I}_C[v]=\int_0^1(1-u_2)\Big(-\frac{1}{2}\phi(1-u_2)v^2+cv\varphi\Big)\,\mathrm{d}u_1=\int_0^1(1-u_2)u_1(1-u_1)\varphi\,\mathrm{d}u_1, \tag{3.49}$$

subject to

$$\frac{\mathrm{d}u_2}{\mathrm{d}u_1} = -\frac{\lambda u_1 u_2}{cv}, \qquad u_2(0) = \nu \qquad \text{and} \qquad u_2(1) = 0, \tag{3.50}$$

where we now consider all of the variables as functions of u_1 only for ease. Introducing the adjoint variable $\omega(u_1)$, we have the adjoint equation defined by

$$\frac{\mathrm{d}\omega}{\mathrm{d}u_1} = v(c\varphi - \tfrac{1}{2}\phi v(1-u_2)) - \frac{\lambda u_1 \omega}{c}, \tag{3.51}$$

and the state equation for stationarity, as described by Equation (3.33), given by

$$-\phi(1-u_2)^2 v^3 + c(1-u_2)\varphi v^2 + \frac{\lambda}{c}\omega u_2 u_1 = 0. \tag{3.52}$$

Together, Equations (3.50)–(3.52) define the constraints on the optimal control problem which seeks to satisfy the integral identity given by Equation (3.48).

If we now consider a weak coupling between the PDE and ODE governing dynamics, in the sense that variations in the secondary state induce only higher-order corrections to the diffusion and reaction terms, then the adjoint-dependent contributions to the optimality condition are not dominant. In this regime, the leading-order behaviour of the wave speed is captured by the variational approximation described by Equation (3.38). We first need to solve the ODE described in Equation (3.39) for $f_C(u_1, u_2) = u_1(1-u_1)$ and $D(u_1, u_2) = 1 - u_2$:

$$\frac{\mathrm{d}u_2}{\mathrm{d}u_1} = -\frac{\lambda\beta}{c^2}\frac{u_2(1-u_2)}{1-u_1}.$$

Rearranging and integrating gives

$$u_2(u_1) = \frac{\nu(1-u_1)^{\frac{\lambda\beta}{c^2}}}{\nu(1-u_1)^{\frac{\lambda\beta}{c^2}} + 1 - \nu}. \tag{3.53}$$

The ODE for $u_2(u_1)$ is well-posed on $[0,1]$ and $v(u_1) > 0$, ensuring a unique smooth solution for each c and β.

For $f_C(u_1, u_2) = u_1(1-u_1)$ and $D(u_1, u_2) = 1 - u_2$, we use Equation (3.53) to find that Equation (3.40) becomes

$$N_C(\beta) = \int_0^1 \frac{u_1^{1-\beta}(1-u_1)^{1+\beta}}{\frac{\nu}{1-\nu}(1-u_1)^{\frac{\lambda\beta}{c^2}} + 1}\mathrm{d}u_1, \tag{3.54}$$

and the lower bound on the travelling wave speed can be subsequently be found by solving

$$\frac{1}{2}c^2 \geq \sup_{\beta\in[0,2)} \beta\,\frac{N_C(\beta)}{\mathrm{B}(2-\beta, 2+\beta)}\,. \tag{3.55}$$

Fig. 3.4A shows the result of numerically solving Equation (3.55) for the lower bound on the travelling wave speed, c, which, when compared to the linearly predicted travelling wave speed, $c_{\text{linear}} \geq 2\sqrt{1-\nu}$ (horizontal lines that match the speed in Figure 3.4A for small $\lambda \to 0^+$), demonstrates a vast improvement. To the best of our knowledge, the linear travelling wave speed is the current best estimate of the speed of invasion in the literature, so the prediction derived in this chapter using the variational approach drastically improves the lower bound on the wave speed.

In the case when the reaction terms for the first species also depend on the density of the second species, as described by $f_B(u_1, u_2) = u_1(1 - u_1 - u_2)$, then Equations (3.50), (3.51) and (3.52) still hold, but instead we have a cost functional defined by

$$\mathcal{I}_B[v] = \int_0^1 (1-u_2)\Big(- u_1(1-u_1)\varphi - \frac{1}{2}\phi(1-u_2)v^2 + cv\varphi\Big)\mathrm{d}u_1 = \int_0^1 -u_1^2(1-u_1)\mathrm{d}u_1. \tag{3.56}$$

In the case when λ is small, so that the system is only weakly coupled, the adjoint contributions to the stationarity equation are very small, and we can again take a variational approach to find the travelling wave speed. The solution to Equation (3.39) is given by Equation (3.53), and thus

$$\begin{aligned} N_B(\beta) &= \int_0^1 (1-u_2(u_1))(1-u_1-u_2(u_1))\, u_1^{1-\beta}(1-u_1)^{\beta}\,\mathrm{d}u_1, \\ &= \int_0^1 \frac{u_1^{1-\beta}(1-u_1)^{\beta} - u_1^{2-\beta}(1-u_1)^{\beta} - A(1-u_1)^{(}1-\beta)(1-u_1)^{\beta+\eta}}{(1+A(1-u_1)^{\eta})^2}, \end{aligned} \tag{3.57}$$

where we have defined $A = \nu/(1-\nu)$ and $\eta = \lambda\beta/c^2$. Substituting $N_B(\beta)$ into equation (3.38) gives the travelling wave speed plotted by the dashed line in Figure 3.4B, which again provides a vast improvement on that predicted by linear stability analysis, namely $c \geq 2(1-\nu)$.

3.3.4 Extension to models with taxis terms

Benguria et al. have, more recently, considered the addition of advection terms in the governing equations and their impact on the resulting travelling wave speed [30]. As is natural, we therefore present a simple extension to the results in Section 3.3.1 that is applicable to populations undergoing taxis, or following the gradient of external

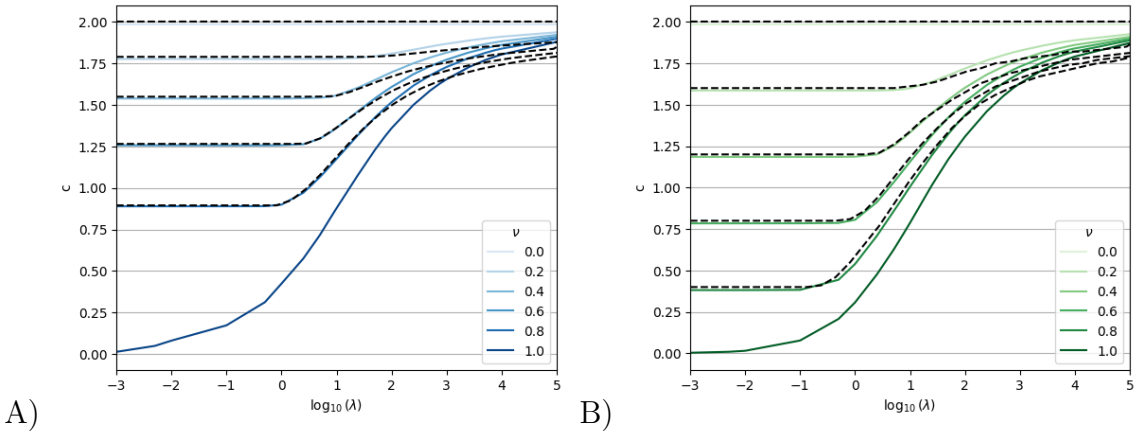

Figure 3.4: Plot of numerically simulated minimum travelling wave speed (solid lines) and analytically calculated travelling wave speed of Equations (3.44) and (3.45) using the variational principle (Equation (3.38), dashed lines) with $N_C(\beta)$ as calculated in Equation (3.54) (left-hand side) and $N_B(\beta)$ as calculated in Equation (3.57) (right-hand side) for various initial ECM densities $\nu \geq 0$. The numerically estimated travelling wave speed is obtained by tracing the point $X(t)$ such that $u_1(X(t), t) = 0.1$.

signals, for example, chemicals or substrates. The model of interest will take the form

$$\frac{\partial \rho_1}{\partial t} = \frac{\partial}{\partial x}\left(D(\rho_1, \rho_2)\frac{\partial \rho_1}{\partial x} - \chi(\rho_1, \rho_2)\frac{\partial \rho_2}{\partial x}\right) + f(\rho_1, \rho_2)\,, \tag{3.58}$$

$$\frac{\partial \rho_2}{\partial t} = g(\rho_1, \rho_2)\,, \tag{3.59}$$

where the boundary conditions of the system are given by

$$\begin{cases} \rho_1(x,t) \to \rho_{1,-}, \quad \rho_2(x,t) \to \rho_{2,-}, \quad \text{as } x \to -\infty, \\ \rho_1(x,t) \to \rho_{1,+}, \quad \rho_2(x,t) \to \rho_{2,+}, \quad \text{as } x \to +\infty, \end{cases}$$

This model differs from the previously studied simple reaction–diffusion system by the inclusion of an advection term in Equation (3.58).

For a right-moving travelling wave with $c > 0$ to exist, the boundary states $\rho_{1,-}, \rho_{2,-}$ and $\rho_{1,+}, \rho_{2,+}$ must be ordered so that the invaded state lies to the left, as $z \to -\infty$, and the uninvaded state to the right, as $z \to \infty$. That is, we require $f(\rho_{1,+}, \rho_{2,+}) = 0$ with $\mathrm{d}/\mathrm{d}\rho_1(f(\rho_{1,+}, \rho_{2,+})) > 0$ to ensure instability ahead of the front, and $f(\rho_{1,-}, \rho_{2,-}) = 0$ with $\mathrm{d}/\mathrm{d}\rho_1(f(\rho_{1,-}, \rho_{2,-})) < 0$ to ensure stability behind the front. These sign conditions guarantee invasion from the stable state toward the unstable one, forcing the wave to propagate with positive speed.

Therefore, for $D(\rho_1, \rho_2) \geq 0$ the density-dependent diffusivity, $\chi(\rho_1, \rho_2) \geq 0$ the tactic sensitivity, and $f(\rho_1, \rho_2)$ and $g(\rho_1, \rho_2)$ the reaction terms, we can then seek a

monotonically decreasing function of $z = x - ct$, defining $\rho_1(x,t) = u_1(x-ct)$ and $\rho_2(x,t) = u_2(x-ct)$, where

$$\begin{cases} u_1(z) \to \rho_{1,-}, & u_2(z) \to \rho_{2,-}, & \text{as } z \to -\infty, \\ u_1(z) \to \rho_{1,+}, & u_2(z) \to \rho_{2,+}, & \text{as } z \to \infty. \end{cases}$$

If we consider a weak coupling between the external cue and the population of interest, we can again follow the variational principal approach detailed in Section 3.3.1 (a natural simplification of the optimal control framework detailed in Section 3.3). Namely, setting $\mathrm{d}u_1/\mathrm{d}z = -v$, we have

$$\begin{aligned} \frac{\mathrm{d}u_1}{\mathrm{d}z} &= -v, \\ \frac{\mathrm{d}}{\mathrm{d}z}\left(D(u_1,u_2)v + \chi(u_1,u_2)\frac{\mathrm{d}u_2}{\mathrm{d}z}\right) &= f(u_1,u_2) - cv, \\ \frac{\mathrm{d}u_2}{\mathrm{d}z} &= -\frac{1}{c}g(u_1,u_2). \end{aligned}$$

If we replace all instances of $D(u_1,u_2)v$ by $D(u_1,u_2)v + \chi(u_1,u_2)\mathrm{d}u_2/\mathrm{d}z$ in Equations (3.28)–(3.35), we can show the existence of a variational principle for taxis-style models. Excluding these computations for brevity, we find

$$\frac{1}{2}c^2 \geq \frac{\int_0^1 f(u_1,u_2(u_1))\varphi(u_1)\left(D(u_1,u_2) - \chi(u_1,u_2)\frac{\mathrm{d}u_2}{\mathrm{d}u_1}\right)\mathrm{d}u_1}{\int_0^1 \varphi(u_1)^2/\phi(u_1)\,\mathrm{d}u_1},$$

with $\phi(u_1) = -\varphi'(u_1)$, as long as

$$\frac{\mathrm{d}u_2}{\mathrm{d}u_1} = \frac{\phi(u_1)\left(D(u_1,u_2) - \chi(u_1,u_2)\frac{\mathrm{d}u_2}{\mathrm{d}u_1}\right)g(u_1,u_2)}{c^2\varphi(u_1)}. \tag{3.60}$$

The same test functions and proof can be used to show that this is indeed a variational principle. However due to lack of relevant examples in the literature, this method is not shown to be applied to any models in this thesis, but instead detailed here to demonstrate the utility and wider applicability of the approach to these models, where they do indeed exist.

3.4 Discussion

In this chapter, a unified framework for understanding travelling wave speed selection in nonlinear reaction–diffusion systems has been developed by recasting the classical

variational approach within the language of optimal control theory [26, 27, 28, 176]. This interpretation reveals that the standard one-species variational principle is a particularly tractable instance of a more general control problem, in which the travelling wave profile plays the role of a state, the local flux acts as the control, and the speed emerges as the optimal value that satisfies a stationarity condition. One immediate benefit of this perspective is conceptual: it exposes the variational principle as the outcome of a pointwise minimisation of an associated Hamiltonian, rather than as an algebraic manipulation tailored to scalar PDEs. More importantly, this viewpoint provides a natural route to generalising the theory to coupled systems, where the traditional variational arguments are not explicitly solvable.

By exploiting this optimal control formulation, it is shown how a tractable variational approximation can be obtained for two-species systems in regimes where the coupling between species' is either weak or sufficiently smooth across the wave front. The resulting lower bounds are a substantial improvement on those predicted by linear stability analysis, particularly in models where nonlinearities in the diffusion or in cross-species interactions play a decisive role in wave speed selection. In the context of cell invasion into ECM, these refined bounds allow for continued analytic progress in parameter regimes where numerical simulations have long indicated nonlinear selection but where no analytical estimates were previously available [69, 100]. The resulting bounds capture key qualitative trends, such as the increasing impact of nonlinear effects as the degradation rate increases, or the shift in the onset of nonlinear selection as the initial ECM density varies.

The examples considered in this chapter demonstrate that the method is flexible enough to incorporate distinct biological mechanisms, ranging from density-dependent mobility, to the Allee effect with a fixed wave speed, to moving-boundary problems such as the Fisher–Stefan model. In every case, the approach provides either an exact variational principle (for one-species equations) or an approximate but effective lower bound (for multi-species models), and these predictions can be easily determined using numerical methods. The extension to chemotactic systems further reveals that the general structure of the theory is robust: the inclusion of taxis terms modifies the effective flux but does not fundamentally alter the variational architecture.

Despite the breadth of these results, the method still has limitations, reducing its applicability. The core assumption throughout this chapter is the existence of a monotonic travelling wave with a fixed profile and constant speed, which allows the reduction of the PDE system to an ODE in the travelling wave co-ordinate, z, and the

subsequent recasting of the main species variable, u, as a spatial parameter. When the travelling wave does not have a constant speed or is non-monotonic, this reduction fails and the variational principle ceases to hold. Similarly, the two-species variational approximation relies on a weak-coupling or slowly varying second species coupling assumption such that the adjoint contributions in the optimal control framework can be considered negligible. As such, this ensures that the adjoint terms arising in the full optimality system may be treated as higher order corrections. In regimes where the coupling is strong, where reaction terms generate sharp internal layers, or where diffusion coefficients depend sensitively on the second species, this validity of the approximation may deteriorate. Relatedly, the lower bounds obtained for multi-species systems are not expected to be tight, and in certain parameter regions the approximation may only capture the correct qualitative dependence on parameters rather than providing an approximate prediction of the selected speed.

These limitations indicate natural directions for future research. The optimal control formulation provides a template that could, with further development, accommodate more general travelling waves or systems with multiple interacting controls. Instead of relying on weak coupling or slowly varying assumptions, this optimal control problem may be able to be solved explicitly, thus finding an exact wave speed prediction in the two-species case. For multi-species systems, the most promising route likely involves asymptotic analysis of the full optimality system in singular regimes, allowing the adjoint contributions to be incorporated rather than neglected. Finally, the structure uncovered here may prove compatible with data-driven approaches. Since the variational principle relates the travelling wave speed to explicit integral quantities, it potentially offers a means to benchmark or regularise machine-learned PDE models, particularly when the underlying biological mechanisms give rise to nonlinear speed selection [20, 46, 278, 287].

In summary, the optimal control interpretation presented in this chapter clarifies why the classical one-species variational principle works so effectively, and presents a systematic approach for extending these ideas beyond the one-species setting. In the context of biological applications, the framework developed here provides new analytic tools for understanding invasion dynamics in heterogeneous environments and for identifying precisely when nonlinear interactions dominate the spread. Although the method is not universally applicable, it significantly broadens the range of systems for which rigorous or semi-rigorous estimates of the travelling wave speed can be obtained, and it lays the groundwork for future analysis of multi-species models in mathematical biology.

Chapter 4

Phenotypic switching and its impact on cell migration under the 'go-or-grow' hypothesis

This work, entitled "Phenotypic switching mechanisms determine the structure of cell migration into extracellular matrix under the 'go-or-grow' hypothesis", was completed in collaboration with K. J. Painter, T. Lorenzi, P. K. Maini and R. E. Baker and is published in Mathematical Biosciences [79]. The author was responsible for all aspects of the research and writing presented in this chapter.

4.1 Introduction

Phenotypic heterogeneity profoundly impacts tumour behaviour and is a hallmark feature driving post-treatment recurrence [322]. Collective cell migration, which can be crucial in understanding various stages of tumour progression, often involves distinct cell phenotypes with varying motility and proliferative capacities [122, 123]. Mathematical approaches to studying these processes often involve models formulated as reaction–diffusion equations with phenotypic structuring [192, 194], where cells are assumed to undergo random, undirected movement and grow logistically to some maximum capacity, similar to the classical Fisher–KPP model [109, 174].

It is often observed experimentally that individual cells are either proliferative or motile [175], but not both [50], representing a trade-off known as the 'go-or-grow' hypothesis [173, 264]. Numerous mathematical models have been proposed to study the migration-proliferation dichotomy of a phenotypically structured population of cells, such as those that simulate glioblastoma growth [120]. Models of this nature consider various phenotypic switching mechanisms, derive analytical expressions for the

minimum travelling wave speed of the migrating front [82, 121, 318, 334], or include more complex nonlinear diffusion terms [70]. Crucially, however, cell phenotypes and their functions are fundamentally dependent on external cues, such as contact with neighbouring cells or interactions with the ECM [34, 73, 356]. Despite the increasing evidence for phenotypically structured cell populations, few studies have included the role of the ECM in models for cell migration under the go-or-grow hypothesis, to consider how it might affect the phenotypic structure of invading fronts and their speed. Furthermore, despite several experimental results supporting the existence of leader and follower cell sub-populations during collective cell migration [210, 377], other findings do not support such hypotheses [114, 347], possibly due to the use of different cell types or experimental conditions.

To investigate the role of the ECM in cell migration further, this work compares a model of a homogeneous population of generalist cells that can proliferate, move and degrade ECM (as introduced in Chapter 2) to a model of a heterogeneous population comprising two sub-populations of specialist cells that can either move and degrade ECM, or proliferate. This continuum model for the dynamics of two distinct cell phenotypes and the ECM is derived from first principles (as the limit of an underlying IBM) to accurately account for individual cell mechanisms at the population level. A number of different possible phenotypic switching mechanisms are considered, including random switching and both cell- and ECM- dependent switches, to explore their impact on the speed and phenotypic structure of invading cell populations.

In Section 4.2, we begin by describing the underlying assumptions of the models we study. We start by reviewing the details of the model for a population of homogeneous generalist cells in one spatial dimension that was first derived in Chapter 2. Subsequently, we extend this model to a heterogeneous population comprising two distinct sub-populations of specialist cells that can either move and degrade ECM or proliferate, in order to introduce phenotypic heterogeneity into the cell population (see Section 4.2.2). In Section 4.3, the solutions of these models are studied numerically for a variety of different phenotypic switching functions. We find that specialist cell populations with the ability to switch phenotype may invade faster than a generalist cell population, and that the choice of phenotypic switching mechanism drastically impacts the phenotypic structure of migrating cell fronts, such that the leading cell sub-population differs between switching functions.

4.2 The mathematical models

It is well-known that cells move in response to gradients in local cell volume fractions, and in response to nearby environmental features, such as the ECM, by haptotaxis, for example [4]. Many mathematical models have been developed that include nonlinear terms to describe these processes, and non-local reaction terms to describe the proliferation of cells to fill the surrounding available space [43, 119].

To invade into surrounding healthy tissues, many tumours must overcome physical barriers to migration, such as the ECM. In order to do this, tumour cells have developed the ability to remodel, reorient and degrade elements of the ECM [187, 361] through the production of specific matrix degrading enzymes, such as MMPs, that act in very close proximity to their cell of origin before decaying [59, 165, 319]. Since the timescale of ECM degradation is much longer than the timescale of intermediate processes, such as MMP decay, we employ the simplifying assumption that cells directly degrade the ECM [261]. In this chapter, we will investigate the simplest possible problem—that of the role of phenotypic heterogeneity in cell invasion into an ECM that is devoid of cells.

In this section, we present two deterministic, continuum models for cell migration into ECM that have been derived by coarse-graining their underlying IBMs to give rise to a corresponding population-level description (see schematic in Figure 4.1). The first model considers a homogeneous generalist cell population invading into ECM, as is often studied in standard models for collective cell migration [251], and builds on similar models for cell migration into ECM studied in [69, 100]. This model was presented in Chapter 2, but is recapitulated here in order to facilitate robust comparisons. Then, to investigate heterogeneity, we introduce a second model describing a population of specialist cells consisting of two distinct phenotypes, in line with evidence supporting the existence of separate proliferating and migrating populations [175]. This model extends those presented in [63, 264, 292] to include the ECM and its degradation by cells, as well as volume-filling effects.

4.2.1 An adapted model for homogeneous 'generalist' cell invasion into ECM

As seen in Chapter 2, previous models considered cell migration into ECM without heterogeneity or volume-filling assumptions [69, 100]; these models comprised two coupled differential equations with nonlinear cross-dependent diffusion and logistic growth. The differential equation model for generalist cells considered here describes

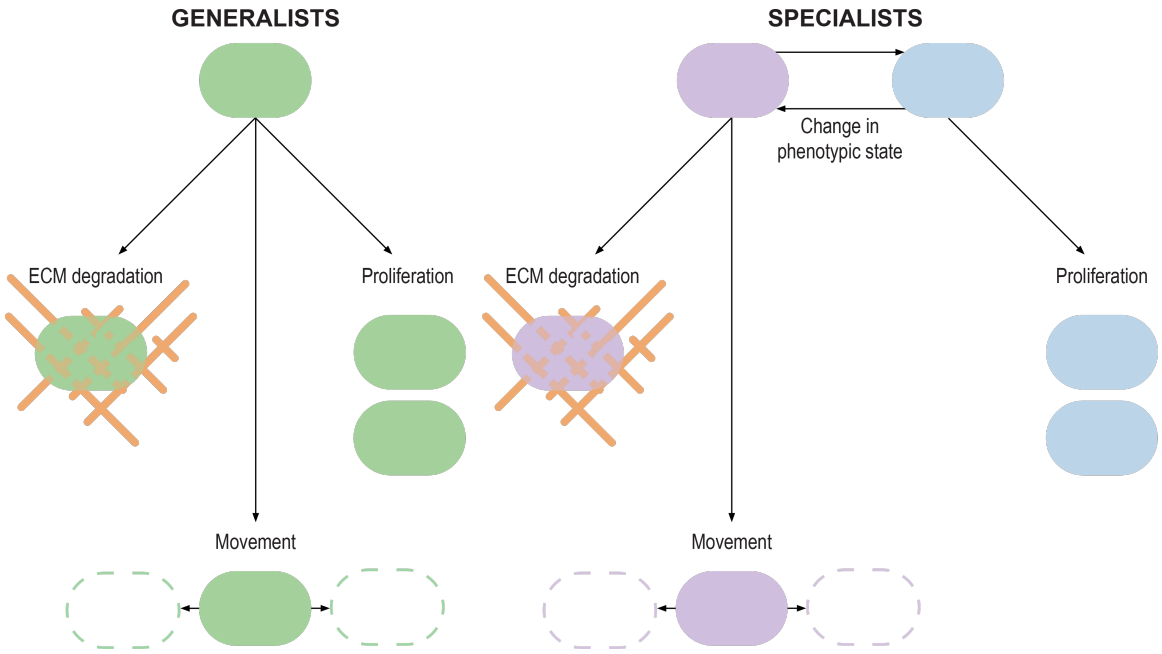

Figure 4.1: Schematic representation of the model for a homogeneous generalist population of cells (green), and the model for a heterogeneous specialist population of cells (purple: ECM degraders and moving cells, blue: proliferative cells).

the evolution of cell and ECM densities under volume-filling assumptions, where the movement and proliferation of cells is reduced in higher volume fraction regions. This model [77] was derived in Chapter 2 from an underlying one-dimensional, on-lattice, IBM, where its travelling wave solutions were studied.

To motivate later comparisons with a heterogeneous cell population, we re-introduce this model here with non-dimensional weightings of the cells towards proliferation, $\theta_{G,P} \in [0,1]$, degradation of ECM, $\theta_{G,D} \in [0,1]$, and movement, $(1-\theta_{G,D}-\theta_{G,P}) \in [0,1]$, that distribute a cells' weighting across different functions. The non-dimensional volume fractions of the generalist cell population and corresponding ECM are denoted as $u_G(x,t)$ and $m_G(x,t)$, respectively, and their dynamics are governed by the following system:

$$\frac{\partial u_G}{\partial t} = (1-\theta_{G,P}-\theta_{G,D})\frac{\partial}{\partial x}\left[(1-u_G-m_G)\frac{\partial u_G}{\partial x} + u_G\frac{\partial}{\partial x}(u_G+m_G)\right] + \theta_{G,P}u_G(1-u_G-m_G), \tag{4.1}$$

$$\frac{\partial m_G}{\partial t} = -\theta_{G,D}\lambda_G m_G u_G, \tag{4.2}$$

where $x \in \mathbb{R}$ and $t \geq 0$. The first term inside the square brackets on the right-hand side of Equation (4.1) models the undirected movement (*i.e.*, diffusion) of the

generalist cells, where this movement is prevented by the presence of other cells and ECM. The second term inside the square brackets models movement of the cells down the gradient of the total volume fraction of both cells and ECM, $u_G + m_G$, whereas the last (reaction) term describes proliferation of cells. The parameter $\lambda_G \in \mathbb{R}_+$ is the rescaled ECM degradation rate and when $\lambda_G = 0$ this model simplifies to a Fisher–KPP model with appropriately rescaled parameters [77]. We employ the following initial conditions:

$$u_G(x,0) = \begin{cases} 1, & \text{if} \quad x < \alpha, \\ 0, & \text{if} \quad x \geq \alpha, \end{cases} \tag{4.3}$$

$$m_G(x,0) = \begin{cases} 0, & \text{if} \quad x < \alpha, \\ m_0, & \text{if} \quad x \geq \alpha, \end{cases} \tag{4.4}$$

where $m_0 \in [0,1)$ corresponds to the volume fraction of ECM ahead of the cells, and $\alpha \in \mathbb{R}$ defines the edge of the region initially occupied by the cells. We complement this model with the following boundary conditions: u_G and $\partial u_G/\partial x \to 0$ as $x \to \infty$. In Chapter 2 and other previous works [69], we saw that under these conditions, travelling wave solutions can be observed, whose speeds depend on the initial volume fraction of ECM ahead of the invading wave of cells and the ECM degradation rate [69, 77]. Examples of these solutions, and their numerically estimated travelling wave speeds as the parameters $\theta_{G,D}$ and $\theta_{G,P}$ vary, are shown in Figure 4.2.

4.2.2 A model for heterogeneous cell invasion into ECM

In this section, we develop a simple one-dimensional model of two distinct cell phenotypes invading into ECM, in the presence of volume-filling effects, such that motility and proliferation of the cells is reduced in regions of space that are more highly occupied by other cells and ECM. Volume-filling effects warrant being accounted for in this way since research shows that cells and ECM both regulate cell movement and proliferation [92, 241, 354, 369]. This model derivation follows directly from the ideas in Chapter 2 extended to multiple cell sub-populations, and uses mean-field assumption ideas [257] from multi-species exclusion processes [311]. Following the go-or-grow assumption studied in [139, 318], we consider a phenotypic trade-off between cells that are degrading the ECM and migrating, and those that are proliferating, where those that proliferate do not move. We then coarse-grain this model to formally derive a corresponding PDE model that comprises a system of coupled differential equations for the densities of cells and ECM.

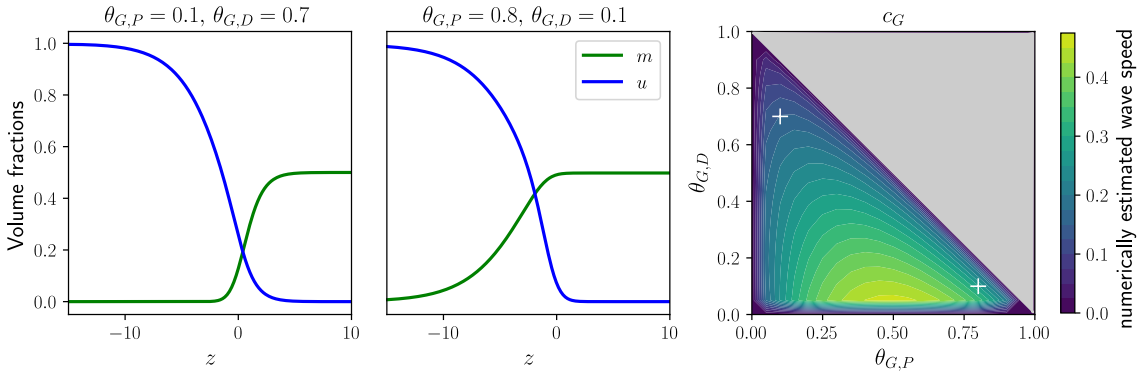

Figure 4.2: Plots of the travelling wave profile of the system (4.1) and (4.2) subject to the initial conditions for the cells as in Equation (4.3) and for the ECM as in Equation (4.4) for different values of $\theta_{G,D}$ and $\theta_{G,P}$ and translated into the travelling wave co-ordinate, $z = x - c_G t$, where c_G is the numerically estimated travelling wave speed. The far right shows the contour plot of the numerically observed travelling wave speeds as we vary $\theta_{G,D}$ and $\theta_{G,P}$. The initial ECM volume fraction ahead of the cells is $m_0 = 0.5$, the ECM degradation rate is $\lambda = 1$, and $\alpha = 1$ across all simulations. The parameter grid used for simulations in the contour plot was equally spaced at 0.05 intervals, forming a steep contour around the edge of the triangle where a travelling wave with positive speed quickly emerges when either of the parameters exceeds zero. For more information regarding the numerical methods used see Section 4.3.1.

4.2.2.1 Individual-based model

We begin with a one-dimensional, on-lattice IBM of two distinct cell phenotypes invading into ECM. In this model, cells are represented as discrete individuals with finite volume. We consider cells that have the ability to change phenotypic state, that move randomly on a one-dimensional uniform lattice, which constitutes the spatial domain, that proliferate, and that degrade the surrounding ECM, which is composed of discrete elements of the same volume as the cells.

Let the number of cells of phenotype $i = \{1, 2\}$ and ECM elements in lattice site $j = 1, \ldots, J$ of width Δ, at time $\tilde{t} \in \mathbb{R}_+$ of realisation $r = 1, \ldots, R$ of the model be denoted by $u^r_{i,j}(\tilde{t})$ and $m^r_j(\tilde{t})$, respectively.

Occupancy level of the lattice sites In order to incorporate volume-filling effects into the model, we prescribe that each lattice site has a maximum total occupancy level of N cells and ECM elements, so that

$$0 \leq \sum_{i=1}^{2} u^r_{i,j}(\tilde{t}) + m^r_j(\tilde{t}) \leq N.$$

Probability of cell movement A cell with phenotype i will attempt a movement event in a time step τ with probability $p_{i,\mathrm{m}} \in [0,1]$, whereby the attempted movement from lattice site j to either of the neighbouring lattice sites $j \pm 1$ occurs with equal probability 1/2. We assume that the probability of a successful move decreases linearly with the occupancy level of the target site, such that it is zero when the target site is full, and one when it is empty. Hence, we define the probability of a movement to the left, $T_{j-}^{i,\mathrm{m}^r}(\tilde{t})$, or right, $T_{j+}^{i,\mathrm{m}^r}(\tilde{t})$, in $[\tilde{t}, \tilde{t}+\tau)$ of realisation r as

$$T_{j\pm}^{i,\mathrm{m}^r}(\tilde{t}) = \frac{p_{i,\mathrm{m}}}{2}\left(1 - \frac{\sum_{i=1}^{2} u_{i,j\pm1}^r(\tilde{t}) + m_{j\pm1}^r(\tilde{t})}{N}\right).$$

Zero flux boundary conditions are implemented such that any attempted move outside of the spatial domain is aborted.

Probability of cell proliferation A cell with phenotype i in lattice site j attempts a proliferation event, placing a daughter cell of equal size into the same lattice site, during time step τ with probability $p_{i,\mathrm{p}} \in [0,1]$. We assume that the probability of a successful proliferation event, where one cell is replaced by two daughter cells with the same heritable phenotypic state as the parent cell, decreases linearly with the occupancy level of the lattice site, such that the probability of a successful proliferation event, $T_j^{i,\mathrm{p}^r}(\tilde{t})$, in time interval $[\tilde{t}, \tilde{t}+\tau)$ of realisation r of the model is

$$T_j^{i,\mathrm{p}^r}(\tilde{t}) = p_{i,\mathrm{p}}\left(1 - \frac{\sum_{i=1}^{2} u_{i,j}^r(\tilde{t}) + m_j^r(\tilde{t})}{N}\right).$$

Probability of cell phenotype change During time interval $[\tilde{t}, \tilde{t}+\tau)$ of realisation r of the model, a cell with phenotype i will change to phenotype l $(i \neq l)$ with probability

$$f_{i\to l}\big(u_{1,j}^r(\tilde{t}),\, u_{2,j}^r(\tilde{t}),\, m_j^r(\tilde{t})\big),$$

where $f_{i\to l} : \mathbb{R}_+^3 \to [0,1]$.

Probability of ECM degradation During the time interval $[\tilde{t}, \tilde{t}+\tau)$ of realisation r, an element of ECM in lattice site j is degraded by cells with phenotype i in the same lattice site with a probability $p_{i,\mathrm{d}} \in [0,1]$. Therefore the overall degradation rate per unit element of ECM by cells with phenotype i, $T_j^{i,\mathrm{d}^r}(\tilde{t})$, is

$$T_j^{i,\mathrm{d}^r}(\tilde{t}) = p_{i,\mathrm{d}} u_{i,j}^r(\tilde{t}).$$

4.2.2.2 Coarse-grained model

To derive a coarse-grained description of the IBM, we introduce the average occupancy of each lattice site j by cells of type $i = \{1, 2\}$ and ECM at time $\tilde{t}$ over R total realisations of the model as

$$\langle u_{i,j}(\tilde{t})\rangle = \frac{1}{R}\sum_{r=1}^{R} u_{i,j}^{r}(\tilde{t}) \;\text{ and }\; \langle m_j(\tilde{t})\rangle = \frac{1}{R}\sum_{r=1}^{R} m_j^{r}(\tilde{t}).$$

Coarse-grained model of cell dynamics We write a conservation equation using mean-field approximations and independence of lattice sites by considering changes in the average occupancy in the lattice site j during the time interval $[\tilde{t}, \tilde{t}+\tau)$ to give:

$$\begin{aligned}
\langle u_{i,j}(\tilde{t}+\tau)\rangle = \langle u_{i,j}(\tilde{t})\rangle &+ \frac{p_{i,\mathrm{m}}}{2}\langle u_{i,j+1}(\tilde{t})\rangle\left(1 - \frac{\sum_{i=1}^{2}\langle u_{i,j}(\tilde{t})\rangle + \langle m_j(\tilde{t})\rangle}{N}\right)\\
&+ \frac{p_{i,\mathrm{m}}}{2}\langle u_{i,j-1}(\tilde{t})\rangle\left(1 - \frac{\sum_{i=1}^{2}\langle u_{i,j}(\tilde{t})\rangle + \langle m_j(\tilde{t})\rangle}{N}\right)\\
&- \frac{p_{i,\mathrm{m}}}{2}\langle u_{i,j}(\tilde{t})\rangle\left(1 - \frac{\sum_{i=1}^{2}\langle u_{i,j+1}(\tilde{t})\rangle + \langle m_{j+1}(\tilde{t})\rangle}{N}\right)\\
&- \frac{p_{i,\mathrm{m}}}{2}\langle u_{i,j}(\tilde{t})\rangle\left(1 - \frac{\sum_{i=1}^{2}\langle u_{i,j-1}(\tilde{t})\rangle + \langle m_{j-1}(\tilde{t})\rangle}{N}\right)\\
&+ p_{i,\mathrm{p}}\langle u_{i,j}(\tilde{t})\rangle\left(1 - \frac{\sum_{i=1}^{2}\langle u_{i,j}(\tilde{t})\rangle + \langle m_j(\tilde{t})\rangle}{N}\right)\\
&- f_{i\to l}\big(\langle u_{1,j}(\tilde{t})\rangle,\, \langle u_{2,j}(\tilde{t})\rangle,\, \langle m_j(\tilde{t})\rangle\big)\langle u_{i,j}(\tilde{t})\rangle\\
&+ f_{l\to i}\big(\langle u_{1,j}(\tilde{t})\rangle,\, \langle u_{2,j}(\tilde{t})\rangle,\, \langle m_j(\tilde{t})\rangle\big)\langle u_{l,j}(\tilde{t})\rangle. \qquad (4.5)
\end{aligned}$$

Rearranging Equation (4.5), and dividing by τ, we find

$$\begin{aligned}
&\frac{\langle u_{i,j}(\tilde{t}+\tau)\rangle - \langle u_{i,j}(\tilde{t})\rangle}{\tau} \\
&= \frac{p_{i,\mathrm{m}}\Delta^2}{2\tau}\left(1 - \frac{\sum_{i=1}^{2}\langle u_{i,j}(\tilde{t})\rangle + \langle m_j(\tilde{t})\rangle}{N}\right)\left[\frac{\langle u_{i,j+1}(\tilde{t})\rangle - 2\langle u_{i,j}(\tilde{t})\rangle + \langle u_{i,j-1}(\tilde{t})\rangle}{\Delta^2}\right] \\
&\quad + \frac{p_{i,\mathrm{m}}\Delta^2}{2\tau}\langle u_{i,j}(\tilde{t})\rangle \times \frac{1}{\Delta^2}\Big[\Big(\sum_{i=1}^{2}\langle u_{i,j+1}(\tilde{t})\rangle + \langle m_{j+1}(\tilde{t})\rangle\Big) \\
&\qquad\qquad - 2\Big(\sum_{i=1}^{2}\langle u_{i,j}(\tilde{t})\rangle + \langle m_j(\tilde{t})\rangle\Big) \\
&\qquad\qquad + \Big(\sum_{i=1}^{2}\langle u_{i,j-1}(\tilde{t})\rangle + \langle m_{j-1}(\tilde{t})\rangle\Big)\Big] \\
&\quad + \frac{p_{i,\mathrm{p}}}{\tau}\langle u_{i,j}(\tilde{t})\rangle\left(1 - \frac{\sum_{i=1}^{2}\langle u_{i,j}(\tilde{t})\rangle + \langle m_j(\tilde{t})\rangle}{N}\right) \\
&\quad - \frac{1}{\tau}f_{i\to l}\big(\langle u_{1,j}(\tilde{t})\rangle,\, \langle u_{2,j}(\tilde{t})\rangle,\, \langle m_j(\tilde{t})\rangle\big)\langle u_{i,j}(\tilde{t})\rangle \\
&\quad + \frac{1}{\tau}f_{l\to i}\big(\langle u_{1,j}(\tilde{t})\rangle,\, \langle u_{2,j}(\tilde{t})\rangle,\, \langle m_j(\tilde{t})\rangle\big)\langle u_{l,j}(\tilde{t})\rangle. \qquad (4.6)
\end{aligned}$$

Dividing Equation (4.6) by Δ and Taylor expanding both sides, before taking limits as Δ, $\tau \to 0$, we obtain a description for cell density dynamics in terms of variables $\tilde{u}_i(\tilde{x},\tilde{t})$ and $\tilde{m}(\tilde{x},\tilde{t})$, which are the continuum equivalents of the number density of cells, $\langle u_{i,j}(\tilde{t})\rangle/\Delta$, and the density of ECM, $\langle m_j(\tilde{t})\rangle/(\tilde{\mu}\Delta)$, at $\tilde{x} \in \mathbb{R}$, $\tilde{t} \in \mathbb{R}_+$, respectively, where $\tilde{\mu}$ represents the number of cells equivalent to a unit mass of ECM and serves as a conversion factor between the density of ECM, as defined by mass of ECM per unit volume, and the number density of ECM elements, given by $\tilde{\mu}\tilde{m}(\tilde{x},\tilde{t})$. Under the following scalings:

$$\begin{aligned}
&\lim_{\Delta\to 0}\frac{N}{\Delta} = \tilde{K}, \qquad \lim_{\tau\to 0}\frac{p_{i,\mathrm{p}}}{\tau} = \tilde{r_i}, \qquad \lim_{\Delta,\tau\to 0}\frac{p_{i,\mathrm{m}}\Delta^2}{\tau} = \tilde{D}_i, \\
&\lim_{\tau\to 0}\frac{1}{\tau}f_{i\to l}\big(\langle u_{1,j}(\tilde{t})\rangle,\, \langle u_{2,j}(\tilde{t})\rangle,\, \langle m_j(\tilde{t})\rangle\big) = \tilde{\gamma_{il}}(\tilde{u_1},\, \tilde{u_2},\, \tilde{m}),
\end{aligned}$$

for all i, $l = \{1, 2\}$, $i \neq l$, we find

$$\begin{aligned}\frac{\partial \tilde{u}_i}{\partial \tilde{t}} &= \tilde{D}_i \left[\left(1 - \frac{\sum_{i=1}^{2} \tilde{u}_i + \tilde{\mu}\tilde{m}}{\tilde{K}} \right) \frac{\partial^2 \tilde{u}_i}{\partial \tilde{x}^2} + \tilde{u}_i \frac{\partial^2}{\partial \tilde{x}^2} \left(\frac{\sum_{i=1}^{2} \tilde{u}_i + \tilde{\mu}\tilde{m}}{\tilde{K}} \right) \right] \\ &\quad + \tilde{r}_i \tilde{u}_i \left(1 - \frac{\sum_{i=1}^{2} \tilde{u}_i + \tilde{\mu}\tilde{m}}{\tilde{K}} \right) - \tilde{\gamma}_{il}(\tilde{u}_1,\, \tilde{u}_2,\, \tilde{m})\tilde{u}_i + \tilde{\gamma}_{li}(\tilde{u}_1,\, \tilde{u}_2,\, \tilde{m})\tilde{u}_l, \\ &= \tilde{D}_i \frac{\partial}{\partial \tilde{x}} \left[\left(1 - \frac{\sum_{i=1}^{2} \tilde{u}_i + \tilde{\mu}\tilde{m}}{\tilde{K}} \right) \frac{\partial \tilde{u}_i}{\partial \tilde{x}} + \tilde{u}_i \frac{\partial}{\partial \tilde{x}} \left(\frac{\sum_{i=1}^{2} \tilde{u}_i + \tilde{\mu}\tilde{m}}{\tilde{K}} \right) \right] \\ &\quad + \tilde{r}_i \tilde{u}_i \left(1 - \frac{\sum_{i=1}^{2} \tilde{u}_i + \tilde{\mu}\tilde{m}}{\tilde{K}} \right) - \tilde{\gamma}_{il}(\tilde{u}_1,\, \tilde{u}_2,\, \tilde{m})\tilde{u}_i + \tilde{\gamma}_{li}(\tilde{u}_1,\, \tilde{u}_2,\, \tilde{m})\tilde{u}_l,\end{aligned}$$

where $\tilde{x} \in \mathbb{R}$ and $\tilde{t} \in \mathbb{R}_+$.

Coarse-grained model of ECM dynamics In the same way, using probabilistic assumptions of mean-field type, we can write the following conservation equation for the evolution of ECM elements in a lattice site j over the time interval $[\tilde{t}, \tilde{t} + \tau)$:

$$\langle m_j(\tilde{t} + \tau) \rangle = \langle m_j(\tilde{t}) \rangle - \sum_{i=1}^{2} p_{i,\mathrm{d}} \langle u_{i,j}(\tilde{t}) \rangle \langle m_j(\tilde{t}) \rangle. \tag{4.7}$$

By rearranging Equation (4.7), dividing by Δ and τ and taking limits as Δ, $\tau \to 0$ under the scaling

$$\tilde{\lambda}_i = \lim_{\Delta, \tau \to 0} \frac{\Delta p_{i,\mathrm{d}}}{\tau},$$

we obtain the following differential equation governing the dynamics of ECM density $\tilde{m}(\tilde{x}, \tilde{t})$ over time:

$$\frac{\partial \tilde{m}}{\partial \tilde{t}} = -\sum_{i=1}^{2} \tilde{\lambda}_i \tilde{m} \tilde{u}_i,$$

where $\tilde{x} \in \mathbb{R}$ and $\tilde{t} \in \mathbb{R}_+$. The parameter $\tilde{\lambda}_i \geq 0$ describes the degradation rate of ECM per cell in phenotypic state i.

Full system of equations For phenotypes $i = \{1, 2\}$, $i \neq l$, the full system of equations at population level is given by

$$\begin{aligned}\frac{\partial \tilde{u}_i}{\partial \tilde{t}} &= \tilde{D}_i \frac{\partial}{\partial \tilde{x}} \left[\left(1 - \frac{\sum_{i=1}^{2} \tilde{u}_i + \tilde{\mu}\tilde{m}}{\tilde{K}} \right) \frac{\partial \tilde{u}_i}{\partial \tilde{x}} + \tilde{u}_i \frac{\partial}{\partial \tilde{x}} \left(\frac{\sum_{i=1}^{2} \tilde{u}_i + \tilde{\mu}\tilde{m}}{\tilde{K}} \right) \right] \\ &\quad + \tilde{r}_i \tilde{u}_i \left(1 - \frac{\sum_{i=1}^{2} \tilde{u}_i + \tilde{\mu}\tilde{m}}{\tilde{K}} \right) - \tilde{\gamma}_{il}(\tilde{u}_1,\, \tilde{u}_2,\, \tilde{m})\tilde{u}_i + \tilde{\gamma}_{li}(\tilde{u}_1,\, \tilde{u}_2,\, \tilde{m})\tilde{u}_l, \\ \frac{\partial \tilde{m}}{\partial \tilde{t}} &= -\sum_{i=1}^{2} \tilde{\lambda}_i \tilde{m} \tilde{u}_i.\end{aligned}$$

4.2.2.3 Two populations of 'specialist' cells

Now consider the case where cells in phenotypic state 1 can move and degrade ECM only, and cells in phenotypic state 2 are only able to proliferate, following the go-or-grow hypothesis [139]. We write $\tilde{\lambda_i}$ as the rate of ECM degradation and $\tilde{r_i}$ as the proliferation rate of cells in phenotypic state i and assume

$$\tilde{\lambda_2} = 0, \quad \tilde{r_1} = 0, \quad \tilde{D}_2 = 0, \quad \tilde{\lambda_1} \in \mathbb{R}_+ \quad \text{and} \quad \tilde{r_2} \in \mathbb{R}_+.$$

Without loss of generality, we assume that cells have a total weighting to distribute across the mechanisms governing the cells' migration, irrelevant of their phenotypic state. As such, we introduce $\theta_{S,D} \in [0,1]$ to describe the weighting of cells in phenotypic state 1 towards degrading ECM, and $(1-\theta_{S,D})$ describes the remaining weighting for movement.

Under the aforementioned assumptions, the following system of differential equations describes the evolution of cell and ECM densities over time:

$$\begin{aligned}\frac{\partial \tilde{u_1}}{\partial \tilde{t}} &= (1-\theta_{S,D})\tilde{D}_1 \frac{\partial}{\partial \tilde{x}}\left[\left(1 - \frac{\sum_{i=1}^{2}\tilde{u}_i + \tilde{\mu}\tilde{m}}{K}\right)\frac{\partial \tilde{u}_1}{\partial \tilde{x}} + \tilde{u}_1 \frac{\partial}{\partial \tilde{x}}\left(\frac{\sum_{i=1}^{2}\tilde{u}_i + \tilde{\mu}\tilde{m}}{K}\right)\right] \\ &\quad + \tilde{u}_2\tilde{\gamma}_{21}(\tilde{u}_1, \tilde{u}_2, \tilde{m}) - \tilde{u}_1\tilde{\gamma}_{12}(\tilde{u}_1, \tilde{u}_2, \tilde{m}), \qquad (4.8) \\ \frac{\partial \tilde{u_2}}{\partial \tilde{t}} &= \tilde{r_2}\tilde{u}_2\left(1 - \frac{\sum_{i=1}^{2}\tilde{u}_i + \tilde{\mu}\tilde{m}}{K}\right) \\ &\quad - \tilde{u}_2\tilde{\gamma}_{21}(\tilde{u}_1, \tilde{u}_2, \tilde{m}) + \tilde{u}_1\tilde{\gamma}_{12}(\tilde{u}_1, \tilde{u}_2, \tilde{m}), \qquad (4.9) \\ \frac{\partial \tilde{m}}{\partial \tilde{t}} &= -\theta_{S,D}\tilde{\lambda_1}\tilde{m}\tilde{u_1}, \qquad (4.10)\end{aligned}$$

where the diffusion coefficient of cells of type 1 is given by $\tilde{D}_1 \in \mathbb{R}_+$, the switching between cell phenotypic states $p \in \{1,2\}$ is given by the functions

$$\tilde{\gamma}_{12} : \mathbb{R}_+^3 \to \mathbb{R}_+ \qquad \text{and} \qquad \tilde{\gamma}_{21} : \mathbb{R}_+^3 \to \mathbb{R}_+,$$

where $\tilde{\gamma}_{12}$ represents the rate of switching from phenotypic state 1 to 2, and $\tilde{\gamma}_{21}$ represents the rate of switching from phenotypic state 2 to 1, $\theta_{S,D} \in [0,1]$ is the weighting of cells in phenotypic state 1 to degrade ECM, and $\tilde{\lambda_1} \in \mathbb{R}$ is the ECM degradation rate by cells in phenotypic state 1 and $\tilde{x} \in \mathbb{R}$.

4.2.2.4 Non-dimensionalised model

Without loss of generality, we introduce the following non-dimensional variables:

$$x = \sqrt{\frac{\tilde{r_2}}{\tilde{D}_1}}\tilde{x}, \quad t = \tilde{r_2}\tilde{t}, \quad u_1 = \frac{\tilde{u_1}}{K}, \quad u_2 = \frac{\tilde{u_2}}{K}, \quad m = \frac{\mu\tilde{m}}{K},$$

alongside the following non-dimensional form of the switching functions

$$\gamma_{12} = \gamma_{12}(u_1, u_2, m) = \frac{1}{\tilde{r_2}}\tilde{\gamma_{12}}(\tilde{u_1}, \tilde{u_2}, \tilde{m}),$$
$$\gamma_{21} = \gamma_{21}(u_1, u_2, m) = \frac{1}{\tilde{r_2}}\tilde{\gamma_{21}}(\tilde{u_1}, \tilde{u_2}, \tilde{m}),$$

which, substituting into Equations (4.8)–(4.10), yields the non-dimensional system

$$\begin{aligned}
\frac{\partial u_1}{\partial t} &= (1-\theta_{S,D})\frac{\partial}{\partial x}\left[\big(1-u_1-u_2-m\big)\frac{\partial u_1}{\partial x} + u_1\frac{\partial}{\partial x}\big(u_1+u_2+m\big)\right] \\
&\qquad + u_2\gamma_{21}(u_1,u_2,m) - u_1\gamma_{12}(u_1,u_2,m), \\
\frac{\partial u_2}{\partial t} &= u_2\big(1-u_1-u_2-m\big) \qquad (4.11)\\
&\qquad - u_2\gamma_{21}(u_1,u_2,m) + u_1\gamma_{12}(u_1,u_2,m), \qquad (4.12)\\
\frac{\partial m}{\partial t} &= -\theta_{S,D}\lambda m u_1, \qquad (4.13)
\end{aligned}$$

where we have introduced the non-dimensional parameter

$$\lambda = \frac{\tilde{\lambda_1}K}{\tilde{r_2}},$$

representing the rescaled ECM degradation rate, noting that $\theta_{S,D} \in [0,1]$ is already a dimensionless parameter.

Similar to the generalist model, the first term inside the square brackets in Equation (4.11) captures undirected movement (*i.e.*, diffusion) of the cells in phenotypic state 1, inhibited by the presence of other cells and ECM, and the second term inside the square brackets describes the movement of cells in phenotypic state 1 down the gradient in the total volume fraction of cells and ECM. The first term on the right-hand side of Equation (4.12) describes the growth of cells in phenotypic state 2, as limited by the presence of other cells and ECM. The single term on the right-hand side of Equation (4.13) describes degradation of ECM by cells in phenotypic state 1. We consider a number of phenotypic switching functions that incorporate different types of behaviour with a view to understanding their impact on the speed of cell invasion as well as on the structure of the invading wave.

Phenotypic switching functions The phenotypic switching functions we consider, $\gamma_{12}(u_1,u_2,m)$ and $\gamma_{21}(u_1,u_2,m)$, are listed in Table 4.1. The first is constant switching, at rate $s \in \mathbb{R}_+$, between the two sub-populations. It is important to note that when $s = 0$ there is no switching between the phenotypic states; in this case the

model does not permit travelling wave solutions and invasion is not observed. The second, ECM-dependent phenotypic switching, entails cells switching from phenotypic state 1 (2) to phenotypic state 2 (1) at a rate that decreases (increases) linearly with ECM volume fraction, and describes a higher rate of switching to the ECM degrading phenotypic state in regions of higher ECM volume fractions. The third, space-dependent phenotypic switching, is defined such that the rate of switching from phenotypic state 1 (2) to phenotypic state 2 (1) increases (decreases) with the available space, $1 - u_1 - u_2 - m$. Finally, cell-dependent phenotypic switching assumes that only the total cell volume fraction impacts switching, with the ECM playing no direct role in driving phenotypic switching. Note that, since $0 \leq u_1 + u_2 + m \leq 1$, all the switching functions given in Table 4.1 are non-negative.

Name	$\gamma_{12}(u_1, u_2, m)$	$\gamma_{21}(u_1, u_2, m)$
Constant	s	s
ECM-dependent	$s(1-m)$	sm
Space-dependent	$s(1-u_1-u_2-m)$	$s(u_1+u_2+m)$
Cell-dependent	$s(1-u_1-u_2)$	$s(u_1+u_2)$

Table 4.1: Table listing the phenotypic switching functions considered in this work.

In order to satisfy the boundary conditions of the coarse-grained continuum model, we implement that $u_1 \to 0$ and $\partial u_1/\partial x \to 0$ as $x \to \infty$. For consistency with the volume-filling assumptions of the model, we also initially assume that

$$u_i(x,0) = \begin{cases} 0.5, & \text{if} \quad x < \alpha, \\ 0, & \text{if} \quad x \geq \alpha, \end{cases} \tag{4.14}$$

for $i \in \{1, 2\}$. The initial conditions for the ECM volume fraction, $m(x, 0)$, are prescribed by Equation (4.4), where α and m_0 have the same interpretation as in Section 4.2.1. The spatially homogeneous steady states of the system are found in Appendix A.1. Furthermore, we note here that the system (4.11)–(4.13) and the system (4.1) and (4.2) are solved numerically on the domain $x \in [0, L]$, where L is chosen differently between simulations to be sufficiently large such that convergence in the travelling wave speed is observed, and the boundary conditions do not have an effect (see Section 4.3.1). For more information on the method for the numerical calculation of the travelling wave speed, c, for each simulation, see Section 4.3.1, and note that the solutions in Figures 4.3, 4.5 and 4.7 are translated into the travelling wave co-ordinate $z = x - ct$.

4.3 Numerical exploration of invasion dynamics

In this section, we use numerical simulations to investigate how the models introduced in Section 4.2 give rise to travelling wave solutions and how the speed and structure of these waves depend on phenotypic heterogeneity and environmental context. We begin by outlining the numerical methods employed to solve the system and to estimate invasion speeds (Section 4.3.1). We then compare the migration behaviour of homogeneous generalist populations with that of heterogeneous specialist populations under constant phenotypic switching. Subsequently, we extend the analysis to consider environmentally-dependent switching mechanisms, examining how cues from the ECM, spatial availability, or local cell density influence the organisation of invading fronts. Finally, we explore the impact of key biological and environmental parameters on the invasion speed, identifying the regimes in which phenotypic heterogeneity enhances or hinders migration.

4.3.1 Numerical methods

The system (4.11)–(4.13) subject to the boundary conditions and initial conditions for the cells as in Equation (4.14), and for the ECM as in Equation (4.4), are solved numerically using the method of lines on a one-dimensional spatial domain $x \in [0, L]$, where $L > 200$ is chosen to be sufficiently large to remove the impact of the boundary conditions and enable convergence to a constant speed, constant profile travelling wave solution.

To employ the method of lines, the spatial domain is uniformly discretised into Q spatial points, with separation h. An explicit central differencing scheme, as described in [321], is then employed to solve the system, taking the following form:

$$\frac{\partial}{\partial x}\left[D\frac{\partial a}{\partial x}\right] \approx \frac{1}{2h^2}\Big[(D_{q-1}+D_q)a_{q-1} + (D_i + D_{q+1})a_{q+1} - (D_{q-1}+2D_q+D_{q+1})a_q\Big],$$

where a_q represents the value of the function a at the spatial point x_q. The system (4.11)–(4.13) can then be re-written as a system of $3Q$ ODEs, solved by simulating ghost points x_{-1} and x_{Q+1} outside of the initial spatial domain, as described in [221]. The remaining system of equations, which has been discretised in space, is then solved numerically in python using the built-in solver `scipy.integrate.solve_ivp` with the explicit Runge-Kutta integration method of order 5 and time step $\Delta t = 0.1$.

To estimate the wave speeds numerically, for each time point that we save a solution, we interpolate to find $X(t)$ such that

$$u_1(X(t), t) + u_2(X(t), t) = y^* \in (0, 1),$$

where we choose $y^* = 0.1$ and then calculate

$$c_{\text{estimated}}(t, t + \Delta t) = \frac{X(t + \Delta t) - X(t)}{\Delta t}.$$

When the calculated wave speeds are observed to have converged to a constant speed, such that the difference between two subsequent measurements is of an order smaller than the order of error of the numerical scheme, we record this as the travelling wave speed estimated numerically.

4.3.2 Comparing the migration speed of the generalist and specialist population with constant phenotypic switching

In reality, many different cell types are known to co-operate to create robust migration, often through the emergence of leader and follower cell phenotypes. For example, in neural crest cell migration, the leader and follower phenotypes generate streams of invading cells [204, 209, 210], whilst during angiogenesis, tip and stalk cells aid in the branching process [52]. To investigate whether robust cell invasion can be observed in a phenotypically heterogeneous specialist population, we explore the dynamics of Equations (4.11)–(4.13), first in the case of constant switching, at rate $s > 0$.

In this case, travelling wave solutions can be observed in the ECM and cell volume fractions, where the two sub-populations of cell types are well mixed along the invading front (see the top row of Figure 4.3). The bottom row of Figure 4.3 shows the difference between the numerically estimated travelling wave speed of a homogeneous generalist population, c_G, and a heterogeneous specialist population, c_S, such that green indicates regions of parameter space where the travelling wave speed of the heterogeneous cell population exceeds that of the homogeneous counterpart. The numerically estimated travelling wave speeds c_G and c_S are computed as described in Section 4.3.1. The maximum observed speed from either cell population in these model simulations was 0.5, when $\theta_{G,P} = 0.5$, $\theta_{G,D} = 0$ and $m_0 < 1$, and thus the differences between the travelling wave speeds of the homogeneous and heterogeneous populations are observed to be of the same order of magnitude as the numerically estimated travelling wave speeds. We find that specialist cell populations with constant, phenotypic switching have a faster travelling wave speed in all cases except when generalists heavily weight their abilities towards cell motility, rather than ECM degradation. However, it is important to note that the maximum possible travelling wave speed for the two models is the same, as explained later in Section 4.4.3.

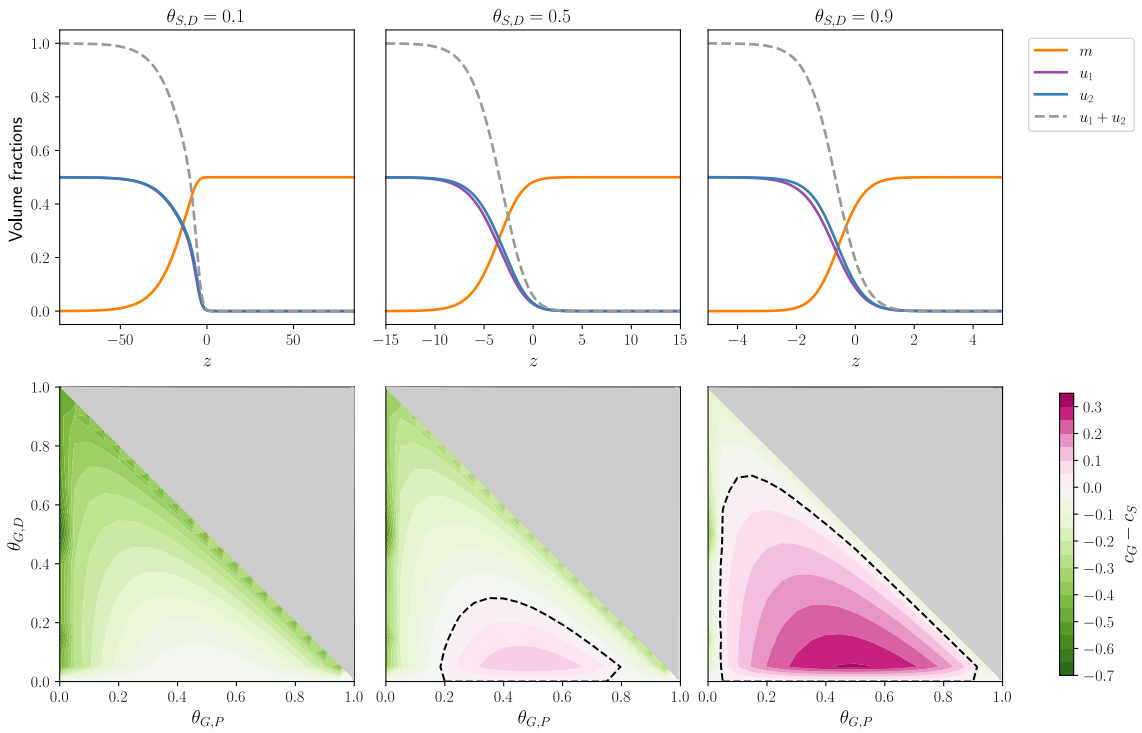

Figure 4.3: The top row shows plots of the travelling wave profile of the system of Equations (4.11)–(4.13) subject to the initial conditions for the cells as in Equation (4.14) and for the ECM as in Equation (4.4) for different values of $\theta_{S,D}$ for the case of constant switching (see Table 4.1) and translated into the travelling wave co-ordinate, $z = x - c_S t$, where c_S is the numerically estimated travelling wave speed of the specialised, heterogeneous population. The bottom row shows contour plots displaying the difference between the numerically observed travelling wave speed of the homogeneous generalist population, c_G, as given by system (4.1) and (4.2) subject to the initial conditions (4.3) and (4.4), and the speed of the heterogeneous specialist population, c_S, as defined by the system (4.11)–(4.13) subject to the initial conditions for the cells as in Equation (4.14) and for the ECM as in Equation (4.4). The regions coloured in pink display the parameter regimes where the difference between the travelling wave speeds, $c_G - c_S$, is positive, meaning generalist cells invade faster than specialists, and regions are coloured in green when this difference is negative (*i.e.*, $c_S > c_G$). The dashed black line is plotted at $c_G = c_S$. The initial ECM volume fraction ahead of the cells is $m_0 = 0.5$, the ECM degradation rate is $\lambda = 1$, the switching rate is $s = 1$, and $\alpha = 1$ across all simulations. For more information regarding the numerical methods used see Section 4.3.1.

4.3.3 The impact of environmentally-dependent phenotypic switching functions on migrating fronts

In reality, cells are able to sense their environment and neighbouring cells, which can both provide cues for directed migration. Variations in the surrounding cells and

environment can also cause phenotypic changes within cells that affect their behaviour [89] and thus we extend our study of the heterogeneous specialist population of cells to consider the impact of ECM-, space- and cell-dependent phenotypic switching functions, as defined in Table 4.1.

ECM-dependent switching We first consider the scenario where cells are able to sense the volume fraction of surrounding ECM which then influences the rate of phenotypic switching. In this case, we find that spatial heterogeneity appears within the travelling wave front (see Figure 4.4). Instead of a well-mixed population of cells that degrade and proliferate throughout the invading wave (as is observed for constant phenotypic switching between heterogeneous specialist populations, or for homogeneous generalist populations) we find that the cells in phenotypic state 1 (*i.e.*, the ECM degraders) concentrate at the wave front in the form of a travelling pulse, whereas the bulk of the invading wave is filled by a travelling front of proliferative cells in phenotypic state 2.

Space-dependent switching In line with the volume-filling principles underlying this model (*i.e.*, the fact that cells are unable to move or proliferate in a region that has no available space), we introduce phenotypic switching from phenotypic state 1 (ECM degrader) to phenotypic state 2 (proliferator) at an increasing rate as available space increases (see Table 4.1). By inspecting Figure 4.4 it is clear that the bulk of the travelling wave consists of cells in phenotypic state 1, whereas cells in phenotypic state 2 concentrate at the migrating front, which is opposite to what is observed for ECM-dependent switching. As a result, there is increased ECM degradation due to a larger proportion of cells in phenotypic state 1, and we see a sharper transition between $m = 0$ and $m = m_0$ in the travelling wave as compared to constant, or ECM-dependent, switching (see Figure 4.4).

Cell-dependent switching When cells change phenotypic state according to the cell-dependent phenotypic switching function defined in Table 4.1, a qualitatively similar cell distribution is observed as in the space-dependent switching case, with proliferating cells (phenotypic state 2) at the migrating front and ECM-degrading cells (phenotypic state 1) in the bulk. Subtle differences between the travelling wave profile for ECM- and space-dependent switching can be observed in Figure 4.4, including a higher maximum volume fraction of proliferating cells and a steeper travelling

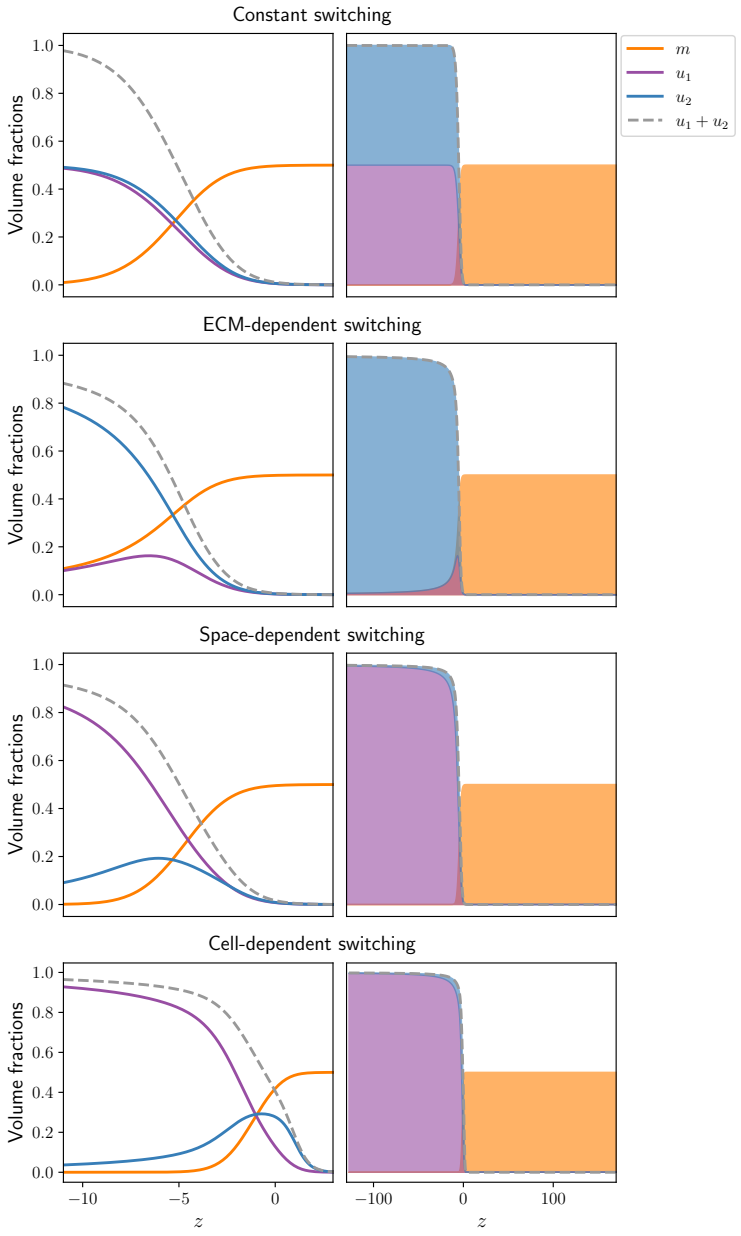

Figure 4.4: Plots showing the proportion of cells in each phenotypic state and their position in the travelling wave when simulating the system (4.11)–(4.13) subject to the initial conditions for the cells as in Equation (4.14) and for the ECM as in Equation (4.4) and translated into the travelling wave co-ordinate, $z = x - ct$, where c is the numerically estimated travelling wave speed, for each of the phenotypic switching functions in Table 4.1. In the first column, we show zoomed in profiles at the front of the travelling wave. In the second column, we show the full travelling wave profile, with each constituent shaded. In all plots, the initial ECM volume fraction ahead of the cells is $m_0 = 0.5$ and the ECM degradation rate is $\lambda = 1$. The width of the region initially invaded by cells is $\alpha = 1$, the weighting of specialists towards degradation is $\theta_{S,D} = 0.5$ and the switching rate for all functions is $s = 1$. For more information regarding the numerical methods used see Section 4.3.1.

wave front in both ECM and total cell volume fractions.

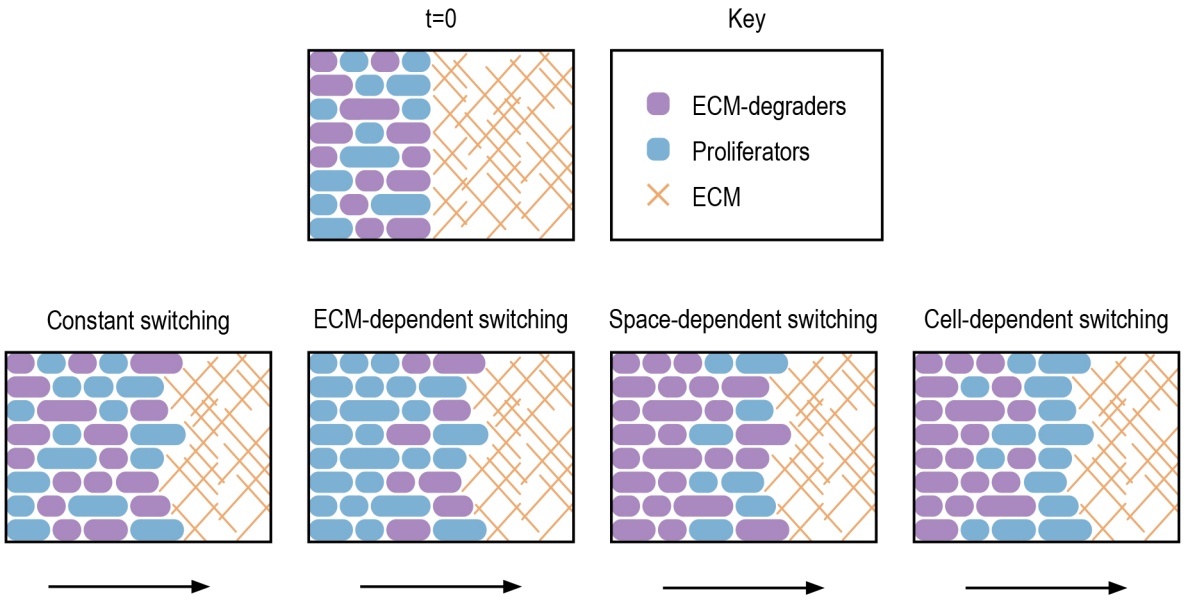

Figure 4.5: Schematics demonstrating the possible cell distributions in the travelling waves observed in this numerical study. Note that there are subtle differences between space- and cell-dependent switching, namely that proliferating cells lead the wave of invasion in isolation during cell-dependent switching.

Overall, it is clear that constant speed travelling wave solutions can be observed for the system (4.11)–(4.13) subject to any of the switching functions described in Table 4.1. Furthermore, the functional form of phenotypic switching mechanism chosen influences the distribution of cell phenotypic states within the travelling wave, as schematised in Figure 4.5.

4.3.4 Exploring the influence of model parameters on migration speed

We now analyse how variations in the model parameters impact the numerically observed travelling wave speed, and determine whether similar trends are observed in the generalist and specialist cell populations. In particular, we will consider:

- manipulations of biological parameters specific to the cells, such as the phenotypic switching rate and ECM degradation rate;
- manipulations of the environmental conditions, specifically the ECM volume fraction ahead of the invading wave.

4.3.4.1 Manipulations of cell parameters

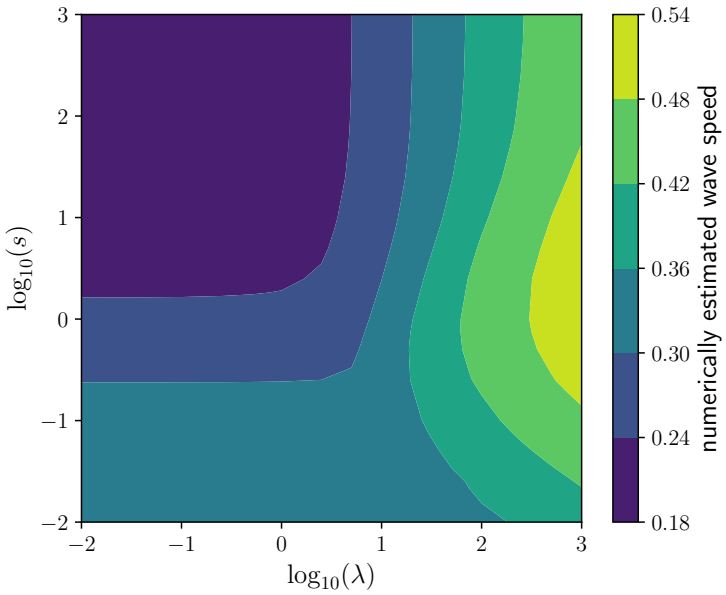

Figure 4.6: The numerically estimated speed of travelling wave solutions of the system (4.11)–(4.13) subject to the initial conditions for the cells in Equation (4.14) and for the ECM in Equation (4.4), with cell-dependent phenotypic switching (see Table 4.1), as a function of the phenotypic switching rate, s, and the rescaled ECM degradation rate, λ. The numerically estimated travelling wave speed was calculated for all combinations of values of $\lambda, s \in \{1, 2.5, 5, 7.5, 10, 25, 50, 75, 100, 250, 500, 750, 1000\}$. The initial ECM volume fraction ahead of the cells is $m_0 = 0.5$, the weighting of the specialists towards degrading ECM is $\theta_{S,D} = 0.5$ and $\alpha = 1$. For more information regarding the numerical methods used see Section 4.3.1.

Variations in the phenotypic switching rate Numerical simulations suggest that variations in the magnitude of the switching rate generally have a small impact on the cell migration speed when we employ constant, ECM- or space-dependent switching (for further details, see Appendix A.2, where phenotypic switching at different rates in either direction is also considered). However, when considering cell-dependent switching we find that changing the size of the switching rate has a significant impact on the speed of invasion. Figure 4.6 reveals that for low values of the ECM degradation rate, λ, the travelling wave speed increases as the switching rate, s, decreases and the wave front becomes smoother (see Figure 4.7). However, for sufficiently large λ, ECM degradation dominates over phenotypic switching to determine the travelling wave speed, and the maximum invasion speed is reached when $s = 1$. The opti-

mal switching rate, in terms of the fastest speed of invasion, is found analytically in Section 4.4.3.

When considering fast phenotypic switching, following ideas in Chapter 2, we can formally find expressions for the travelling wave speed in asymptotic regimes of the ECM degradation rate, $\lambda \to 0^+$ and $\lambda \to \infty$, that match the wave speed observed in numerical solutions to Equations (4.11)–(4.13) subject to the initial conditions for the cells in Equation (4.14) and for the ECM in Equation (4.4). This will be explored in more detail in Section 4.4.

Despite the phenotypic switching rate, s, having little quantitative impact on the travelling wave speed for most moderate parameter values across most phenotypic switching mechanisms considered, changing s does significantly impact the distribution of cells within the travelling wave front in all cases. By looking at the top row of Figure 4.7, when we have constant phenotypic switching, we see that increasing the switching rate balances the proportion of cells throughout the wave. As the phenotypic switching rate decreases, however, there is a larger proportion of cells in phenotypic state 2 at the front of the wave, and a wider travelling wave profile. For ECM-dependent switching, increasing the switching rate increases the proportion of cells in phenotypic state 1 at the wave front, and concentrates them to the front, leading to sharper travelling wave profiles. Alternatively, for space- and cell-dependent phenotypic switching mechanisms, increasing the switching rate reduces the volume fraction of cells in phenotypic state 2. This reduction is larger with space-dependent phenotypic switching. Qualitatively, the travelling wave profiles for space- and cell-dependent switching are almost identical (see the bottom two rows of Figure 4.7), and increasing the switching rate decreases the maximum volume fraction of cells in phenotypic state 2 at the front of the wave, shortening the tail of the travelling pulse and leading to sharper travelling wave profiles in the total volume fraction of cells.

Variations in ECM degradation rate In the model of a homogeneous cell type invading into ECM (as defined by Equations (4.1) and (4.2) in Chapter 2), the shape and speed of the travelling wave changes as the ECM degradation rate varies. Relationships between the rescaled ECM degradation rate in asymptotic regions and travelling wave speed were established in Chapter 2 for the fully non-dimensional model, without weightings, and it was shown that $c \to 2^-$ as $\lambda \to \infty$ and $c \to 2(1 - m_0)$ as $\lambda \to 0^+$ [77].

By inspecting Figure 4.8, we can see that for specialist cell populations, across all the switching functions that we consider in this work, an increase in the ECM

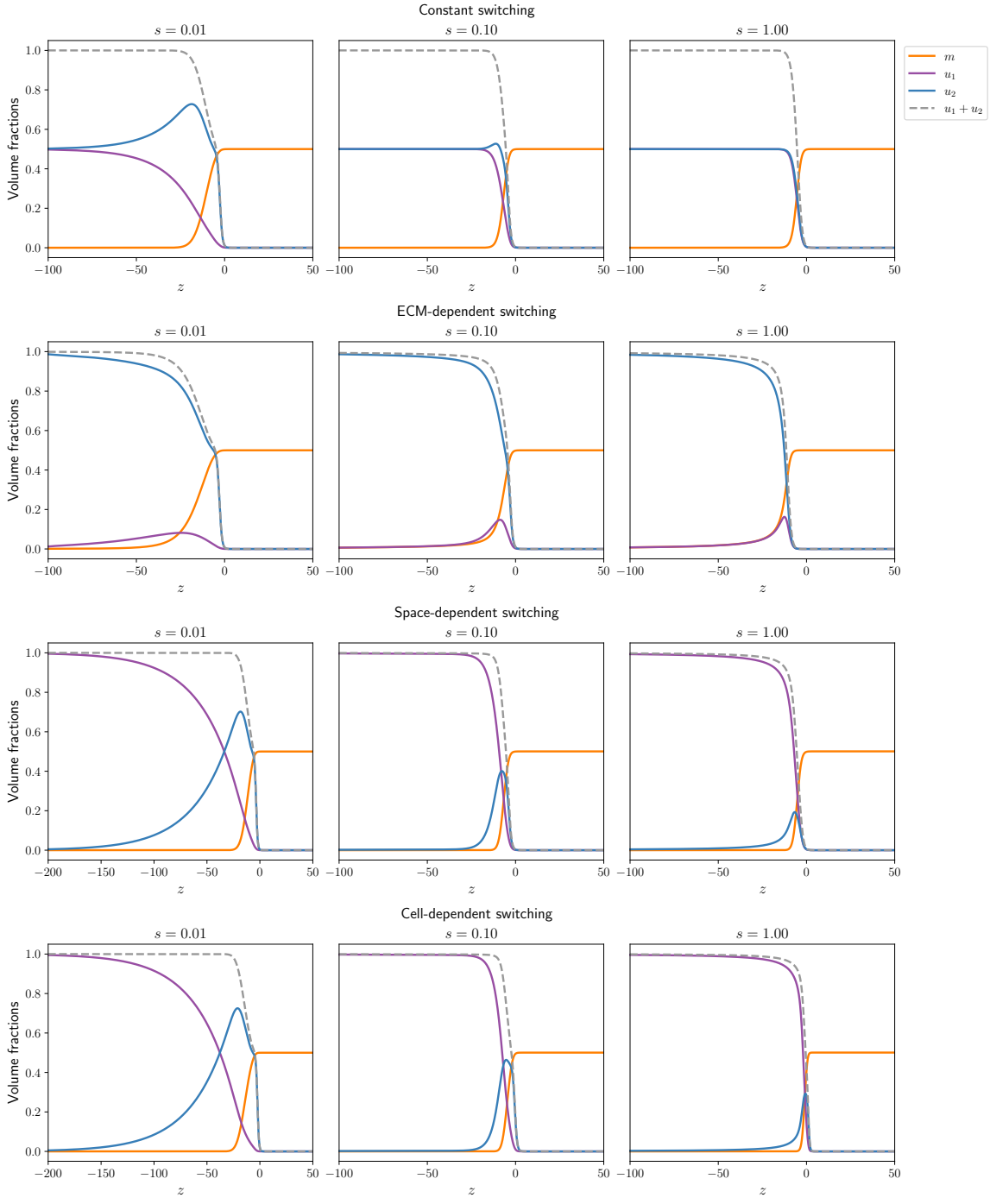

Figure 4.7: Travelling wave profiles of the solutions of the specialist system (4.11)–(4.13) subject to the initial conditions for the cells in Equation (4.14) and for the ECM in Equation (4.4), plotted as a function of the travelling wave variable $z = x - ct$, where c is the numerically observed constant travelling wave speed. These solutions show that changing the switching rate changes the distribution of the cell phenotypes within the invading wave front. Here, the initial ECM volume fraction ahead of the cells is $m_0 = 0.5$, the ECM degradation rate is $\lambda = 1$, the weighting of the specialists towards degrading ECM is $\theta_{S,D} = 0.5$ and The width of the region initially invaded by cells is $\alpha = 1$ in all cases. For more information regarding the numerical methods used see Section 4.3.1.

degradation rate leads to an increase in the numerically estimated travelling wave speed when ECM degradation rates are above a critical value. We also see that as $\lambda \to \infty$ there is convergence in the travelling wave speed observed for constant, space-dependent and cell-dependent switching mechanisms, but to different specific values. In Section 4.4, we perform formal analysis of the system (4.11)–(4.13) for general switching rates that can differ in either direction. In the particular case where the switching rate is the same in either direction, we show analytically in Section 4.4 that, in the fast phenotypic switching regime, $c \to (1-m_0)\sqrt{1-\theta_{S,D}}$ as $\lambda \to 0^+$ and $c \to \sqrt{1-\theta_{S,D}}$ as $\lambda \to \infty$ for constant phenotypic switching. Convergence of the numerically estimated travelling wave speed to these values can be seen in Figure 4.9 in Section 4.4. In contrast to the other switching functions, ECM-dependent switching is far less sensitive to changes in ECM degradation rates at low initial ECM volume fractions ahead of the cells, and convergence of the travelling wave speed to a constant value is not observed within the parameter ranges considered in this work.

From a biological perspective, the limit $\lambda \to \infty$ is relevant in describing cells in an aggressive tumour that have a very high ability to degrade ECM, which may enable them to invade much faster. Alternatively, for a sufficiently small product $\theta_{S,D}\lambda$, specialist cells in phenotypic state 1 should focus more of their ability on movement (*i.e.*, decrease $\theta_{S,D}$) in order to increase migration speed, since a small change in the ECM degradation rate alone, when below some threshold value, will minimally impact the migration speed.

4.3.4.2 Manipulations of the environmental conditions

For a generalist cell population, it can be shown analytically and numerically that the speed of the travelling wave of migrating cells increases as the initial ECM volume fraction ahead of the cells decreases [77]. Specifically, for small ECM degradation rates (*i.e.*, $\lambda \to 0^+$), there is a linear relationship between the travelling wave speed,

$$c_G = 2(1-m_0)\sqrt{\theta_{G,P}(1-\theta_{G,D}-\theta_{G,P})},$$

and the initial volume fraction of ECM ahead of the wave, $m_0 \in [0,1)$.

Examining Figure 4.8, it is clear that, across all four phenotypic switching functions considered, the speed of the travelling wave of the specialist cell population increases as the initial ECM volume fraction ahead of the cells decreases. Biologically, a lower ECM volume fraction corresponds to a less densely packed region of ECM ahead of the cells which facilitates faster cell invasion.

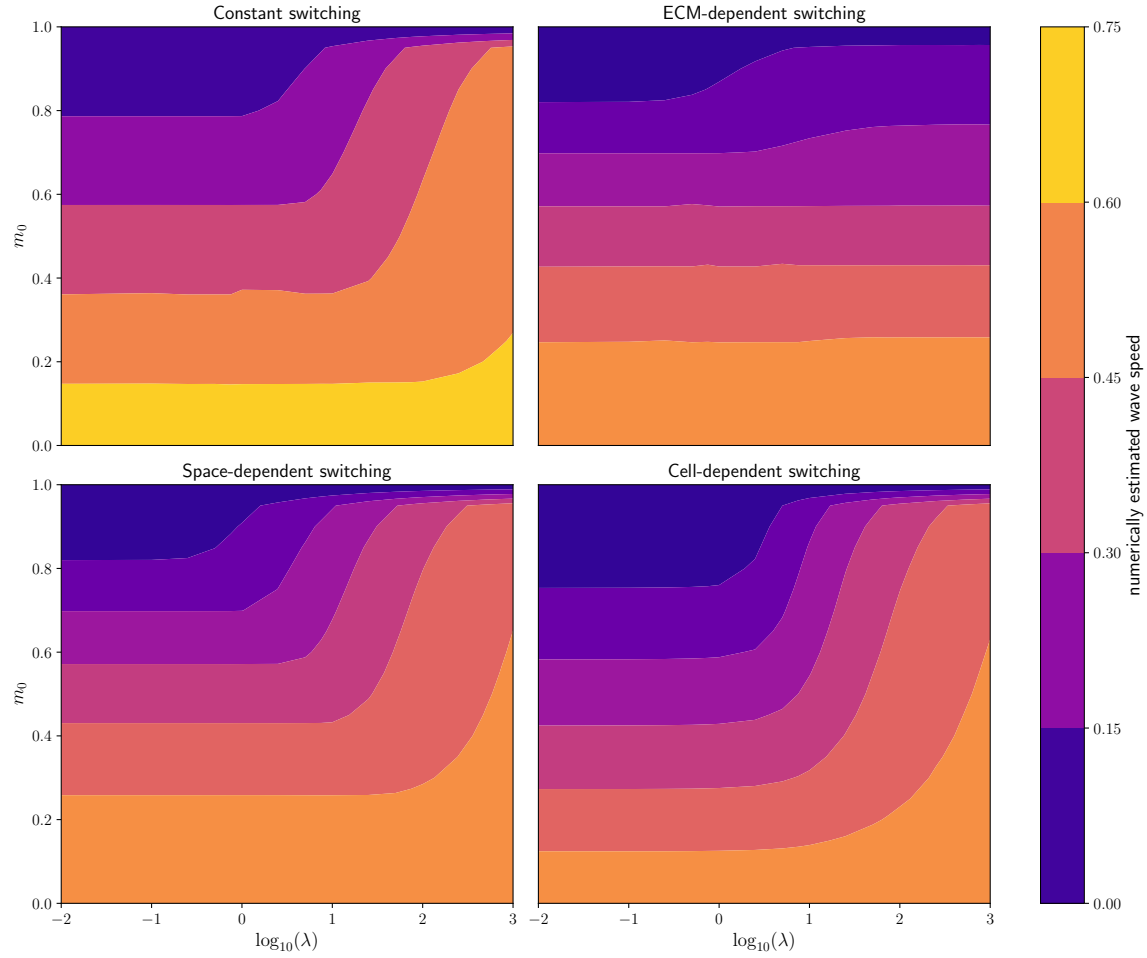

Figure 4.8: Plots summarising the relationship between the numerically estimated speed of travelling wave solutions of Equations (4.11)–(4.13) subject to the initial conditions for the cells in Equation (4.14) and for the ECM in Equation (4.4), the initial volume fraction of ECM ahead of the cells, $m_0 \in \{0, 0.05, 0.1, \ldots, 0.95, 1\}$, and ECM degradation rate, $\lambda \in \{10^{-2}, 10^{-1}, \ldots, 10^{2}, 10^{3}\}$. Here, the weighting of the specialists towards degrading ECM is $\theta_{S,D} = 0.5$, the switching rate is $s = 1$ and $\alpha = 1$ in all cases. For more information regarding the numerical methods used see Section 4.3.1.

4.4 Travelling wave analysis in the fast phenotypic switching regime

To explore the travelling wave speed at extreme parameter values, we consider a regime where phenotypic switching is faster than cell motility and proliferation. To do this, we study the following rescaled model with variables u_{1_ϵ}, u_{2_ϵ} and $m_\epsilon \in [0,1]$ as before

$$\begin{aligned}
\frac{\partial u_{1_\epsilon}}{\partial t} &= (1-\theta_{S,D})\frac{\partial}{\partial x}\left[\big(1-u_{1_\epsilon}-u_{2_\epsilon}-m_\epsilon\big)\frac{\partial u_1}{\partial x} + u_{1_\epsilon}\frac{\partial}{\partial x}\big(u_{1_\epsilon}+u_{2_\epsilon}+m_\epsilon\big)\right] \\
&\qquad + \frac{1}{\epsilon}u_{2_\epsilon}\gamma_{21}(u_{1_\epsilon},u_{2_\epsilon},m_\epsilon) - \frac{1}{\epsilon}u_{1_\epsilon}\gamma_{12}(u_{1_\epsilon},u_{2_\epsilon},m_\epsilon), \\
\frac{\partial u_{2_\epsilon}}{\partial t} &= \theta_{S,P}u_{2_\epsilon}\big(1-u_{1_\epsilon}-u_{2_\epsilon}-m_\epsilon\big) \\
&\qquad - \frac{1}{\epsilon}u_{2_\epsilon}\gamma_{21}(u_{1_\epsilon},u_{2_\epsilon},m_\epsilon) + \frac{1}{\epsilon}u_{1_\epsilon}\gamma_{12}(u_{1_\epsilon},u_{2_\epsilon},m_\epsilon), \\
\frac{\partial m_\epsilon}{\partial t} &= -\theta_{S,D}\lambda m u_{1_\epsilon},
\end{aligned}$$

where $\epsilon \ll 1$, $x \in \mathbb{R}$ and $t \in \mathbb{R}_+$. Here, $\lambda \in \mathbb{R}_+$ remains as the rescaled rate of ECM degradation by cells in phenotypic state 1, whilst $\theta_{S,D} \in [0,1]$ describes the weighting of cells in phenotypic state 1 towards degrading ECM, $\theta_{S,P} = 1$ describes the weighting of cells in phenotypic state 2 towards proliferation, and

$$\gamma_{12} : \mathbb{R}^3_+ \to \mathbb{R}_+ \qquad \text{and} \qquad \gamma_{21} : \mathbb{R}^3_+ \to \mathbb{R}_+,$$

are the same non-dimensional phenotypic switching functions as in earlier parts of this chapter.

Simulations reveal that the model admits constant profile, constant speed travelling wave solutions, so we introduce the travelling wave ansatz

$$\begin{aligned}
U_{i_\epsilon}(z) &= U_{i_\epsilon}(x-ct) = u_{i_\epsilon}(x,t), \\
M_\epsilon(z) &= M_\epsilon(x-ct) \ = m_\epsilon(x,t),
\end{aligned}$$

for $i = \{1, 2\}$, where $c \in \mathbb{R}_+$, that satisfy the following system of ODEs:

$$-c\frac{\mathrm{d}U_{1_\epsilon}}{\mathrm{d}z} = (1-\theta_{S,D})\frac{\mathrm{d}}{\mathrm{d}z}\left[(1-U_{1_\epsilon}-U_{2_\epsilon}-M_\epsilon)\frac{\mathrm{d}U_{1_\epsilon}}{\mathrm{d}z} + U_{1_\epsilon}\frac{\mathrm{d}}{\mathrm{d}z}(U_{1_\epsilon}+U_{2_\epsilon}+M_\epsilon)\right] - \frac{1}{\epsilon}U_{1_\epsilon}\gamma_{12}(U_{1_\epsilon}, U_{2_\epsilon}, M_\epsilon) + \frac{1}{\epsilon}U_{2_\epsilon}\gamma_{21}(U_{1_\epsilon}, U_{2_\epsilon}, M_\epsilon), \tag{4.15}$$

$$-c\frac{\mathrm{d}U_{2_\epsilon}}{\mathrm{d}z} = U_{2_\epsilon}(1-U_{1_\epsilon}-U_{2_\epsilon}-M_\epsilon) + \frac{1}{\epsilon}U_{1_\epsilon}\gamma_{12}(U_{1_\epsilon}, U_{2_\epsilon}, M_\epsilon) - \frac{1}{\epsilon}U_{2_\epsilon}\gamma_{21}(U_{1_\epsilon}, U_{2_\epsilon}, M_\epsilon), \tag{4.16}$$

$$-c\frac{\mathrm{d}M_\epsilon}{\mathrm{d}z} = -\lambda\theta_{S,D}U_{1_\epsilon}M_\epsilon. \tag{4.17}$$

Combining Equations (4.15) and (4.16) we find that the total cell volume fraction

$$U_\epsilon(z) = U_{1_\epsilon}(z) + U_{2_\epsilon}(z),$$

satisfies the ODE

$$-c\frac{\mathrm{d}U_\epsilon}{\mathrm{d}z} = (1-\theta_{S,D})\frac{\mathrm{d}}{\mathrm{d}z}\left[(1-U_\epsilon-M_\epsilon)\frac{\mathrm{d}U_{1_\epsilon}}{\mathrm{d}z} + U_{1_\epsilon}\frac{\mathrm{d}}{\mathrm{d}z}(U_\epsilon+M_\epsilon)\right] + U_{2_\epsilon}(1-U_\epsilon-M_\epsilon), \tag{4.18}$$

for $z \in \mathbb{R}$. We now consider constant phenotypic switching given by

$$\gamma_{12}(u_{1_\epsilon}, u_{2_\epsilon}, m_\epsilon) = s_{12} \geq 0 \qquad \text{and} \qquad \gamma_{21}(u_{1_\epsilon}, u_{2_\epsilon}, m_\epsilon) = s_{21} \geq 0,$$

and look for an analytical expression for the travelling wave speed in asymptotic regions of λ, following the ideas in [77].

By considering the asymptotic expansions around U_ϵ, U_{1_ϵ}, U_{2_ϵ} and M_ϵ such that the leading-order terms are given by U, U_1, U_2 and M, respectively, then as $\epsilon \to 0^+$ we formally find, from Equations (4.15) and (4.16), that

$$U_i(z) = \omega_i U, \tag{4.19}$$

where

$$\omega_1 = \frac{s_{21}}{s_{12}+s_{21}},$$
$$\omega_2 = \frac{s_{12}}{s_{12}+s_{21}}.$$

By substitution, Equation (4.18) becomes

$$-c\frac{\mathrm{d}U}{\mathrm{d}z} = \omega_1(1-\theta_{S,D})\frac{\mathrm{d}}{\mathrm{d}z}\left[(1-U-M)\frac{\mathrm{d}U}{\mathrm{d}z} + U\frac{\mathrm{d}}{\mathrm{d}z}(U+M)\right] + \omega_2 U(1-U-M),$$

which can be expanded and written as

$$\omega_1(1-\theta_{S,D})(1-M)\frac{\mathrm{d}^2U}{\mathrm{d}z^2}+c\frac{\mathrm{d}U}{\mathrm{d}z}+\omega_2U(1-U-M)=-\omega_1(1-\theta_{S,D})U\frac{\mathrm{d}^2M}{\mathrm{d}z^2}. \quad (4.20)$$

Furthermore, Equation (4.17) can be written as

$$c\frac{\mathrm{d}M}{\mathrm{d}z}=\lambda\theta_{S,D}\omega_1UM, \quad (4.21)$$

which yields

$$\begin{aligned}\frac{\mathrm{d}^2M}{\mathrm{d}z^2}&=\frac{\lambda\theta_{S,D}\omega_1}{c}\left(U\frac{\mathrm{d}M}{\mathrm{d}z}+M\frac{\mathrm{d}U}{\mathrm{d}z}\right)\\&=\frac{\lambda\theta_{S,D}\omega_1}{c}\left(\frac{\lambda\theta_{S,D}\omega_1}{c}MU^2+M\frac{\mathrm{d}U}{\mathrm{d}z}\right).\end{aligned} \quad (4.22)$$

Substituting Equation (4.22) into Equation (4.20) we find

$$\begin{aligned}&\omega_1(1-\theta_{S,D})(1-M)\frac{\mathrm{d}^2U}{\mathrm{d}z^2}+c\frac{\mathrm{d}U}{\mathrm{d}z}+\omega_2U(1-U-M)\\&\quad=-\frac{\lambda\theta_{S,D}\omega_1^2(1-\theta_{S,D})}{c}MU\left[\frac{\lambda\theta_{S,D}\omega_1}{c}U^2+\frac{\mathrm{d}U}{\mathrm{d}z}\right].\end{aligned} \quad (4.23)$$

Moreover, solving Equation (4.21) subject to the boundary condition $M(z)\to m_0$ as $z\to\infty$, where $m_0\in[0,1]$, gives

$$M(z)=m_0\exp\left\{-\frac{\lambda\theta_{S,D}\omega_1}{c}\int_z^\infty U(s)\mathrm{d}s\right\}. \quad (4.24)$$

Under the boundary conditions $U_i(z)\to 0$ as $z\to\infty$ for $i=1,\,2$ we have $U(z)\to 0$ as $z\to\infty$. At the migrating front of the travelling wave (*i.e.*, for $z\in(\ell,\infty)$ with $1\ll\ell<\infty$), we can use the ansatz

$$U(z)\approx\exp\{-\beta z\},$$

where $\beta\in(0,\infty)$ to give

$$M(z)=m_0\exp\left\{-\frac{\lambda\theta_{S,D}\omega_1}{\beta c}U(z)\right\}, \quad (4.25)$$

for $z\in(\ell,\infty)$.

4.4.1 Formal asymptotic analysis for $\lambda \to 0^+$

Using Equation (4.25), it is clear that

$$M(z) \approx m_0 \exp\left\{ -\frac{\lambda\theta_{S,D}\omega_1}{\beta c}U(z)\right\} \to m_0 \qquad \text{as} \qquad \lambda \to 0^+, \tag{4.26}$$

for $z \in (\ell, \infty)$. In the asymptotic regime $\lambda \to 0^+$, since $0 < U(z) < 1$ and $\mathrm{d}U(z)/\mathrm{d}z \approx -\beta U(z)$ for $z \in (\ell, \infty)$, substituting Equation (4.26) into the right-hand side of Equation (4.23) entails

$$-\frac{\lambda\theta_{S,D}\omega_1^2(1-\theta_{S,D})}{c}U(z)m_0 \exp\left\{ -\frac{\lambda\theta_{S,D}\omega_1}{\beta c}U(z)\right\}\left[\frac{\lambda\theta_{S,D}\omega_1}{c}U^2(z) + \frac{\mathrm{d}U(z)}{\mathrm{d}z}\right] \to 0,$$

as $\lambda \to 0^+$ for $z \in (\ell, \infty)$. Formally, we find from the left-hand side of Equation (4.23) that

$$\omega_1(1-\theta_{S,D})(1-M)\frac{\mathrm{d}^2U}{\mathrm{d}z^2} + c\frac{\mathrm{d}U}{\mathrm{d}z} + \omega_2 U(1-U-M) \approx 0, \tag{4.27}$$

for $z \in (\ell, \infty)$. We notice that Equation (4.27) is equivalent to the Fisher–KPP model [109, 174] in travelling wave co-ordinates:

$$\tilde{D}\frac{\mathrm{d}^2\tilde{U}(z)}{\mathrm{d}z^2} + \tilde{c}\frac{\mathrm{d}\tilde{U}}{\mathrm{d}z} + \tilde{r}\tilde{U}\left(1 - \frac{\tilde{U}}{\tilde{K}}\right) = 0, \tag{4.28}$$

where we have $\tilde{D} = \omega_1(1-\theta_{S,D})(1-m_0)$, $\tilde{r} = \omega_2(1-m_0)$ and $\tilde{K} = (1-m_0)$. This correctly predicts (see Figure 4.9), as $\lambda \to 0^+$, a minimum travelling wave speed given by

$$c_{\min} = 2(1-m_0)\sqrt{\omega_1\omega_2(1-\theta_{S,D})}, \tag{4.29}$$

which can be observed in Figure 4.9 to agree with the numerically estimated wave speed for the system (4.11)–(4.13) when $\lambda \to 0^+$.

4.4.2 Formal asymptotic analysis for $\lambda \to \infty$

Revisiting the semi-explicit solution for M given by Equation (4.25), we find

$$M(z) \approx m_0 \exp\left\{ -\frac{\lambda\theta_{S,D}\omega_1}{\beta c}U(z)\right\} \to 0 \qquad \text{as} \qquad \lambda \to \infty, \tag{4.30}$$

for $z \in (\ell, \infty)$. In the asymptotic regime $\lambda \to \infty$, since $0 < U(z) < 1$ and $\mathrm{d}U(z)/\mathrm{d}z \approx -\beta U(z)$ for $z \in (\ell, \infty)$, substituting Equation (4.30) into the right-hand side of Equation (4.23) shows that

$$-\frac{\lambda\theta_{S,D}\omega_1^2(1-\theta_{S,D})}{c}U(z)m_0 \exp\left\{ -\frac{\lambda\theta_{S,D}\omega_1}{\beta c}U(z)\right\}\left[\frac{\lambda\theta_{S,D}\omega_1}{c}U^2(z) + \frac{\mathrm{d}U(z)}{\mathrm{d}z}\right] \to 0,$$

as $\lambda \to \infty$ for $z \in (\ell, \infty)$. Looking at the left-hand side of Equation (4.23) we then find

$$\omega_1(1-\theta_{S,D})\frac{\mathrm{d}^2U}{\mathrm{d}z^2} + c\frac{\mathrm{d}U}{\mathrm{d}z} + \omega_2 U(1-U) \approx 0, \tag{4.31}$$

for $z \in (\ell, \infty)$. In this case, Equation (4.31) is equivalent to the Fisher–KPP model (see Equation (4.28)) with parameters $\tilde{D} = \omega_1(1-\theta_{S,D})$, $\tilde{r} = \omega_2$ and $\tilde{K} = 1$, so when $\lambda \to \infty$ we have

$$c_{\min} = 2\sqrt{\omega_1\omega_2(1-\theta_{S,D})}. \tag{4.32}$$

Figure 4.10 visually demonstrates that the shape of the travelling wave of the system (4.11)–(4.13) converges as $\lambda \to \infty$ to the shape of the travelling wave solution of the Fisher–KPP model (see Equation (4.28)) with parameters $\tilde{D} = \omega_1(1-\theta_{S,D})$, $\tilde{r} = \omega_2$ and $\tilde{K} = 1$ as $\lambda \to \infty$.

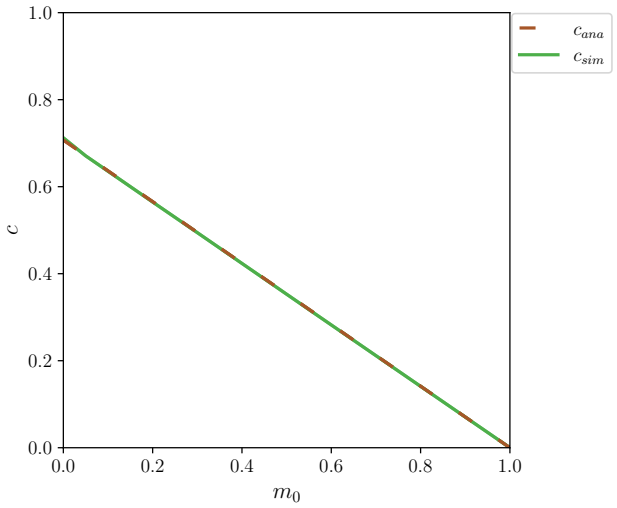

Figure 4.9: Plot showing the analytically predicted minimum travelling wave speed, c_{ana}, and the numerically estimated travelling wave speed, c_{sim}, of solutions of the system (4.11)–(4.13) subject to the initial conditions for the cells as in Equation (4.14), and for the ECM as in Equation (4.4), for very low ECM degradation rates, $\lambda \to 0^+$, and various initial ECM volume fractions ahead of the cells, m_0, in fast phenotypic switching regimes (defined by $s_{12} = s_{21} = s = 10^4$). m_0 took values between 0 and 1 inclusive, and simulations were run at 0.05 intervals in ECM density. The numerically estimated travelling wave speeds plotted in green (three curves) are for simulations with $\lambda = \{10^{-2}, 10^{-3}, 10^{-4}\}$, while the analytical wave speeds plotted in red are given by Equation (4.29). The weighting of specialists towards degradation is $\theta_{S,D} = 0.5$ and $\alpha = 1$ across all simulations. For more information regarding the numerical methods used see Section 4.3.1.

4.4.3 Maximising the travelling wave speed

In both $\lambda \to 0^+$ and $\lambda \to \infty$ regimes, the travelling wave speeds, determined by Equation (4.29) and Equation (4.32), respectively, are maximised when $\omega_1\omega_2$ is maximised. By considering

$$\omega_1\omega_2 = \frac{s_{12}s_{21}}{(s_{12}+s_{21})^2}, \tag{4.33}$$

and differentiating twice with respect to s_{12}, it is clear that

$$\max(\omega_1\omega_2) = 0.25, \tag{4.34}$$

which is obtained when $s_{12} = s_{21}$. As such, we can conclude that the travelling wave speed is always maximised in these limits when phenotypic switching between states 1 and 2 occurs at the same rate in either direction. Intuitively, we believe this is due to a balance between both diffusivity and growth required between the two populations in order to invade at an optimum speed.

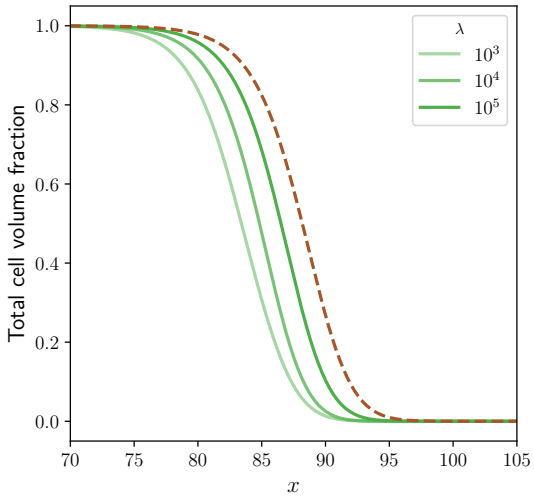

Figure 4.10: Plot of the total cell volume fraction obtained through numerical simulations of Equations (4.11)–(4.13) with constant phenotypic switching (see Table 4.1) subject to the initial conditions for the cells as in Equation (4.14) and for the ECM as in Equation (4.4), for large values of λ (solid lines), and numerical simulations of the Fisher–KPP model given by Equation (4.31) (dashed red line). In all simulations, solutions are shown at $t = 100$ and the initial ECM volume fraction ahead of the cells is $m_0 = 0.1$. The width of the region initially invaded by cells is $\alpha = 1$, the weighting of specialists towards degradation is $\theta_{S,D} = 0.1$ and the switching rate for all functions is $s = \lambda$. Qualitatively, the same behaviour is observed for all $m_0 \in [0, 1)$. For more information regarding the numerical methods used see Section 4.3.1.

4.5 Discussion

Recently, population heterogeneity has been recognised as an important driver of collective cell migration and has attracted significant attention [341]. Here, we have extended the model for a homogeneous generalist population of cells migrating into ECM developed in Chapter 2 to explicitly incorporate phenotypic heterogeneity under the migration-proliferation dichotomy. We considered how distinct phenotypic switching mechanisms impact population structure, and the dependence of the travelling wave speed on different biological parameters. Specifically, we considered constant switching, ECM-dependent switching, space-dependent switching and cell-dependent switching.

Initially, we compared a homogeneous cell population to a heterogeneous cell population without phenotypic switching. In this case, it is clear that a specialist cell population without the ability to change its phenotypic state can never outcompete a generalist cell population as the model does not admit travelling wave solutions. Conversely, analysis of the model for a specialist cell population invading into ECM that includes phenotypic switching shows that a heterogeneous cell population can produce travelling waves of invasion with a faster speed than homogeneous generalist populations that weight their ability towards movement. Moreover, we confirmed that the travelling wave speed in the specialist model, irrespective of the phenotypic switching mechanism, depends qualitatively on the ECM degradation rate and initial ECM volume fraction ahead of the cells in the same manner as for the generalist model.

This work considers phenotypic switching that is equal in either direction, and in this case the travelling wave speed is shown to be independent of the switching rate for constant, ECM- and space-dependent switching. When asymmetric switching rates are considered, both switching rates impact the speed and distribution of phenotypes in the invading cell population (see Appendix A.2 and Section 4.4). In the case of cell-dependent switching, increasing the switching rate decreases the travelling wave speed.

Biologically, there exist a number of factors that could drive phenotypic switching. For example, direct contact between cell surfaces, or contact on the cell surface from molecules released by neighbouring cells or ECM, which provide information about surrounding cell and ECM volume fractions, might cause phenotypic change [332, 274]. This work demonstrates that the mechanism determining the form of phenotypic switching function has a profound impact on the phenotypic structure of the invading

cell population. When there is no environmental dependence, a well-mixed population of cells invades into ECM, with the ratio of cells in phenotypic state 1 (moving and degrading cells) to phenotypic state 2 (proliferative cells) in the bulk being determined by the ratio between the switching rates. In ECM-dependent switching models, ECM degrading cells in phenotypic state 1 occupy the migrating front, with proliferating cells in the bulk, as observed in many examples of leader-follower dynamics [175] where both sub-populations play an important role in driving cell invasion. Space- and cell-dependent switching models exhibit the opposite distribution, with proliferating cells leading the way, as observed *in vivo* and *in vitro* in melanoma spheroid growth, where proliferative clusters form on the outer edges of a larger bulk cluster consisting of migratory cells [50]. This could be favourable for the population, for example, if there are lower energetic requirements for proliferation in low volume cell density regions. A further extension of this work would be to consider a general switching function combining the influence of both available space and ECM, to find critical parameter values that determine the transition between the degraders or the proliferators leading the invasive population.

The speed of invasion of a cell population alone does not necessarily allow us to distinguish the mechanisms governing collective cell migration. In the case of space- and cell-dependent switching, changing the phenotypic switching rate and observing the changes in the resulting travelling wave speed may indeed be sufficient to distinguish between the two switching mechanisms, where differences in the travelling wave profile are otherwise very subtle. However, in other cases, the spatial structure of each cell sub-population within an invading wave might provide further insights. For instance, examination of a detailed population profile from a tissue biopsy in the direction of migration could be used as a predictive tool to distinguish the mechanism underlying cell phenotypic switching which could, in turn, be used to help develop therapeutic treatments. Moreover, if such a biopsy revealed cells in one phenotypic state only at the front of the invading wave, this may indicate that the rate of phenotypic switching in one direction far exceeds the rate of phenotypic switching in the opposite direction. For example, in the case of ECM-dependent switching, a population of primarily ECM-degrading cells observed at the migrating front would suggest that phenotypic switching from proliferative to degrading phenotypic state is much faster than from degrading to proliferative. Simulations in Appendix A.2 simultaneously revealed a trade-off between population structuring and the travelling wave speed, such that when a single population leads the migrating front, the invasion speed is reduced. This suggests that both the spatial profile and speed of invasion

(measured experimentally using the distance travelled over a given time) are required to distinguish the underlying phenotypic switching mechanisms and that further possible therapeutic treatments could be developed to slow tumour growth by preventing switching in one direction, or to speed up wound healing or developmental processes by initiating symmetrical switching.

In conclusion, understanding the phenotypic structure of invading cellular collectives is an important objective, attracting significant research over recent years. The model presented in this study, whilst clearly simple, provides compelling insights associated with the speed and structure of heterogeneous cell invasion into ECM under various phenotypic switching mechanisms, and provides a basis for more complex and detailed model development and analysis in the future.

Chapter 5

A model of a continuum of cell phenotypes invading into the local environment

This work was completed in collaboration with P. K. Maini and R. E. Baker and is published in the Bulletin of Mathematical Biology [76]. The author was responsible for all aspects of the research and writing presented in this chapter.

5.1 Introduction

As seen in Chapter 4, when there are a discrete set of known cell phenotypes with distinct behaviours, it may be most appropriate to model the phenotypic state as a discrete variable, typically using integer values. However, when there are numerous (or potentially infinite) phenotypes with incremental differences or smoothly transitioning behaviours between them, then a continuously structured phenotype model might be more pertinent. In this case, the resulting evolution equation for the cell population density often takes the form of a non-local reaction–diffusion equation [17, 31, 191], or a non-local advection–reaction–diffusion equation [56, 57, 192].

One critical issue arises when these non-local continuum models describing a continuous phenotype are constructed phenomenologically, without derivation from an underlying IBM. In such cases, we may not fully understand the biological significance of the terms within the model, especially in realising their connection to the behaviours of the individual cells and their interactions with the local environment, leaving the meaning of various terms in the model ambiguous. Moreover, we may unknowingly be making intrinsic assumptions about cell behaviour that could be invalid.

This chapter, therefore, focuses on constructing a general continuum model that is explicitly derived from the individual-based behaviours of the cells and their interactions with their surrounding environment. The underlying IBM has a discrete structure in time, physical space and phenotype, whereas the resulting PDE found by taking appropriate limits has continuously structured space, time and phenotype variables. This approach ensures that each term in the resulting continuum model has a well-defined interpretation in relation to the underlying cell dynamics, providing clarity and a deeper understanding of the biological processes being modelled. This chapter, however, is not concerned with a detailed quantitative comparison between individual- and population-level models as this is already well explored elsewhere in the literature [15, 48, 66, 190, 191, 195, 228, 229, 293, 316].

As such, we hereafter present a methodology for deriving a continuously structured PDE model for general cell migration into the local environment that is derived from an underlying IBM that takes into account the individual interactions between cells and their local environment. Using this approach, the model's applicability to various biological scenarios is demonstrated, highlighting the flexibility of this general framework and the consistent, coherent connections between the microscale behaviours captured in the IBM and the macroscale descriptions in the resulting PDE model. For the purposes of the derivation of the continuum model, cell invasion into the local environment is chosen in general, but due to this generality, the local environment could be replaced by a variety of substances, such as neighbouring tissues or organs, a tissue engineering scaffold or a wound, during healing. The applications in this chapter are carefully selected to illustrate a wide range of cell behaviours by employing different functional forms that describe the probabilities of movement in both physical and phenotypic spaces, as well as the behaviours governing proliferation at the individual-based level. However these applications were not chosen with the aim to provide detailed biological insights at this stage.

5.2 The individual-based model

In order to incorporate microscopic descriptions of the interactions occurring between cells and their local environment, we formulate a one-dimensional phenotype-structured, on-lattice, IBM for collective cell migration.

In this model, the cells are represented as individual, discrete agents and we assume that the features of the local environment we are interested in, such as the ECM or other cell types, occupy some finite volume in a limited space. In order to

fit in with the individual-based framework, we therefore choose to model the local environment as being composed of individual, discrete elements of the same finite volume as the cells. Depending on the phenotype of the individual cell, and the number of cells and elements of the local environment in the same lattice site, each individual cell has a capacity to undergo random, undirected movement, heritable phenotypic changes and proliferation, that can be adapted to the specific biological application of interest by employing appropriate individual-based rules to describe these changes. Furthermore, we assume that each individual cell can also interact with the surrounding environment, and that the cell's capacity to impact their local environment depends on the phenotype of the cell.

Considering a one-dimensional spatial domain, we allow the cells and the local environment to be distributed in the region $x \in [X_{\min}, X_{\max}]$. We describe the phenotypic state of each individual cell on a one-dimensional domain through a structured variable $y \in [Y_{\min}, Y_{\max}]$. We discretise the time variable $t \in \mathbb{R}^+$, as $t_h = h\Delta_t$ with $h \in \mathbb{N}$ and $\Delta_t \in \mathbb{R}^+$, we discretise the spatial variable into an integer number of lattice sites $x_i = X_{\min} + \Delta_x(i-1)$ for $\Delta_x \in \mathbb{R}^+$ and $i = 1, \ldots, N_x + 1$, and we discretise the phenotype variable using $y_j = Y_{\min} + \Delta_y(j-1)$ for $\Delta_y \in \mathbb{R}^+$ and $j = 1, \ldots, N_y + 1$. In this case, $\Delta_t, \Delta_x, \Delta_y \in \mathbb{R}^+$ are the time-, space- and phenotype-steps, respectively.

We introduce the dependent variable $n_i^j(t_h) \in \mathbb{N}_0$ to model the number of cells that occupy position $x_i \times y_j \in [X_{\min}, X_{\max}] \times [Y_{\min}, Y_{\max}]$ on the lattice at time t_h, where $\mathbb{N}_0$ represents the natural numbers, including zero. Then, we define the total cell number at a spatial position x_i at time t_h as $N_i(t_h) = \sum_{j=1}^{N_y+1} n_i^j(t_h) \in \mathbb{N}_0$ and the number of elements of the local environment at spatial position x_i at time t_h is denoted by $e_i(t_h) \in \mathbb{N}_0$.

5.2.1 Modelling the dynamics of the cells

We denote by $p(\mathbf{n}, \mathbf{e}, t_h)$ the joint probability that the number of cells in spatial position x_i in phenotypic state y_j (for $i = 1, \ldots, N_x + 1$ and $j = 1, \ldots, N_y + 1$) at time t_h ($h \in \mathbb{R}$) is given by $\mathbf{n} = ([n_1^1, \ldots, n_{N_x+1}^1], \ldots, [n_1^{N_y+1}, \ldots, n_{N_x+1}^{N_y+1}])$ and that the number of discrete, constitutive elements of the local environment in spatial position x_i for $i = 1, \ldots, N_x + 1$ is given by $\mathbf{e} = [e_1, \ldots, e_{N_x+1}]$.

Between a time step t_h and t_{h+1} (equivalently described by $t_h + \Delta_t$), each cell in phenotypic state $y_j \in [Y_{\min}, Y_{\max}]$ at position $x_i \in [X_{\min}, X_{\max}]$ can undergo random movement, heritable phenotypic changes and cell proliferation independently of time and according to the following assumptions in this section. We note here that we write $n_i^j(t_h)$ as n_i^j and $e_i(t_h)$ as e_i for simplicity going forward.

5.2.1.1 Random cell movement

We model cell movement in space as an on-lattice, biased random walk between neighbouring lattice sites. We assume the probability of cell movement can depend on a number of factors, such as the local environment or the phenotype of the cell, but it is easy to relax or vary these assumptions to consider other variables of interest. In particular, we introduce the following two changes in state vector that describe movement to the left or right in physical space of a single cell in phenotypic state y_j into position $x_{i\pm1}$ from position x_i:

$$\begin{aligned} L^{\mathrm{m}}_{i,j} : [n^j_1, \ldots, n^j_{i-1}, n^j_i, \ldots, n^j_{N_x+1}] &\longrightarrow [n^j_1, \ldots, n^j_{i-1}+1, n^j_i-1, \ldots, n^j_{N_x+1}], \\ &\text{for} \quad i = 2, \ldots, N_x+1, \quad j = 1, \ldots, N_y+1, \\ R^{\mathrm{m}}_{i,j} : [n^j_1, \ldots, n^j_i, n^j_{i+1}, \ldots, n^j_{N_x+1}] &\longrightarrow [n^j_1, \ldots, n^j_i-1, n^j_{i+1}+1, \ldots, n^j_{N_x+1}], \\ &\text{for} \quad i = 1, \ldots, N_x, \quad j = 1, \ldots, N_y+1, \end{aligned}$$

where $L^{\mathrm{m}}_{i,j}, R^{\mathrm{m}}_{i,j} : \mathbb{N}^{N_x+1} \to \mathbb{N}^{N_x+1}$. We assume that the probability of cell movement depends on the phenotype of the cell and on the number of cells and elements of the local environment in the target site, rather than the lattice site that they are currently in. As such, we define the probability of movement to the left, to spatial position x_{i-1} from x_i, during a single time step Δ_t, as

$$\beta_-(j, N_{i-1}, e_{i-1}) \in [0,1], \qquad i = 2, \ldots, N_x+1, \quad j = 1, \ldots, N_y+1,$$

and the probability of movement to the right, to spatial position x_{i+1} from x_i, during a single time step Δ_t, as described by the change to the state vector $R^{\mathrm{m}}_{i,j}$, as

$$\beta_+(j, N_{i+1}, e_{i+1}) \in [0,1], \qquad i = 1, \ldots, N_x, \quad j = 1, \ldots, N_y+1,$$

which both depend on the phenotypic state of the cell, j, the number of elements of the local environment and the total number of cells in the target site. In order to ensure that cells cannot move to a physical site "outside of the domain" $x \in [X_{\min}, X_{\max}]$, we assume that cells cannot move left out of site $i = 1$, or right out of site $i = N_x+1$, so that $\beta_-(j, N_0, e_0) = 0$ and $\beta_+(j, N_{N_x+2}, e_{N_x+2}) = 0$. For $i = 1, \ldots, N_x+1$, $j = 1, \ldots, N_y+1$, cells remain in their current site (*i.e.*, do not move) with probability

$$1 - \beta_+(j, N_{i+1}, e_{i+1}) - \beta_-(j, N_{i-1}, e_{i-1}) \in [0,1].$$

5.2.1.2 Cell proliferation

In order to model cell proliferation, we assume that a dividing cell is instantaneously replaced by two identical cells of equal volume to one another and the parent cell, such that the daughter cells inherit the spatial position and phenotypic state of the parent cell. As such, the corresponding change in state vector during a time step Δ_t can be written as

$$G_{i,j} : [n_1^j, \ldots, n_i^j, \ldots, n_{N_x+1}^j] \longrightarrow [n_1^j, \ldots, n_i^j - 1, \ldots, n_{N_x+1}^j], \quad \text{for} \quad i = 1, \ldots, N_x + 1, \quad j = 1, \ldots, N_y + 1.$$

To represent phenotype-dependent cell proliferation, we assume that the probability of a proliferation event is dependent on the phenotypic state of the cell, and the total number of cells and elements of the local environment in the same physical site as the cell that is dividing. Therefore, we define the probability that a cell in site i with phenotype j proliferates during time step Δ_t as

$$\gamma(j, N_i, e_i) \in [0, 1], \qquad i = 1, \ldots, N_x + 1, \quad j = 1, \ldots, N_y + 1.$$

The probability of a cell not undergoing proliferation during a time step Δ_t can then be written as

$$1 - \gamma(j, N_i, e_i) \in [0, 1], \qquad i = 1, \ldots, N_x + 1, \quad j = 1, \ldots, N_y + 1.$$

5.2.1.3 Cell phenotypic changes

During a single time step, Δ_t, we model transitions in phenotype space from state y_j to $y_{j\pm1}$ via the following changes in state vectors:

$$D_{i,j}^{\mathrm{p}} : [n_i^1, \ldots, n_i^{j-1}, n_i^j, \ldots, n_i^{N_y+1}] \longrightarrow [n_i^1, \ldots, n_i^{j-1} + 1, n_i^j - 1, \ldots, n_i^{N_y+1}], \quad \text{for} \quad i = 1, \ldots, N_x + 1, \quad j = 1, \ldots, N_y + 1,$$

$$U_{i,j}^{\mathrm{p}} : [n_i^1, \ldots, n_i^j, n_i^{j+1}, \ldots, n_i^{N_y+1}] \longrightarrow [n_i^1, \ldots, n_i^j - 1, n_i^{j+1} + 1, \ldots, n_i^{N_y+1}], \quad \text{for} \quad i = 1, \ldots, N_x + 1, \quad j = 1, \ldots, N_y + 1,$$

where $D_{i,j}^{\mathrm{p}}, U_{i,j}^{\mathrm{p}} : \mathbb{N}^{N_y+1} \to \mathbb{N}^{N_y+1}$. A cell in site i transitions from phenotypic state y_j to y_{j+1} during time step Δ_t with a probability that depends on the phenotypic state of the cell and the total number of cells and elements of the local environment in site i. Therefore, we can write that this transition, described by the change in state vector $U_{i,j}^{\mathrm{p}}$, occurs with a probability

$$\mu_+(j, N_i, e_i) \in [0, 1], \qquad i = 1, \ldots, N_x + 1, \quad j = 1, \ldots, N_y + 1.$$

Similarly, a cell in site i transitions from phenotypic state y_j to y_{j-1} during time step Δ_t with a probability that depends on the phenotypic state of the cell and the total number of cells and elements of the local environment in site i. Therefore, we can write that this transition, described by the change in state vector $D^{\mathrm{p}}_{i,j}$, occurs with probability

$$\mu_-(j, N_i, e_i) \in [0, 1], \qquad i = 1, \ldots, N_x + 1, \quad j = 1, \ldots, N_y + 1.$$

In order to ensure cells cannot transition to phenotypic states "outside of the domain" $y_j \in [Y_{\min}, Y_{\max}]$, we implement $\mu_-(1, N_i, e_i) = 0$ and $\mu_+(N_y + 1, N_i, e_i) = 0$. Taking this into consideration, the probability that a cell in phenotypic state y_j and spatial position x_i will not change phenotype during a time step Δ_t is given by

$$1 - \mu_+(j, N_i, e_i) - \mu_-(j, N_i, e_i) \in [0, 1], \qquad i = 1, \ldots, N_x + 1, \quad j = 1, \ldots, N_y + 1.$$

5.2.2 Modelling the dynamics of the local environment

We model degradation of elements of the local environment through contact with cells in the same physical site. Other cell-environment interactions, such as haptotaxis, or environmental changes such as production by cells, could also be considered here. The same methodology as presented in these sections can be followed to determine the corresponding population-level PDE. In particular, the change in state vector $H_i : \mathbb{N}^{N_x+1} \to \mathbb{N}^{N_x+1}$ describes degradation of an element of the local environment in spatial position x_i as

$$H_i : [e_1, \ldots, e_i, \ldots, e_{N_x+1}] \longrightarrow [e_1, \ldots, e_i + 1, \ldots, e_{N_x+1}], \qquad i = 1, \ldots, N_x + 1.$$

We assume that the probability of degradation of an element of the local environment during time step Δ_t depends on the number of cells in each phenotypic state j in the same spatial position x_i. As such, we define the probability of a cell in site i of phenotype j degrading an element of the local environment during a time step Δ_t as

$$\lambda(j, n_i^j) \in [0, 1], \qquad i = 1, \ldots, N_x + 1, \quad j = 1, \ldots, N_y + 1.$$

5.3 Formal derivation of the corresponding continuum model

In order to derive the corresponding continuum model describing the dynamics of the entire population of cells and the local environment over time, we now coarse-grain

this IBM. Assuming that the probability of two or more events occurring in time step Δ_t is sufficiently small that it can be ignored, the master equation, which describes the evolution of the probability density over time, is given by

$$
\begin{aligned}
&\Delta_t \frac{\partial}{\partial t} p(\mathbf{n}, \mathbf{e}, t_h) + \mathcal{O}(\Delta_t^2) \\
&= \sum_{i=1}^{N_x+1} \sum_{j=1}^{N_y} \mu_-(j+1, N_i, e_i) \left\{ (n_i^{j+1}+1) p(U_{i,j}^{\mathrm{p}} \mathbf{n}, \mathbf{e}, t_h) - n_i^{j+1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&\quad + \sum_{i=1}^{N_x+1} \sum_{j=2}^{N_y+1} \mu_+(j-1, N_i, e_i) \left\{ (n_i^{j-1}+1) p(D_{i,j}^{\mathrm{p}} \mathbf{n}, \mathbf{e}, t_h) - n_i^{j-1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&\quad + \sum_{i=1}^{N_x} \sum_{j=1}^{N_y+1} \beta_-(j, N_i, e_i) \left\{ (n_{i+1}^{j}+1) p(R_{i,j}^{\mathrm{m}} \mathbf{n}, \mathbf{e}, t_h) - n_{i+1}^{j} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&\quad + \sum_{i=2}^{N_x+1} \sum_{j=1}^{N_y+1} \beta_+(j, N_i, e_i) \left\{ (n_{i-1}^{j}+1) p(L_{i,j}^{\mathrm{m}} \mathbf{n}, \mathbf{e}, t_h) - n_{i-1}^{j} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&\quad + \sum_{i=1}^{N_x+1} \sum_{j=1}^{N_y+1} \left\{ \gamma(j, N_i - 1, e_i)(n_i^j - 1) p(G_{i,j} \mathbf{n}, \mathbf{e}, t_h) - \gamma(j, N_i, e_i) n_i^j p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&\quad + \sum_{i=1}^{N_x+1} \sum_{j=1}^{N_y+1} \lambda(j, n_i^j) \left\{ (e_i+1) p(\mathbf{n}, H_i \mathbf{e}, t_h) - e_i p(\mathbf{n}, \mathbf{e}, t_h) \right\}. \qquad (5.1)
\end{aligned}
$$

Briefly, the first two lines on the right-hand side correspond to changes in the phenotypic state of the cell, the second two correspond to changes in the physical position of the cell, the penultimate line describes proliferation of the cell and the final line describes degradation of the local environment.

5.3.1 Equation for the cell density

As is standard in the literature, we define the ensemble average for the function, f, of the number of cells at position $i = 1, \ldots, N_x+1$ in state $j = 1, \ldots, N_y+1$ and number of elements of local environment in lattice site $i = 1, \ldots, N_x+1$ in the following way:

$$
\langle f(n_i^j, e_i) \rangle = \sum_{\mathbf{n}} \sum_{\mathbf{e}} f(n_i^j, e_i) p(\mathbf{n}, \mathbf{e}, t_h). \qquad (5.2)
$$

Now, returning to Equation (5.1), we multiply by n_s^q, where $s = 1, \ldots, N_x+1$ and $q = 1, \ldots, N_y+1$, and take the sum over the vectors $\mathbf{n}$ and $\mathbf{e}$. We note that the final term in Equation (5.1) corresponds to the evolution of the number of elements of the local environment, and does not contribute to the cell dynamics. Then, to begin with,

we consider only the first remaining term on the right-hand side, which we denote

$$I = \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{i=1}^{N_x+1} \sum_{j=1}^{N_y} \mu_-(j+1, N_i, e_i) \left\{ (n_i^{j+1}+1) p(U_j^{\mathrm{p}} \mathbf{n}, \mathbf{e}, t_h) - n_i^{j+1} p(\mathbf{n}, \mathbf{e}, t_h) \right\}.$$

To continue, we must now consider the following cases in turn:

- I_1: $i \neq s$, $j \neq q, q-1$;
- I_2: $i \neq s$ and $j = q$;
- I_3: $i \neq s$ and $j = q-1$;
- I_4: $i = s$ and $j \neq q, q-1$;
- I_5: $i = s$ and $j = q$;
- I_6: $i = s$ and $j = q-1$.

We begin with the case I_1 where $i \neq s$ and $j \neq q, q-1$. In this scenario, we find

$$\begin{aligned} I_1 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \sum_{\substack{j=1,\\ j\neq q,q-1}}^{N_y} \mu_-(j+1, N_i, e_i) \left\{ (n_i^{j+1}+1) p(U_j^{\mathrm{p}} \mathbf{n}, \mathbf{e}, t_h) - n_i^{j+1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\ &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \sum_{\substack{j=1,\\ j\neq q,q-1}}^{N_y} \mu_-(j+1, N_i, e_i) \times \\ &\qquad \left\{ (n_i^{j+1}+1) p([n_i^1, \ldots, n_i^j - 1, n_i^{j+1}+1, \ldots, n_i^{N_y+1}], \mathbf{e}, t_h) - n_i^{j+1} p(\mathbf{n}, \mathbf{e}, t_h) \right\}. \end{aligned}$$

Under the change of variables $\bar{n}_i^j = n_i^j - 1$ and $\bar{n}_i^{j+1} = n_i^{j+1} + 1$ in the first term, we have that

$$\begin{aligned} I_1 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \sum_{\substack{j=1,\\ j\neq q,q-1}}^{N_y} \mu_-(j+1, N_i, e_i) \times \\ &\qquad \left\{ \bar{n}_i^{j+1} p([n_i^1, \ldots, \bar{n}_i^j, \bar{n}_i^{j+1}, \ldots, n_i^{N_y+1}], \mathbf{e}, t_h) - n_i^{j+1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\ &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \sum_{\substack{j=1,\\ j\neq q,q-1}}^{N_y} \mu_-(j+1, N_i, e_i) \left\{ n_i^{j+1} p(\mathbf{n}, \mathbf{e}, t_h) - n_i^{j+1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\ &= 0, \end{aligned}$$

where we can arbitrarily drop the bars after the change of variables. We will do this throughout what follows.

Next, we consider the case where $i \neq s$ and $j = q$, and apply the same change of variables argument (with $\bar{n}_i^q = n_i^q - 1$ and $\bar{n}_i^{q+1} = n_i^{q+1} + 1$) so that

$$\begin{aligned}
I_2 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \mu_-(q+1, N_i, e_i) \times \\
&\qquad \left\{ (n_i^{q+1}+1) p([n_i^1, \ldots, n_i^q - 1, n_i^{q+1}+1, \ldots, n_i^{N_y+1}], \mathbf{e}, t_h) - n_i^{q+1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \mu_-(q+1, N_i, e_i) \times \\
&\qquad \left\{ \bar{n}_i^{q+1} p([n_i^1, \ldots, \bar{n}_i^q, \bar{n}_i^{q+1}, \ldots, n_i^{N_y+1}], \mathbf{e}, t_h) - n_i^{q+1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \mu_-(q+1, N_i, e_i) \left\{ n_i^{q+1} p(\mathbf{n}, \mathbf{e}, t_h) - n_i^{q+1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= 0.
\end{aligned}$$

Finally we consider the last case for $i \neq s$, namely, where $j = q - 1$. Once again, by changing the variables using $\bar{n}_i^{q-1} = n_i^{q-1} - 1$ and $\bar{n}_i^q = n_i^q + 1$, we have

$$\begin{aligned}
I_3 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \mu_-(q, N_i, e_i) \times \\
&\qquad \left\{ (n_i^q+1) p([n_i^1, \ldots, n_i^{q-1} - 1, n_i^q+1, \ldots, n_i^{N_y+1}], \mathbf{e}, t_h) - n_i^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \mu_-(q, N_i, e_i) \times \\
&\qquad \left\{ \bar{n}_i^q p([n_i^1, \ldots, \bar{n}_i^{q-1}, \bar{n}_i^q, \ldots, n_i^{N_y+1}], \mathbf{e}, t_h) - n_i^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \mu_-(q, N_i, e_i) \left\{ n_i^q p(\mathbf{n}, \mathbf{e}, t_h) - n_i^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= 0.
\end{aligned}$$

As such, we can see that there are no contributions arising from the terms for which $i \neq s$.

Now we can consider the case where $i = s$. First off, we consider $j \neq q, q-1$, and

apply the change of variables $\bar{n}_s^j = n_s^j - 1$ and $\bar{n}_s^{j+1} = n_s^{j+1} + 1$:

$$\begin{aligned}
I_4 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{j=1,\\ j\neq q,q-1}}^{N_y} \mu_-(j+1, N_s, e_s)\times \\
&\qquad \left\{(n_s^{j+1}+1)p([n_s^1, \ldots, n_s^j - 1, n_s^{j+1}+1, \ldots, n_s^{N_y+1}], \mathbf{e}, t_h) - n_s^{j+1} p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{j=1,\\ j\neq q,q-1}}^{N_y} \mu_-(j+1, N_s, e_s)\times \\
&\qquad \left\{\bar{n}_s^{j+1} p([n_s^1, \ldots, \bar{n}_s^j, \bar{n}_s^{j+1}, \ldots, n_s^{N_y+1}], \mathbf{e}, t_h) - n_s^{j+1} p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{j=1,\\ j\neq q,q-1}}^{N_y} \mu_-(j+1, N_s, e_s) \left\{n_s^{j+1} p(\mathbf{n}, \mathbf{e}, t_h) - n_s^{j+1} p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&= 0.
\end{aligned}$$

Then, considering $i = s$ and $j = q$, we see that

$$\begin{aligned}
I_5 = \sum_{\mathbf{n}} \sum_{\mathbf{e}} & n_s^q \mu_-(q+1, N_s, e_s)\times \\
&\left\{(n_s^{q+1}+1)p([n_s^1, \ldots, n_s^q - 1, n_s^{q+1}+1, \ldots, n_s^{N_y+1}], \mathbf{e}, t_h) - n_s^{q+1} p(\mathbf{n}, \mathbf{e}, t_h)\right\}.
\end{aligned}$$

Now we want to separate the terms and apply the change of variables $\bar{n}_s^q = n_s^q - 1$ and $\bar{n}_s^{q+1} = n_s^{q+1} + 1$ in the first term to see

$$\begin{aligned}
I_5 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \mu_-(q+1, N_s, e_s)\times \\
&\qquad \left\{n_s^q (n_s^{q+1}+1)p([n_s^1, \ldots, n_s^q - 1, n_s^{q+1}+1, \ldots, n_s^{N_y+1}], \mathbf{e}, t_h)\right. \\
&\qquad\qquad \left. - n_s^q n_s^{q+1} p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \mu_-(q+1, N_s, e_s)\times \\
&\qquad \left\{(\bar{n}_s^q + 1)\bar{n}_s^{q+1} p([n_s^1, \ldots, \bar{n}_s^q, \bar{n}_s^{q+1}, \ldots, n_s^{N_y+1}], \mathbf{e}, t_h) - n_s^q n_s^{q+1} p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \mu_-(q+1, N_s, e_s) \left\{n_s^{q+1}(n_s^q + 1)p(\mathbf{n}, \mathbf{e}, t_h) - n_s^q n_s^{q+1} p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \mu_-(q+1, N_s, e_s) n_s^{q+1} p(\mathbf{n}, \mathbf{e}, t_h) \\
&= \langle \mu_-(q+1, N_s, e_s) n_s^{q+1} \rangle.
\end{aligned}$$

Finally, considering the case when $i = s$ and $j = q - 1$ we have

$$\begin{aligned}
I_6 = \sum_{\mathbf{n}} \sum_{\mathbf{e}} & n_s^q \mu_-(q, N_s, e_s)\times \\
&\left\{(n_s^q + 1)p([n_s^1, \ldots, n_s^{q-1} - 1, n_s^q + 1, \ldots, n_s^{N_y+1}], \mathbf{e}, t_h) - n_s^q p(\mathbf{n}, \mathbf{e}, t_h)\right\}.
\end{aligned}$$

Separating out the two terms and applying the change of variables $\bar{n}_s^q = n_s^q + 1$ and $\bar{n}_s^{q-1} = n_s^{q-1} - 1$ and then dropping the bars in the first term, we see

$$\begin{aligned}
I_6 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \mu_-(q, N_s, e_s) \times \\
&\qquad \left\{ n_s^q (n_s^q + 1) p([n_s^1, \ldots, n_s^{q-1} - 1, n_s^q + 1, \ldots, n_s^{N_y+1}], \mathbf{e}, t_h) - n_s^q n_s^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \mu_-(q, N_s, e_s) \times \\
&\qquad \left\{ (\bar{n}_s^q - 1) \bar{n}_s^q p([n_s^1, \ldots, \bar{n}_s^{q-1}, \bar{n}_s^q, \ldots, n_s^{N_y+1}], \mathbf{e}, t_h) - n_s^q n_s^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \mu_-(q, N_s, e_s) \left\{ n_s^q (n_s^q - 1) p(\mathbf{n}, \mathbf{e}, t_h) - (n_s^q)^2 p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= - \sum_{\mathbf{n}} \sum_{\mathbf{e}} \mu_-(q, N_s, e_s) n_s^q p(\mathbf{n}, \mathbf{e}, t_h) \\
&= - \langle \mu_-(q, N_s, e_s) n_s^q \rangle.
\end{aligned}$$

Putting these all back together, we find that

$$\begin{aligned}
I &= I_1 + I_2 + I_3 + I_4 + I_5 + I_6 \\
&= \langle \mu_-(q+1, N_s, e_s) n_s^{q+1} \rangle - \langle \mu_-(q, N_s, e_s) n_s^q \rangle.
\end{aligned}$$

Now we repeat this process for the second term in Equation (5.1), which we call J:

$$\begin{aligned}
J &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{i=1}^{N_x+1} \sum_{j=2}^{N_y+1} \mu_+(j-1, N_i, e_i) \left\{ (n_i^{j-1} + 1) p(D_{i,j}^{\mathrm{p}} \mathbf{n}, \mathbf{e}, t_h) - n_i^{j-1} p(\mathbf{n}, \mathbf{e}, t_h) \right\}, \\
&= J_1 + J_2 + J_3 + J_4 + J_5 + J_6
\end{aligned}$$

and consider the six cases: $J_1, \ldots, J_6$ (defined in the same way as before) where we have $i = s, i \neq s, j = q, j = q+1$, and $j \neq q, q+1$. Starting with $i \neq s$ and $j \neq q, q+1$

we see, using the change of variables $\bar{n}_i^j = n_i^j - 1$ and $\bar{n}_i^{j-1} = n_i^{j-1} + 1$, that we have

$$
\begin{aligned}
J_1 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \sum_{\substack{j=2,\\ j\neq q,q-1}}^{N_y+1} \mu_+(j-1, N_i, e_i) \left\{ (n_i^{j-1}+1) p(D_{i,j}^{\mathrm{p}} \mathbf{n}, \mathbf{e}, t_h) - n_i^{j-1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \sum_{\substack{j=2,\\ j\neq q,q-1}}^{N_y+1} \mu_+(j-1, N_i, e_i) \times \\
&\qquad \left\{ (n_i^{j-1}+1) p([n_i^1, \ldots, n_i^{j-1}+1, n_i^j - 1, \ldots, n_i^{N_y+1}], \mathbf{e}, t_h) - n_i^{j-1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \sum_{\substack{j=2,\\ j\neq q,q-1}}^{N_y+1} \mu_+(j-1, N_i, e_i) \times \\
&\qquad \left\{ \bar{n}_i^{j-1} p([n_i^1, \ldots, \bar{n}_i^{j-1}, \bar{n}_i^j, \ldots, n_i^{N_y+1}], \mathbf{e}, t_h) - n_i^{j-1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \sum_{\substack{j=2,\\ j\neq q,q-1}}^{N_y+1} \mu_+(j-1, N_i, e_i) \left\{ n_i^{j-1} p(\mathbf{n}, \mathbf{e}, t_h) - n_i^{j-1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= 0.
\end{aligned}
$$

Next, we consider the case when $i \neq s$ and $j = q$, and apply the same change of variables argument (with $\bar{n}_i^q = n_i^q - 1$ and $\bar{n}_i^{q-1} = n_i^{q-1} + 1$) to the first term:

$$
\begin{aligned}
J_2 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \mu_+(q-1, N_i, e_i) \times \\
&\qquad \left\{ (n_i^{q-1}+1) p([n_i^1, \ldots, n_i^{q-1}+1, n_i^q - 1, \ldots, n_i^{N_y+1}], \mathbf{e}, t_h) - n_i^{q-1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \mu_+(q-1, N_i, e_i) \times \\
&\qquad \left\{ \bar{n}_i^{q-1} p([n_i^1, \ldots, \bar{n}_i^{q-1}, \bar{n}_i^q, \ldots, n_i^{N_y+1}], \mathbf{e}, t_h) - n_i^{q-1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \mu_+(q-1, N_i, e_i) \left\{ n_i^{q-1} p(\mathbf{n}, \mathbf{e}, t_h) - n_i^{q-1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= 0.
\end{aligned}
$$

Finally we consider the last remaining situation when $i \neq s$, where $j = q + 1$. Once

again, by changing the variables using $\bar{n}_i^{q+1} = n_i^{q+1} - 1$ and $\bar{n}_i^q = n_i^q + 1$, we have

$$\begin{aligned}
J_3 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \mu_+(q, N_i, e_i) \times \\
&\qquad \left\{(n_i^q + 1)p([n_i^1, \ldots, n_i^q + 1, n_i^{q+1} - 1, \ldots, n_i^{N_y+1}], \mathbf{e}, t_h) - n_i^q p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \mu_+(q, N_i, e_i) \big\{\bar{n}_i^q p([n_i^1, \ldots, \bar{n}_i^q, \bar{n}_i^{q+1}, \ldots, n_i^{N_y+1}], \mathbf{e}, t_h) \\
&\qquad\qquad - n_i^q p(\mathbf{n}, \mathbf{e}, t_h)\big\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \mu_+(q, N_i, e_i) \left\{n_i^q p(\mathbf{n}, \mathbf{e}, t_h) - n_i^q p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&= 0.
\end{aligned}$$

As such, we can see that there are no contributions arising from the terms when $i \neq s$.

Now we can consider the case where $i = s$. First, we consider $j \neq q, q+1$, and apply the change of variables $\bar{n}_s^j = n_s^j - 1$ and $\bar{n}_s^{j-1} = n_s^{j-1} + 1$:

$$\begin{aligned}
J_4 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{j=2,\\ j\neq q, q-1}}^{N_y+1} \mu_+(j-1, N_s, e_s) \times \\
&\qquad \left\{(n_s^{j-1} + 1)p([n_s^1, \ldots, n_s^{j-1} + 1, n_s^j - 1, \ldots, n_s^{N_y+1}], \mathbf{e}, t_h) - n_s^{j-1} p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{j=2,\\ j\neq q, q-1}}^{N_y+1} \mu_+(j-1, N_s, e_s) \times \\
&\qquad \left\{\bar{n}_s^{j-1} p([n_s^1, \ldots, \bar{n}_s^{j-1}, \bar{n}_s^j, \ldots, n_s^{N_y+1}], \mathbf{e}, t_h) - n_s^{j-1} p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{j=2,\\ j\neq q, q-1}}^{N_y+1} \mu_+(j-1, N_s, e_s) \left\{n_s^{j-1} p(\mathbf{n}, \mathbf{e}, t_h) - n_s^{j-1} p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&= 0.
\end{aligned}$$

Then, considering the case when $i = s$ and $j = q$, we see that applying the change of

variables $\bar{n}_s^q = n_s^q - 1$ and $\bar{n}_s^{q-1} = n_s^{q-1} + 1$ in the first term, we have

$$\begin{aligned}
J_5 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \mu_+(q-1, N_s, e_s) \times \\
&\qquad \left\{ (n_s^{q-1}+1) p([n_s^1, \ldots, n_s^{q-1}+1, n_s^q - 1, \ldots, n_s^{N_y+1}], \mathbf{e}, t_h) - n_s^{q-1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \mu_+(q-1, N_s, e_s) \times \\
&\qquad \left\{ n_s^q (n_s^{q-1}+1) p([n_s^1, \ldots, n_s^{q-1}+1, n_s^q - 1, \ldots, n_s^{N_y+1}], \mathbf{e}, t_h) \right. \\
&\qquad\qquad \left. - n_s^q n_s^{q-1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \mu_+(q-1, N_s, e_s) \times \\
&\qquad \left\{ (\bar{n}_s^q + 1) \bar{n}_s^{q-1} p([n_s^1, \ldots, \bar{n}_s^{q-1}, \bar{n}_s^q, \ldots, n_s^{N_y+1}], \mathbf{e}, t_h) - n_s^q n_s^{q-1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \mu_+(q-1, N_s, e_s) \left\{ n_s^{q-1} (n_s^q + 1) p(\mathbf{n}, \mathbf{e}, t_h) - n_s^q n_s^{q-1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \mu_+(q-1, N_s, e_s) n_s^{q-1} p(\mathbf{n}, \mathbf{e}, t_h) \\
&= \langle \mu_+(q-1, N_s, e_s) n_s^{q-1} \rangle.
\end{aligned}$$

Finally, considering the case when $i = s$ and $j = q + 1$ we consider term J_6. Using the change of variables $\bar{n}_s^q = n_s^q + 1$ and $\bar{n}_s^{q+1} = n_s^{q+1} - 1$ and then dropping the bars in the first term, we see that

$$\begin{aligned}
J_6 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \mu_+(q, N_s, e_s) \times \\
&\qquad \left\{ (n_s^q + 1) p([n_s^1, \ldots, n_s^q + 1, n_s^{q+1} - 1, \ldots, n_s^{N_y+1}], \mathbf{e}, t_h) - n_s^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \mu_+(q, N_s, e_s) \times \\
&\qquad \left\{ n_s^q (n_s^q + 1) p([n_s^1, \ldots, n_s^q + 1, n_s^{q+1} - 1, \ldots, n_s^{N_y+1}], \mathbf{e}, t_h) - n_s^q n_s^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \mu_+(q, N_s, e_s) \times \\
&\qquad \left\{ (\bar{n}_s^q - 1) \bar{n}_s^q p([n_s^1, \ldots, \bar{n}_s^q, \bar{n}_s^{q+1}, \ldots, n_s^{N_y+1}], \mathbf{e}, t_h) - n_s^q n_s^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \mu_+(q, N_s, e_s) \left\{ n_s^q (n_s^q - 1) p(\mathbf{n}, \mathbf{e}, t_h) - (n_s^q)^2 p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= - \sum_{\mathbf{n}} \sum_{\mathbf{e}} \mu_+(q, N_s, e_s) n_s^q p(\mathbf{n}, \mathbf{e}, t_h) \\
&= - \langle \mu_+(q, N_s, e_s) n_s^q \rangle.
\end{aligned}$$

Putting these all back together, it is clear that

$$\begin{aligned}
J &= J_1 + J_2 + J_3 + J_4 + J_5 + J_6 \\
&= \langle \mu_+(q-1, N_s, e_s) n_s^{q-1} \rangle - \langle \mu_+(q, N_s, e_s) n_s^q \rangle.
\end{aligned}$$

Next, we consider the terms modelling movement in physical space in time step Δ_t. Returning to Equation (5.1), we look first at the third term on the right-hand side, which we call K. For this term, we once again consider the cases $j = q, j \neq q, i = s, i = s-1$ and $i \neq s, s-1$ in turn, so

$$\begin{aligned} K &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{i=1}^{N_x} \sum_{j=1}^{N_y+1} \beta_-(j, N_i, e_i) \left\{ (n_{i+1}^j + 1) p(R_{i,j}^{\mathrm{m}} \mathbf{n}, \mathbf{e}, t_h) - n_{i+1}^j p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\ &= K_1 + K_2 + K_3 + K_4 + K_5 + K_6. \end{aligned}$$

Considering first the case where $j \neq q$ and $i \neq s, s-1$, and applying the change of variables $\bar{n}_i^j = n_i^j - 1$ and $\bar{n}_{i+1}^j = n_{i+1}^j + 1$ and dropping the bars, we have

$$\begin{aligned} K_1 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1, \\ i \neq s, s-1}}^{N_x} \sum_{\substack{j=1, \\ j \neq q}}^{N_y+1} \beta_-(j, N_i, e_i) \times \\ &\qquad \left\{ (n_{i+1}^j + 1) p([n_1^j, \ldots, n_i^j - 1, n_{i+1}^j + 1, \ldots, n_{N_x+1}^j], \mathbf{e}, t_h) - n_{i+1}^j p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\ &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1, \\ i \neq s, s-1}}^{N_x} \sum_{\substack{j=1, \\ j \neq q}}^{N_y+1} \beta_-(j, N_i, e_i) \times \\ &\qquad \left\{ \bar{n}_{i+1}^j p([n_1^j, \ldots, \bar{n}_i^j, \bar{n}_{i+1}^j, \ldots, n_{N_x+1}^j], \mathbf{e}, t_h) - n_{i+1}^j p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\ &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1, \\ i \neq s, s-1}}^{N_x} \sum_{\substack{j=1, \\ j \neq q}}^{N_y+1} \beta_-(j, N_i, e_i) \left\{ n_{i+1}^j p(\mathbf{n}, \mathbf{e}, t_h) - n_{i+1}^j p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\ &= 0. \end{aligned}$$

Now consider $j \neq q$ with $i = s$. Applying the change of variables $\bar{n}_s^j = n_s^j - 1$ and $\bar{n}_{s+1}^j = n_{s+1}^j + 1$ and dropping the bars, we have

$$\begin{aligned} K_2 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{j=1, \\ j \neq q}}^{N_y+1} \beta_-(j, N_s, e_s) \times \\ &\qquad \left\{ (n_{s+1}^j + 1) p([n_1^j, \ldots, n_s^j - 1, n_{s+1}^j + 1, \ldots, n_{N_x+1}^j], \mathbf{e}, t_h) - n_{s+1}^j p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\ &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{j=1, \\ j \neq q}}^{N_y+1} \beta_-(j, N_s, e_s) \times \\ &\qquad \left\{ \bar{n}_{s+1}^j p([n_1^j, \ldots, \bar{n}_s^j, \bar{n}_{s+1}^j, \ldots, n_{N_x+1}^j], \mathbf{e}, t_h) - n_{s+1}^j p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\ &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{j=1, \\ j \neq q}}^{N_y+1} \beta_-(j, N_s, e_s) \left\{ n_{s+1}^j p(\mathbf{n}, \mathbf{e}, t_h) - n_{s+1}^j p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\ &= 0. \end{aligned}$$

Next we look at $j \neq q$ with $i = s-1$. If we apply the change of variables in the first term and drop the bars ($\bar{n}_s^j = n_s^j + 1$ and $\bar{n}_{s-1}^j = n_{s-1}^j - 1$), we have

$$\begin{aligned}
K_3 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{j=1,\\ j\neq q}}^{N_y+1} \beta_-(j, N_{s-1}, e_{s-1}) \times \\
&\qquad \left\{ (n_s^j + 1) p([n_1^j, \ldots, n_{s-1}^j - 1, n_s^j + 1, \ldots, n_{N_x+1}^j], \mathbf{e}, t_h) - n_s^j p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{j=1,\\ j\neq q}}^{N_y+1} \beta_-(j, N_{s-1}, e_{s-1}) \times \\
&\qquad \left\{ \bar{n}_s^j p([n_1^j, \ldots, \bar{n}_{s-1}^j, \bar{n}_s^j, \ldots, n_{N_x+1}^j], \mathbf{e}, t_h) - n_s^j p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{j=1,\\ j\neq q}}^{N_y+1} \beta_-(j, N_{s-1}, e_{s-1}) \left\{ n_s^j p(\mathbf{n}, \mathbf{e}, t_h) - n_s^j p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= 0.
\end{aligned}$$

Now we consider the three cases when $j = q$. First, when $i \neq s, s-1$, we apply the change of variables $\bar{n}_i^q = n_i^q - 1$ and $\bar{n}_{i+1}^q = n_{i+1}^q + 1$ and dropping the bars, we have

$$\begin{aligned}
K_4 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s,s-1}}^{N_x} \beta_-(q, N_i, e_i) \times \\
&\qquad \left\{ (n_{i+1}^q + 1) p([n_1^q, \ldots, n_i^q - 1, n_{i+1}^q + 1, \ldots, n_{N_x+1}^q], \mathbf{e}, t_h) - n_{i+1}^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s,s-1}}^{N_x} \beta_-(q, N_i, e_i) \times \\
&\qquad \left\{ \bar{n}_{i+1}^q p([n_1^q, \ldots, \bar{n}_i^q, \bar{n}_{i+1}^q, \ldots, n_{N_x+1}^q], \mathbf{e}, t_h) - n_{i+1}^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s,s-1}}^{N_x} \beta_-(q, N_i, e_i) \left\{ n_{i+1}^q p(\mathbf{n}, \mathbf{e}, t_h) - n_{i+1}^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= 0.
\end{aligned}$$

Now consider $i = s$ and the change of variables $\bar{n}_s^q = n_s^q - 1$ and $\bar{n}_{s+1}^q = n_{s+1}^q + 1$ to

give

$$
\begin{aligned}
K_5 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \beta_-(q, N_s, e_s) \times \\
&\qquad \left\{ (n_{s+1}^q + 1) p([n_1^q, \ldots, n_s^q - 1, n_{s+1}^q + 1, \ldots, n_{N_x+1}^q], \mathbf{e}, t_h) - n_{s+1}^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \beta_-(q, N_s, e_s) \times \\
&\qquad \left\{ (\bar{n}_s^q + 1) \bar{n}_{s+1}^q p([n_1^q, \ldots, \bar{n}_s^q, \bar{n}_{s+1}^q, \ldots, n_{N_x+1}^q], \mathbf{e}, t_h) - n_s^q n_{s+1}^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \beta_-(q, N_s, e_s) \left\{ (n_s^q + 1) n_{s+1}^q p(\mathbf{n}, \mathbf{e}, t_h) - n_s^q n_{s+1}^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \beta_-(q, N_s, e_s) n_{s+1}^q p(\mathbf{n}, \mathbf{e}, t_h) \\
&= \langle \beta_-(q, N_s, e_s) n_{s+1}^q \rangle.
\end{aligned}
$$

Finally, consider $i = s-1$ with the change of variables $\bar{n}_{s-1}^q = n_{s-1}^q - 1$ and $\bar{n}_s^q = n_s^q + 1$ to give

$$
\begin{aligned}
K_6 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \beta_-(q, N_{s-1}, e_{s-1}) \times \\
&\qquad \left\{ (n_s^q + 1) p([n_1^q, \ldots, n_{s-1}^q - 1, n_s^q + 1, \ldots, n_{N_x+1}^q], \mathbf{e}, t_h) - n_s^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \beta_-(q, N_{s-1}, e_{s-1}) \times \\
&\qquad \left\{ (\bar{n}_s^q - 1) \bar{n}_s^q p([n_1^q, \ldots, \bar{n}_{s-1}^q, \bar{n}_s^q, \ldots, n_{N_x+1}^q], \mathbf{e}, t_h) - n_s^q n_s^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \beta_-(q, N_{s-1}, e_{s-1}) \left\{ (n_s^q - 1) n_s^q p(\mathbf{n}, \mathbf{e}, t_h) - n_s^q n_s^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= - \sum_{\mathbf{n}} \sum_{\mathbf{e}} \beta_-(q, N_{s-1}, e_{s-1}) n_s^q p(\mathbf{n}, \mathbf{e}, t_h) \\
&= - \langle \beta_-(q, N_{s-1}, e_{s-1}) n_s^q \rangle.
\end{aligned}
$$

Bringing back together these terms, we have

$$
\begin{aligned}
K &= K_1 + K_2 + K_3 + K_4 + K_5 + K_6 \\
&= \langle \beta_-(q, N_s, e_s) n_{s+1}^q \rangle - \langle \beta_-(q, N_{s-1}, e_{s-1}) n_s^q \rangle.
\end{aligned}
$$

For the second movement term, we repeat this process, by breaking it down into six cases, defined by $j = q, j \neq q, i = s, i = s+1$ and $i \neq s, s+1$:

$$
\begin{aligned}
\mathcal{L} &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{i=2}^{N_x+1} \sum_{j=1}^{N_y+1} \beta_+(j, N_i, e_i) \left\{ (n_{i-1}^j + 1) p(L_{i,j}^{\mathrm{m}} \mathbf{n}, \mathbf{e}, t_h) - n_{i-1}^j p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 + \mathcal{L}_4 + \mathcal{L}_5 + \mathcal{L}_6.
\end{aligned}
$$

We start by considering $j \neq q$ and $i \neq s, s+1$. Employing the change of variables $\bar{n}_i^j = n_i^j - 1$ and $\bar{n}_{i-1}^j = n_{i-1}^j + 1$ and dropping the bars, we have

$$\begin{aligned}
\mathcal{L}_1 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s,s+1}}^{N_x} \sum_{\substack{j=1,\\ j\neq q}}^{N_y+1} \beta_+(j, N_i, e_i) \times \\
&\qquad \left\{ (n_{i-1}^j + 1) p([n_1^j, \ldots, n_{i-1}^j + 1, n_i^j - 1, \ldots, n_{N_x+1}^j], \mathbf{e}, t_h) - n_{i-1}^j p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s,s+1}}^{N_x} \sum_{\substack{j=1,\\ j\neq q}}^{N_y+1} \beta_+(j, N_i, e_i) \times \\
&\qquad \left\{ \bar{n}_{i-1}^j p([n_1^j, \ldots, \bar{n}_{i-1}^j, \bar{n}_i^j, \ldots, n_{N_x+1}^j], \mathbf{e}, t_h) - n_{i-1}^j p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s,s+1}}^{N_x} \sum_{\substack{j=1,\\ j\neq q}}^{N_y+1} \beta_+(j, N_i, e_i) \left\{ n_{i-1}^j p(\mathbf{n}, \mathbf{e}, t_h) - n_{i-1}^j p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= 0.
\end{aligned}$$

Now consider $j \neq q$ with $i = s$. Then, applying the change of variables $\bar{n}_s^j = n_s^j - 1$ and $\bar{n}_{s-1}^j = n_{s-1}^j + 1$ and dropping the bars, we have

$$\begin{aligned}
\mathcal{L}_2 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{j=1,\\ j\neq q}}^{N_y+1} \beta_+(j, N_s, e_s) \times \\
&\qquad \left\{ (n_{s-1}^j + 1) p([n_1^j, \ldots, n_{s-1}^j + 1, n_s^j - 1, \ldots, n_{N_x+1}^j], \mathbf{e}, t_h) - n_{s+1}^j p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{j=1,\\ j\neq q}}^{N_y+1} \beta_+(j, N_s, e_s) \times \\
&\qquad \left\{ \bar{n}_{s-1}^j p([n_1^j, \ldots, \bar{n}_{s-1}^j, \bar{n}_s^j, \ldots, n_{N_x+1}^j], \mathbf{e}, t_h) - n_{s-1}^j p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{j=1,\\ j\neq q}}^{N_y+1} \beta_+(j, N_s, e_s) \left\{ n_{s+1}^j p(\mathbf{n}, \mathbf{e}, t_h) - n_{s-1}^j p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= 0.
\end{aligned}$$

Next we look at $j \neq q$ with $i = s + 1$. If we apply the change of variables in the first

term and drop the bars ($\bar{n}_s^j = n_s^j + 1$ and $\bar{n}_{s+1}^j = n_{s+1}^j - 1$), we have

$$\begin{aligned}
\mathcal{L}_3 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{j=1,\\ j\neq q}}^{N_y+1} \beta_+(j, N_{s-1}, e_{s-1}) \times \\
&\qquad \left\{(n_s^j + 1) p([n_1^j, \ldots, n_s^j + 1, n_{s+1}^j - 1, \ldots, n_{N_x+1}^j], \mathbf{e}, t_h) - n_s^j p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{j=1,\\ j\neq q}}^{N_y+1} \beta_+(j, N_{s-1}, e_{s-1}) \times \\
&\qquad \left\{\bar{n}_s^j p([n_1^j, \ldots, \bar{n}_s^j, \bar{n}_{s+1}^j, \ldots, n_{N_x+1}^j], \mathbf{e}, t_h) - n_s^j p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{j=1,\\ j\neq q}}^{N_y+1} \beta_+(j, N_{s-1}, e_{s-1}) \left\{n_s^j p(\mathbf{n}, \mathbf{e}, t_h) - n_s^j p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&= 0.
\end{aligned}$$

Now we consider the three cases when $j = q$. First, when $i \neq s, s+1$, applying the change of variables $\bar{n}_i^q = n_i^q - 1$ and $\bar{n}_{i-1}^q = n_{i-1}^q + 1$ and dropping the bars, we have

$$\begin{aligned}
\mathcal{L}_4 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s,s-1}}^{N_x} \beta_+(q, N_i, e_i) \times \\
&\qquad \left\{(n_{i-1}^q + 1) p([n_1^q, \ldots, n_{i-1}^q + 1, n_i^q - 1, \ldots, n_{N_x+1}^q], \mathbf{e}, t_h) - n_{i-1}^q p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s,s-1}}^{N_x} \beta_+(q, N_i, e_i) \times \\
&\qquad \left\{\bar{n}_{i-1}^q p([n_1^q, \ldots, \bar{n}_{i-1}^q, \bar{n}_i^q, \ldots, n_{N_x+1}^q], \mathbf{e}, t_h) - n_{i-1}^q p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s,s-1}}^{N_x} \beta_+(q, N_i, e_i) \left\{n_{i-1}^q p(\mathbf{n}, \mathbf{e}, t_h) - n_{i-1}^q p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&= 0.
\end{aligned}$$

Now for $i = s$ and the change of variables $\bar{n}_s^q = n_s^q - 1$ and $\bar{n}_{s-1}^q = n_{s-1}^q + 1$, we have

$$\begin{aligned}
\mathcal{L}_5 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \beta_+(q, N_s, e_s) \times \\
&\qquad \left\{ (n_{s-1}^q + 1) p([n_1^q, \ldots, n_{s-1}^q + 1, n_s^q - 1, \ldots, n_{N_x+1}^q], \mathbf{e}, t_h) - n_{s-1}^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \beta_+(q, N_s, e_s) \times \\
&\qquad \left\{ (\bar{n}_s^q + 1) \bar{n}_{s-1}^q p([n_1^q, \ldots, \bar{n}_{s-1}^q, \bar{n}_s^q, \ldots, n_{N_x+1}^q], \mathbf{e}, t_h) - n_s^q n_s^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \beta_+(q, N_s, e_s) \left\{ (n_s^q + 1) n_{s-1}^q p(\mathbf{n}, \mathbf{e}, t_h) - n_s^q n_{s-1}^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \beta_+(q, N_s, e_s) n_{s-1}^q p(\mathbf{n}, \mathbf{e}, t_h) \\
&= \langle \beta_+(q, N_s, e_s) n_{s-1}^q \rangle.
\end{aligned}$$

Finally, for $i = s + 1$ with the change of variables $\bar{n}_{s+1}^q = n_{s+1}^q - 1$ and $\bar{n}_s^q = n_s^q + 1$, we have

$$\begin{aligned}
\mathcal{L}_6 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \beta_+(q, N_{s+1}, e_{s+1}) \times \\
&\qquad \left\{ (n_s^q + 1) p([n_1^q, \ldots, n_s^q + 1, n_{s+1}^q - 1, \ldots, n_{N_x+1}^q], \mathbf{e}, t_h) - n_s^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \beta_+(q, N_{s+1}, e_{s+1}) \times \\
&\qquad \left\{ (\bar{n}_s^q - 1) \bar{n}_s^q p([n_1^q, \ldots, \bar{n}_s^q, \bar{n}_{s+1}^q, \ldots, n_{N_x+1}^q], \mathbf{e}, t_h) - n_s^q n_s^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \beta_+(q, N_{s+1}, e_{s+1}) \left\{ (n_s^q - 1) n_s^q p(\mathbf{n}, \mathbf{e}, t_h) - n_s^q n_s^q p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&= - \sum_{\mathbf{n}} \sum_{\mathbf{e}} \beta_+(q, N_{s+1}, e_{s+1}) n_s^q p(\mathbf{n}, \mathbf{e}, t_h) \\
&= - \langle \beta_+(q, N_{s+1}, e_{s+1}) n_s^q \rangle.
\end{aligned}$$

Bringing back together these terms, we have

$$\begin{aligned}
\mathcal{L} &= \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 + \mathcal{L}_4 + \mathcal{L}_5 + \mathcal{L}_6 \\
&= \langle \beta_+(q, N_s, e_s) n_{s-1}^q \rangle - \langle \beta_+(q, N_{s+1}, e_{s+1}) n_s^q \rangle.
\end{aligned}$$

Finally, we must consider the last term in Equation (5.1) (ignoring the degradation

terms, which do not contribute to cell dynamics)

$$
\begin{aligned}
M &= \sum_{\mathbf{n}}\sum_{\mathbf{e}} n_s^q \sum_{i=1}^{N_x+1}\sum_{j=1}^{N_y+1} \big\{\gamma(j,N_i-1,e_i)(n_i^j-1)p(G_{i,j}\mathbf{n},\mathbf{e},t_h) \\
&\qquad\qquad - \gamma(j,N_i,e_i)n_i^j p(\mathbf{n},\mathbf{e},t_h)\big\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} n_s^q \sum_{i=1}^{N_x+1}\sum_{j=1}^{N_y+1} \big\{\gamma(j,N_i-1,e_i)(n_i^j-1)p([n_1^j,\ldots,n_i^j-1,\ldots,n_{N_x+1}^j],\mathbf{e},t_h) \\
&\qquad\qquad - \gamma(j,N_i,e_i)n_i^j p(\mathbf{n},\mathbf{e},t_h)\big\}, \\
&= M_1+M_2+M_3+M_4.
\end{aligned}
$$

Again we consider the cases $i=s$, $i\neq s$, $j=q$, $j\neq q$. First, begin with $i\neq s$ and $j\neq q$, and use the change of variable $\bar{n}_i^j = n_i^j - 1$, noticing that $N_i - 1$ becomes $\bar{N}_i$ when we change the variable. Dropping the bars gives

$$
\begin{aligned}
M_1 &= \sum_{\mathbf{n}}\sum_{\mathbf{e}} n_s^q \sum_{i=1}^{N_x+1}\sum_{j=1}^{N_y+1} \big\{\gamma(j,N_i-1,e_i)(n_i^j-1)p([n_1^j,\ldots,n_i^j-1,\ldots,n_{N_x+1}^j],\mathbf{e},t_h) \\
&\qquad\qquad - \gamma(j,N_i,e_i)n_i^j p(\mathbf{n},\mathbf{e},t_h)\big\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1}\sum_{\substack{j=1,\\ j\neq q}}^{N_y+1} \big\{\gamma(j,\bar{N}_i,e_i)\bar{n}_i{}^j p([n_1^j,\ldots,\bar{n}_i^j,\ldots,n_{N_x+1}^j],\mathbf{e},t_h) \\
&\qquad\qquad - \gamma(j,N_i,e_i)n_i^j p(\mathbf{n},\mathbf{e},t_h)\big\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1}\sum_{\substack{j=1,\\ j\neq q}}^{N_y+1} \gamma(j,N_i,e_i)\left\{n_i^j p(\mathbf{n},\mathbf{e},t_h) - n_i^j p(\mathbf{n},\mathbf{e},t_h)\right\} \\
&= 0.
\end{aligned}
$$

Now consider $i\neq s$ and $j=q$. Using the change of variable $\bar{n}_i^q = n_i^q - 1$, along with

the newly defined $\bar{N}_i$, before dropping the bars we have

$$\begin{aligned}
M_2 &= \sum_{\mathbf{n}}\sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \big\{\gamma(q, N_i-1, e_i)(n_i^q-1)p([n_1^q,\ldots,n_i^q-1,\ldots,n_{N_x+1}^q],\mathbf{e},t_h) \\
&\qquad\qquad - \gamma(q,N_i,e_i)n_i^q p(\mathbf{n},\mathbf{e},t_h)\big\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \big\{\gamma(q, \bar{N}_i, e_i)\bar{n}_i^q p([n_1^q,\ldots,\bar{n}_i^q,\ldots,n_{N_x+1}^q],\mathbf{e},t_h) \\
&\qquad\qquad - \gamma(q,N_i,e_i)n_i^q p(\mathbf{n},\mathbf{e},t_h)\big\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} n_s^q \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1} \gamma(q,N_i,e_i)\left\{n_i^q p(\mathbf{n},\mathbf{e},t_h) - n_i^q p(\mathbf{n},\mathbf{e},t_h)\right\} \\
&= 0.
\end{aligned}$$

Now take $i = s$ and consider $j \neq q$. Using the following change of variables, $\bar{n}_s^j = n_s^j - 1$ and $\bar{N}_s = N_s - 1$, we see that

$$\begin{aligned}
M_3 &= \sum_{\mathbf{n}}\sum_{\mathbf{e}} n_s^q \sum_{\substack{j=1,\\ j\neq q}}^{N_y+1} \big\{\gamma(j, N_s-1, e_s)(n_s^j-1)p([n_1^j,\ldots,n_s^j-1,\ldots,n_{N_x+1}^j],\mathbf{e},t_h) \\
&\qquad\qquad - \gamma(j,N_s,e_s)n_s^j p(\mathbf{n},\mathbf{e},t_h)\big\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} n_s^q \sum_{\substack{j=1,\\ j\neq q}}^{N_y+1} \big\{\gamma(j, \bar{N}_s, e_s)\bar{n}_s^j p([n_1^j,\ldots,\bar{n}_s^j,\ldots,n_{N_x+1}^j],\mathbf{e},t_h) \\
&\qquad\qquad - \gamma(j,N_s,e_s)n_s^j p(\mathbf{n},\mathbf{e},t_h)\big\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} n_s^q \sum_{\substack{j=1,\\ j\neq q}}^{N_y+1} \gamma(j,N_s,e_s)\left\{n_s^j p(\mathbf{n},\mathbf{e},t_h) - n_s^j p(\mathbf{n},\mathbf{e},t_h)\right\} \\
&= 0.
\end{aligned}$$

Finally, consider $i = s$ and $j = q$ with the change of variable $\bar{n}_s^q = n_s^q - 1$ along with

$\bar{N}_s$ in the second term, and then drop the bars to give

$$\begin{aligned}
M_4 &= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \big\{ \gamma(q, N_s - 1, e_s)(n_s^q - 1) p([n_1^q, \ldots, n_s^q - 1, \ldots, n_{N_x+1}^q], \mathbf{e}, t_h) \\
&\qquad\qquad - \gamma(q, N_s, e_s) n_s^q p(\mathbf{n}, \mathbf{e}, t_h) \big\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} \big\{ \gamma(q, \bar{N}_s, e_s) \bar{n}_s^q (\bar{n}_s^q + 1) p([n_1^q, \ldots, \bar{n}_s^q, \ldots, n_{N_x+1}^q], \mathbf{e}, t_h) \\
&\qquad\qquad - \gamma(q, N_s, e_s)(n_s^q)^2 p(\mathbf{n}, \mathbf{e}, t_h) \big\} \\
&= \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^q \gamma(q, N_s, e_s) p(\mathbf{n}, \mathbf{e}, t_h) \\
&= \langle \gamma(q, N_s, e_s) n_s^q \rangle.
\end{aligned}$$

Combining all of these calculations, we can rewrite the master equation, Equation (5.1), as

$$\begin{aligned}
\frac{\partial}{\partial t}\langle n_s^q \rangle &= \frac{1}{\Delta_t}\langle \beta_+(q, N_s, e_s) n_{s-1}^q \rangle + \frac{1}{\Delta_t}\langle \beta_-(q, N_s, e_s) n_{s+1}^q \rangle \\
&\quad - \frac{1}{\Delta_t}\langle \beta_-(q, N_{s-1}, e_{s-1}) n_s^q \rangle - \frac{1}{\Delta_t}\langle \beta_+(q, N_{s+1}, e_{s+1}) n_s^q \rangle \\
&\quad + \frac{1}{\Delta_t}\langle \mu_+(q-1, N_s, e_s) n_s^{q-1} \rangle + \frac{1}{\Delta_t}\langle \mu_-(q+1, N_s, e_s) n_s^{q+1} \rangle \\
&\quad - \frac{1}{\Delta_t}\langle \mu_+(q, N_s, e_s) n_s^q \rangle - \frac{1}{\Delta_t}\langle \mu_-(q, N_s, e_s) n_s^q \rangle \\
&\quad + \frac{1}{\Delta_t}\langle \gamma(q, N_s, e_s) n_s^q \rangle. \qquad (5.3)
\end{aligned}$$

This mean equation is related to a PDE model in the appropriate limits as $\Delta_x \to 0$, $\Delta_y \to 0$ and $\Delta_t \to 0$ simultaneously, and the discrete values of $\langle n_i^j(t_h) \rangle$ and $\langle e_i(t_h) \rangle$ are written in terms of the continuous variables $n(x, y, t)$ and $e(x, t)$, describing the cell and local environment density, respectively, along with

$$\rho(x, t) = \int_{y=Y_{\min}}^{y=Y_{\max}} n(x, y, t) \mathrm{d}y,$$

describing the total cell density. Equation (5.3) can be rewritten as follows, using

mean-field approximations correct to $\mathcal{O}(\Delta t)$,

$$\begin{aligned}
\frac{\partial n(x,y,t)}{\partial t} = & \frac{1}{\Delta_t}\beta_+(y,\rho(x,t),e(x,t))n(x-\Delta_x,y,t) \\
& + \frac{1}{\Delta_t}\beta_-(y,\rho(x,t),e(x,t))n(x+\Delta_x,y,t) \\
& - \frac{1}{\Delta_t}\beta_-(y,\rho(x-\Delta_x,t),e(x-\Delta_x,t))n(x,y,t) \\
& - \frac{1}{\Delta_t}\beta_+(y,\rho(x+\Delta_x,t),e(x+\Delta_x,t))n(x,y,t) \\
& + \frac{1}{\Delta_t}\mu_+(y-\Delta_y,\rho(x,t),e(x,t))n(x,y-\Delta_y,t) \\
& + \frac{1}{\Delta_t}\mu_-(y+\Delta_y,\rho(x,t),e(x,t))n(x,y+\Delta_y,t) \\
& - \frac{1}{\Delta_t}\mu_+(y,\rho(x,t),e(x,t))n(x,y,t) \\
& - \frac{1}{\Delta_t}\mu_-(y,\rho(x,t),e(x,t))n(x,y,t) \\
& + \frac{1}{\Delta_t}\gamma(y,\rho(x,t),e(x,t))n(x,y,t). \qquad (5.4)
\end{aligned}$$

Then, we can employ Taylor series expansions in Equation (5.4) and take limits to

give

$$\begin{aligned}
\frac{\partial n(x,y,t)}{\partial t} &= \frac{1}{\Delta_t}\beta_+(y,\rho(x,t),e(x,t))\left[n(x,y,t) - \Delta_x\frac{\partial}{\partial x}n(x,y,t) + \frac{\Delta_x^2}{2}\frac{\partial^2}{\partial x^2}n(x,y,t)\right] \\
&+ \frac{1}{\Delta_t}\beta_-(y,\rho(x,t),e(x,t))\left[n(x,y,t) + \Delta_x\frac{\partial}{\partial x}n(x,y,t) + \frac{\Delta_x^2}{2}\frac{\partial^2}{\partial x^2}n(x,y,t)\right] \\
&- \frac{1}{\Delta_t}\left[\beta_-(y,\rho,e) - \Delta_x\frac{\partial}{\partial x}\beta_-(y,\rho,e) + \frac{\Delta_x^2}{2}\frac{\partial^2}{\partial x^2}\beta_-(y,\rho,e)\right]n(x,y,t) \\
&- \frac{1}{\Delta_t}\left[\beta_+(y,\rho,e) + \Delta_x\frac{\partial}{\partial x}\beta_+(y,\rho,e) + \frac{\Delta_x^2}{2}\frac{\partial^2}{\partial x^2}\beta_+(y,\rho,e)\right]n(x,y,t) \\
&+ \frac{1}{\Delta_t}\left[\mu_+(y,\rho,e) - \Delta_y\frac{\partial}{\partial y}\mu_+(y,\rho,e) + \frac{\Delta_y^2}{2}\frac{\partial^2}{\partial y^2}\mu_+(y,\rho,e)\right]\times \\
&\qquad\left[n(x,y,t) - \Delta_y\frac{\partial}{\partial y}n(x,y,t) + \frac{\Delta_y^2}{2}\frac{\partial^2}{\partial y^2}n(x,y,t)\right] \\
&+ \frac{1}{\Delta_t}\left[\mu_-(y,\rho,e) + \Delta_y\frac{\partial}{\partial y}\mu_-(y,\rho,e) + \frac{\Delta_y^2}{2}\frac{\partial^2}{\partial y^2}\mu_-(y,\rho,e)\right]\times \\
&\qquad\left[n(x,y,t) + \Delta_y\frac{\partial}{\partial y}n(x,y,t) + \frac{\Delta_y^2}{2}\frac{\partial^2}{\partial y^2}n(x,y,t)\right] \\
&- \frac{1}{\Delta_t}\mu_+(y,\rho(x,t),e(x,t))n(x,y,t) - \frac{1}{\Delta_t}\mu_-(y,\rho(x,t),e(x,t))n(x,y,t) \\
&+ \frac{1}{\Delta_t}\gamma(y,\rho(x,t),e(x,t))n(x,y,t) + \mathcal{O}(\Delta_x^3) + \mathcal{O}(\Delta_y^3) + \mathcal{O}(\Delta_t^2).
\end{aligned}$$

Rearranging and collecting terms, we obtain

$$\begin{aligned}
\frac{\partial}{\partial t}n(x,y,t) &= \frac{\Delta_x}{\Delta_t}\frac{\partial}{\partial x}\Big(\left(\beta_-(y,\rho(x,t),e(x,t)) - \beta_+(y,\rho(x,t),e(x,t))\right)n\Big) \\
&\quad + \frac{\Delta_x^2}{2\Delta_t}\frac{\partial}{\partial x}\Bigg(\Big(\beta_-\left(y,\rho(x,t),e(x,t)\right) + \beta_+\left(y,\rho(x,t),e(x,t)\right)\Big)\frac{\partial}{\partial x}n(x,y,t) \\
&\qquad\qquad - n(x,y,t)\frac{\partial}{\partial x}\Big(\beta_-\left(y,\rho(x,t),e(x,t)\right) \\
&\qquad\qquad\qquad + \beta_+\left(y,\rho(x,t),e(x,t)\right)\Big)\Bigg) \\
&\quad + \frac{\Delta_y}{\Delta_t}\frac{\partial}{\partial y}\Big(\Big(\mu_-\left(y,\rho(x,t),e(x,t)\right) - \mu_+\left(y,\rho(x,t),e(x,t)\right)\Big)n(x,y,t)\Big) \\
&\quad + \frac{\Delta_y^2}{2\Delta_t}\frac{\partial^2}{\partial y^2}\Big(\Big(\mu_-\left(y,\rho(x,t),e(x,t)\right) + \mu_+\left(y,\rho(x,t),e(x,t)\right)\Big)n(x,y,t)\Big) \\
&\quad + \frac{1}{\Delta_t}\gamma\left(y,\rho(x,t),e(x,t)\right)n(x,y,t).
\end{aligned}$$

We take the parabolic limit as Δ_x, Δ_y, $\Delta_t \to 0$ simultaneously (assuming $n(x,y,t) \sim \mathcal{O}(1)$, and the associated constraints on the functions are satisfied), and define

$$\lim_{\Delta_x,\Delta_t\to 0} \frac{\Delta_x}{\Delta_t}\Big(\beta_-(y,\rho(x,t),e(x,t)) - \beta_+(y,\rho(x,t),e(x,t))\Big) = v^m(y,\rho(x,t),e(x,t)), \tag{5.5}$$

$$\lim_{\Delta_x,\Delta_t\to 0} \frac{\Delta_x^2}{2\Delta_t}\Big(\beta_-(y,\rho(x,t),e(x,t)) + \beta_+(y,\rho(x,t),e(x,t))\Big) = D^m(y,\rho(x,t),e(x,t)), \tag{5.6}$$

$$\lim_{\Delta_y,\Delta_t\to 0} \frac{\Delta_y}{\Delta_t}\Big(\mu_-\left(y,\rho(x,t),e(x,t)\right) - \mu_+\left(y,\rho(x,t),e(x,t)\right)\Big) = v^p(y,\rho(x,t),e(x,t)), \tag{5.7}$$

$$\lim_{\Delta_y,\Delta_t\to 0} \frac{\Delta_y^2}{2\Delta_t}\Big(\mu_-\left(y,\rho(x,t),e(x,t)\right) + \mu_+\left(y,\rho(x,t),e(x,t)\right)\Big) = D^p(y,\rho(x,t),e(x,t)), \tag{5.8}$$

$$\lim_{\Delta_t\to 0} \frac{1}{\Delta_t}\gamma\left(y,\rho(x,t),e(x,t)\right) = r(y,\rho(x,t),e(x,t)). \tag{5.9}$$

The final equation for the evolution of the cell density is therefore given by

$$\begin{aligned}\frac{\partial}{\partial t}n(x,y,t) = &\frac{\partial}{\partial x}\Big(v^m(y,\rho(x,t),e(x,t))n\Big)\\ &+\frac{\partial}{\partial x}\Bigg(D^m\left(y,\rho(x,t),e(x,t)\right)\frac{\partial}{\partial x}n(x,y,t)\\ &\qquad - n(x,y,t)\frac{\partial}{\partial x}D^m\left(y,\rho(x,t),e(x,t)\right)\Bigg)\\ &+\frac{\partial}{\partial y}\left(v^p\left(y,\rho(x,t),e(x,t)\right)n(x,y,t)\right)\\ &+\frac{\partial^2}{\partial y^2}\left(D^p\left(y,\rho(x,t),e(x,t)\right)n(x,y,t)\right)\\ &+r\left(y,\rho(x,t),e(x,t)\right)n(x,y,t).\end{aligned} \tag{5.10}$$

The differential equation governing the cell population evolution over time is complemented with the initial condition

$$n(x,y,0) = n_0(x,y), \tag{5.11}$$

and boundary conditions that we define in the following section.

5.3.1.1 Model equations on the boundaries

We can repeat the above analysis for the boundaries of physical and phenotype space in order to retrieve the boundary equations. We note here that when looking at the

boundaries for the cell equation, all terms involving change in the number of elements of the local environment provide no contribution and can be ignored hereon in.

Boundary condition at $x = X_{\min}$. Revisiting Equation (5.1) to derive the equation on the left-most lattice site, we multiply by n_1^q and sum over all possible states $\mathbf{n}$ and $\mathbf{e}$:

$$\begin{aligned}
&\Delta_t \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_1^q \frac{\partial}{\partial t} p(\mathbf{n}, \mathbf{e}, t_h) + \mathcal{O}(\Delta_t^2) \\
&\quad = \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_1^q \sum_{i=1}^{N_x+1} \sum_{j=1}^{N_y} \mu_-(j+1, N_i, e_i) \left\{ (n_i^{j+1}+1) p(U^{\mathrm{p}}_{i,j}\mathbf{n}, \mathbf{e}, t_h) - n_i^{j+1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&\quad + \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_1^q \sum_{i=1}^{N_x+1} \sum_{j=2}^{N_y+1} \mu_+(j-1, N_i, e_i) \left\{ (n_i^{j-1}+1) p(D^{\mathrm{p}}_{i,j}\mathbf{n}, \mathbf{e}, t_h) - n_i^{j-1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&\quad + \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_1^q \sum_{i=1}^{N_x} \sum_{j=1}^{N_y+1} \beta_-(j, N_i, e_i) \left\{ (n_{i+1}^{j}+1) p(R^{\mathrm{m}}_{i,j}\mathbf{n}, \mathbf{e}, t_h) - n_{i+1}^{j} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&\quad + \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_1^q \sum_{i=2}^{N_x+1} \sum_{j=1}^{N_y+1} \beta_+(j, N_i, e_i) \left\{ (n_{i-1}^{j}+1) p(L^{\mathrm{m}}_{i,j}\mathbf{n}, \mathbf{e}, t_h) - n_{i-1}^{j} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&\quad + \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_1^q \sum_{i=1}^{N_x+1} \sum_{j=1}^{N_y+1} \Big\{ \gamma(j, N_i - 1, e_i)(n_i^j - 1) p(G_{i,j}\mathbf{n}, \mathbf{e}, t_h) \\
&\qquad\qquad - \gamma(j, N_i, e_i) n_i^j p(\mathbf{n}, \mathbf{e}, t_h) \Big\}.
\end{aligned} \tag{5.12}$$

From earlier working, we know that the terms describing movement in physical space only give non-zero contributions for certain choices of j. As such, we only consider these going forward. Starting by considering the first term on the right-hand side of Equation (5.12), we consider the non-zero contributions from the cases $j = q$ and $j = q - 1$. Using $j = q$:

$$\begin{aligned}
&\sum_{\mathbf{n}} \sum_{\mathbf{e}} n_1^q \sum_{i=1}^{N_x+1} \mu_-(q+1, N_i, e_i) \left\{ (n_i^{q+1}+1) p(U^{\mathrm{p}}_{i,q}\mathbf{n}, \mathbf{e}, t_h) - n_i^{q+1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&\quad = \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_1^q \sum_{i=1}^{N_x+1} \mu_-(q+1, N_i, e_i) \times \\
&\qquad \left\{ (n_i^{q+1}+1) p([n_i^1, \ldots, n_i^q - 1, n_i^{q+1} + 1, \ldots, n_i^{N_y+1}], \mathbf{e}, t_h) - n_i^{q+1} p(\mathbf{n}, \mathbf{e}, t_h) \right\}.
\end{aligned} \tag{5.13}$$

As per previous calculations, the only contributions here will come from terms where $i = 1$. Looking at these, we see that if we employ the change of variables $\bar{n}_1^{q+1} =$

$n_1^{q+1}+1$ and $\bar{n}_1^q = n_1^q - 1$ in the first term, then we can rewrite the right-hand side of Equation (5.13) as

$$\begin{aligned}
&\sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_-(q+1, N_1, e_1)\times \\
&\qquad \left\{\bar{n}_1^{q+1}(\bar{n}_1^q+1)p([n_1^1,\ldots,\bar{n}_1^q,\bar{n}_1^{q+1},\ldots,n_1^{N_y+1}],\mathbf{e},t_h) - n_1^q n_1^{q+1} p(\mathbf{n},\mathbf{e},t_h)\right\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_-(q+1, N_1, e_1)\left\{n_1^{q+1}(n_1^q+1)p(\mathbf{n},\mathbf{e},t_h) - n_1^q n_1^{q+1} p(\mathbf{n},\mathbf{e},t_h)\right\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_-(q+1, N_1, e_1) n_1^{q+1} p(\mathbf{n},\mathbf{e},t_h) \\
&= \langle \mu_-(q+1, N_1, e_1) n_1^{q+1}\rangle.
\end{aligned}$$

Using $j = q-1$ and $i = 1$, with the change of variables $\bar{n}_1^{q-1} = n_1^{q-1} - 1$ and $\bar{n}_1^q = n_1^q + 1$ in the first term of Equation (5.12), then we have

$$\begin{aligned}
&\sum_{\mathbf{n}}\sum_{\mathbf{e}} n_1^q \mu_-(q, N_1, e_1)\times \\
&\qquad \left\{(n_1^q+1)p([n_1^1,\ldots,n_1^{q-1}-1,n_1^q+1,\ldots,n_1^{N_y+1}],\mathbf{e},t_h) - n_1^q p(\mathbf{n},\mathbf{e},t_h)\right\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_-(q, N_1, e_1)\times \\
&\qquad \left\{\bar{n}_1^q(\bar{n}_1^q-1)p([n_1^1,\ldots,\bar{n}_1^{q-1},\bar{n}_1^q,\ldots,n_1^{N_y+1}],\mathbf{e},t_h) - n_1^q n_1^q p(\mathbf{n},\mathbf{e},t_h)\right\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_-(q, N_1, e_1)\left\{n_1^q(n_1^q-1)p(\mathbf{n},\mathbf{e},t_h) - n_1^q n_1^q p(\mathbf{n},\mathbf{e},t_h)\right\} \\
&= -\sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_-(q, N_1, e_1) n_1^q p(\mathbf{n},\mathbf{e},t_h) \\
&= -\langle \mu_-(q, N_1, e_1) n_1^q\rangle.
\end{aligned}$$

Considering the second term of Equation (5.12), we need to consider the cases $j = q$ and $j = q+1$, for $i = 1$. For $i = 1$ and $j = q$, using the change of variables

$\bar{n}_1^{q-1} = n_1^{q-1} + 1$ and $\bar{n}_1^q = n_1^q - 1$ we have

$$\begin{aligned}
&\sum_{\mathbf{n}}\sum_{\mathbf{e}} n_1^q \mu_+(q-1, N_1, e_1)\big\{(n_1^{q-1}+1)p(D^{\mathrm{p}}_{1,q}\mathbf{n}, \mathbf{e}, t_h) - n_1^{q-1}p(\mathbf{n}, \mathbf{e}, t_h)\big\} \\
&\quad = \sum_{\mathbf{n}}\sum_{\mathbf{e}} n_1^q \mu_+(q-1, N_1, e_1)\times \\
&\qquad \big\{(n_1^{q-1}+1)p([n_1^1, \ldots, n_1^{q-1}+1, n_1^q - 1, \ldots, n_1^{N_y+1}], \mathbf{e}, t_h) - n_1^{q-1}p(\mathbf{n}, \mathbf{e}, t_h)\big\} \\
&\quad = \sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_+(q-1, N_1, e_1)\times \\
&\qquad \big\{\bar{n}_1^{q-1}(\bar{n}_1^q+1)p([n_1^1, \ldots, \bar{n}_1^{q-1}, \bar{n}_1^q, \ldots, n_1^{N_y+1}], \mathbf{e}, t_h) - n_1^q n_1^{q-1}p(\mathbf{n}, \mathbf{e}, t_h)\big\} \\
&\quad = \sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_+(q-1, N_1, e_1)\left\{n_1^{q-1}(n_1^q+1)p(\mathbf{n}, \mathbf{e}, t_h) - n_1^q n_1^{q-1}p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad = \sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_+(q-1, N_1, e_1) n_1^{q-1}p(\mathbf{n}, \mathbf{e}, t_h) \\
&\quad = \langle \mu_+(q-1, N_1, e_1) n_1^{q-1}\rangle.
\end{aligned}$$

For the case $j = q+1$ and $i = 1$, the change of variables $\bar{n}_1^{q+1} = n_1^{q+1} - 1$ and $\bar{n}_1^q = n_1^q + 1$ gives

$$\begin{aligned}
&\sum_{\mathbf{n}}\sum_{\mathbf{e}} n_1^q \mu_+(q, N_1, e_1)\big\{(n_1^q+1)p(D^{\mathrm{p}}_{1,q}\mathbf{n}, \mathbf{e}, t_h) - n_1^q p(\mathbf{n}, \mathbf{e}, t_h)\big\} \\
&\quad = \sum_{\mathbf{n}}\sum_{\mathbf{e}} n_1^q \mu_+(q, N_1, e_1)\times \\
&\qquad \big\{(n_1^q+1)p([n_1^1, \ldots, n_1^q+1, n_1^{q+1} - 1, \ldots, n_1^{N_y+1}], \mathbf{e}, t_h) - n_1^q p(\mathbf{n}, \mathbf{e}, t_h)\big\} \\
&\quad = \sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_+(q, N_1, e_1)\times \\
&\qquad \Big\{\bar{n}_1^q(\bar{n}_1^q-1)p([n_1^1, \ldots, \bar{n}_1^q, \bar{n}_1^{q+1}, \ldots, n_1^{N_y+1}], \mathbf{e}, t_h) - n_1^q n_1^q p(\mathbf{n}, \mathbf{e}, t_h)\Big\} \\
&\quad = \sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_+(q, N_1, e_1)\left\{n_1^q(n_1^q-1)p(\mathbf{n}, \mathbf{e}, t_h) - n_1^q n_1^q p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad = -\sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_+(q, N_1, e_1) n_1^q p(\mathbf{n}, \mathbf{e}, t_h) \\
&\quad = -\langle \mu_+(q, N_1, e_1) n_1^q\rangle.
\end{aligned}$$

Now, looking at the third term of Equation (5.12), which governs movement in physical space, for $j = q$ we have

$$\begin{aligned}
&\sum_{\mathbf{n}}\sum_{\mathbf{e}} n_1^q \sum_{i=1}^{N_x} \beta_-(q, N_i, e_i)\left\{(n_{i+1}^q+1)p(R^{\mathrm{m}}_{i,q}\mathbf{n}, \mathbf{e}, t_h) - n_{i+1}^q p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad = \sum_{\mathbf{n}}\sum_{\mathbf{e}} n_1^q \sum_{i=1}^{N_x} \beta_-(q, N_i, e_i)\times \\
&\qquad \left\{(n_{i+1}^q+1)p([n_1^q, \ldots, n_i^q - 1, n_{i+1}^q + 1, \ldots, n_{N_x+1}^q], \mathbf{e}, t_h) - n_{i+1}^q p(\mathbf{n}, \mathbf{e}, t_h)\right\},
\end{aligned}$$

which only produces non-zero contributions when $i = 1$, namely,

$$\begin{aligned}
&\sum_{\mathbf{n}}\sum_{\mathbf{e}} n_1^q \beta_-(q, N_1, e_1)\left\{(n_2^q+1)p([n_1^q-1, n_2^q+1, \ldots, n_{N_x+1}^q], \mathbf{e}, t_h) - n_2^q p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad = \sum_{\mathbf{n}}\sum_{\mathbf{e}} \beta_-(q, N_1, e_1)\left\{(\bar{n}_1^q+1)\bar{n}_2^q p([\bar{n}_1^q, \bar{n}_2^q, \ldots, n_{N_x+1}^q], \mathbf{e}, t_h) - n_1^q n_2^q p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad = \sum_{\mathbf{n}}\sum_{\mathbf{e}} \beta_-(q, N_1, e_1)\left\{(n_1^q+1)n_2^q p(\mathbf{n}, \mathbf{e}, t_h) - n_1^q n_2^q p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad = \sum_{\mathbf{n}}\sum_{\mathbf{e}} \beta_-(q, N_1, e_1) n_2^q p(\mathbf{n}, \mathbf{e}, t_h) \\
&\quad = \langle \beta_-(q, N_1, e_1) n_2^q \rangle,
\end{aligned}$$

where we used the change of variables $\bar{n}_1^q = n_1^q - 1$ and $\bar{n}_2^q = n_2^q + 1$, and then dropped the bars. Next, this argument can be repeated for $j = q$ and $i = 2$ in the fourth term of Equation (5.12) (chosen such that $i - 1 = 1$, and recalling that terms with $i > 2$ will give zero contributions), using the change of variables $\bar{n}_1^q = n_1^q + 1$ and $\bar{n}_2^q = n_2^q - 1$:

$$\begin{aligned}
&\sum_{\mathbf{n}}\sum_{\mathbf{e}} n_1^q \beta_+(q, N_2, e_2)\left\{(n_1^q+1)p(L_{2,q}^{\mathrm{m}}\mathbf{n}, \mathbf{e}, t_h) - n_1^q p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad = \sum_{\mathbf{n}}\sum_{\mathbf{e}} n_1^q \beta_+(q, N_2, e_2)\times \\
&\qquad\qquad \left\{(n_1^q+1)p([n_1^q+1, n_2^q-1, \ldots, n_{N_x+1}^q], \mathbf{e}, t_h) - n_1^q p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad = \sum_{\mathbf{n}}\sum_{\mathbf{e}} \beta_+(q, N_2, e_2)\left\{\bar{n}_1^q(\bar{n}_1^q-1)p([\bar{n}_1^q, \bar{n}_2^q, \ldots, n_{N_x+1}^q], \mathbf{e}, t_h) - n_1^q n_1^q p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad = \sum_{\mathbf{n}}\sum_{\mathbf{e}} \beta_+(q, N_2, e_2)\left\{n_1^q(n_1^q-1)p(\mathbf{n}, \mathbf{e}, t_h) - n_1^q n_1^q p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad = -\sum_{\mathbf{n}}\sum_{\mathbf{e}} \beta_+(q, N_2, e_2) n_1^q p(\mathbf{n}, \mathbf{e}, t_h) \\
&\quad = -\langle \beta_+(q, N_2, e_2) n_1^q \rangle.
\end{aligned}$$

Now, finally, we look at the last term on the right-hand side of Equation (5.12) which has only non-zero contributions when $j = q$ and $i = 1$:

$$\begin{aligned}
&\sum_{\mathbf{n}}\sum_{\mathbf{e}} n_1^q \left\{\gamma(q, N_1-1, e_1)(n_1^q-1)p(G_{1,q}\mathbf{n}, \mathbf{e}, t_h) - n_1^q \gamma(q, N_1, e_1)p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad = \sum_{\mathbf{n}}\sum_{\mathbf{e}} n_1^q \big\{\gamma(q, N_1-1, e_1)(n_1^q-1)p([n_1^q-1, \ldots, n_{N_x+1}^q], \mathbf{e}, t_h) \\
&\qquad\qquad - n_1^q \gamma(q, N_1, e_1)p(\mathbf{n}, \mathbf{e}, t_h)\big\}. \qquad (5.14)
\end{aligned}$$

Using the change of variable $\bar{n}_1^q = n_1^q - 1$ in the second term of Equation (5.14) we have

$$\begin{aligned}
&\sum_{\mathbf{n}}\sum_{\mathbf{e}}\left\{\bar{n}_1^q(\bar{n}_1^q+1)\gamma(q,\bar{N}_1,e_1)\bar{n}_1^q p([\bar{n}_1^q,\ldots,n_{N_x+1}^q],\mathbf{e},t_h) - (n_1^q)^2\gamma(q,N_1,e_1)p(\mathbf{n},\mathbf{e},t_h)\right\}\\
&\quad=\sum_{\mathbf{n}}\sum_{\mathbf{e}}\left\{n_1^q(n_1^q+1)\gamma(q,N_1,e_1)n_1^q p(\mathbf{n},\mathbf{e},t_h) - (n_1^q)^2\gamma(q,N_1,e_1)p(\mathbf{n},\mathbf{e},t_h)\right\}\\
&\quad=\sum_{\mathbf{n}}\sum_{\mathbf{e}}\gamma(q,N_1,e_1)n_1^q p(\mathbf{n},\mathbf{e},t_h)\\
&\quad=\langle\gamma(q,N_1,e_1)n_1^q\rangle.
\end{aligned}$$

Recompiling these simplified terms, the equation for cell evolution on the left-hand boundary in physical space becomes

$$\begin{aligned}
\frac{\partial}{\partial t}\langle n_1^q\rangle = &\frac{1}{\Delta_t}\langle\mu_-(q+1,N_1,e_1)n_1^{q+1}\rangle + \frac{1}{\Delta_t}\langle\mu_+(q-1,N_1,e_1)n_1^{q-1}\rangle\\
&-\frac{1}{\Delta_t}\langle\mu_-(q,N_1,e_1)n_1^q\rangle - \frac{1}{\Delta_t}\langle\mu_+(q,N_1,e_1)n_1^q\rangle\\
&+\frac{1}{\Delta_t}\langle\beta_-(q,N_1,e_1)n_2^q\rangle - \frac{1}{\Delta_t}\langle\beta_+(q,N_2,e_2)n_1^q\rangle + \frac{1}{\Delta_t}\langle\gamma(q,N_1,e_1)n_1^q\rangle. \quad (5.15)
\end{aligned}$$

Now we wish to find the continuum equivalent of this equation. Recalling the continuum equivalents of the dependent variables, and employing Taylor series expansions around $x = X_{\min}$, we can rewrite Equation (5.15) as (dropping the dependent variables for simplicity)

$$\begin{aligned}
\Delta_t\frac{\partial n}{\partial t} = &\left(\mu_- + \Delta_y\frac{\partial\mu_-}{\partial y} + \frac{\Delta_y^2}{2}\frac{\partial^2\mu_-}{\partial y^2} + \ldots\right)\left(n + \Delta_y\frac{\partial n}{\partial y} + \frac{\Delta_y^2}{2}\frac{\partial^2 n}{\partial y^2} + \ldots\right)\\
&+\left(\mu_+ - \Delta_y\frac{\partial\mu_+}{\partial y} + \frac{\Delta_y^2}{2}\frac{\partial^2\mu_+}{\partial y^2} + \ldots\right)\left(n - \Delta_y\frac{\partial n}{\partial y} + \frac{\Delta_y^2}{2}\frac{\partial^2 n}{\partial y^2} + \ldots\right)\\
&-\mu_+ n - \mu_- n + \gamma n + \beta_-\left(n + \Delta_x\frac{\partial n}{\partial x} + \frac{\Delta_x^2}{2}\frac{\partial^2 n}{\partial x^2} + \ldots\right)\\
&-n\left(\beta_+ + \Delta_x\frac{\partial\beta_+}{\partial x} + \frac{\Delta_x^2}{2}\frac{\partial^2\beta_+}{\partial x^2} + \ldots\right),
\end{aligned}$$

at $x = X_{\min}$, so that we have

$$\begin{aligned}
\frac{\partial n}{\partial t} = &\frac{\Delta_y}{\Delta_t}\frac{\partial}{\partial y}\left((\mu_- - \mu_+)\,n\right) + \frac{\Delta_y^2}{2\Delta_t}\frac{\partial^2}{\partial y^2}\left((\mu_- + \mu_+)\,n\right) + \frac{1}{\Delta_t}n\left(\beta_- - \beta_+\right)\\
&+\frac{\Delta_x}{\Delta_t}\left(\beta_-\frac{\partial n}{\partial x} - n\frac{\partial\beta_+}{\partial x}\right) + \frac{\Delta_x^2}{2\Delta_t}\left(\beta_-\frac{\partial^2 n}{\partial x^2} - n\frac{\partial^2\beta_+}{\partial x^2}\right) + \frac{1}{\Delta_t}\gamma n. \quad (5.16)
\end{aligned}$$

Recalling Equations (5.5)–(5.9) such that

$$\beta_\pm = \frac{D^m\Delta_t}{\Delta_x^2} \mp \frac{v^m\Delta_t}{2\Delta_x}, \quad (5.17)$$

in the limit $\Delta_x, \Delta_t \to 0$, then Equation (5.16) can be rewritten (ignoring higher order corrections) as

$$\begin{aligned}\frac{\partial n}{\partial t} = \frac{\partial}{\partial y}\left(v^p n\right) + \frac{\partial^2}{\partial y^2}\left(D^p n\right) + \frac{1}{\Delta_x} v^m n + \frac{1}{2} v^m \frac{\partial n}{\partial x} + \frac{1}{\Delta_x} D^m \frac{\partial n}{\partial x} \\ - \frac{1}{\Delta_x} n \frac{\partial}{\partial x} D^m - \frac{1}{2} n \frac{\partial}{\partial x} v^m + \frac{1}{2} D^m \frac{\partial^2 n}{\partial x^2} - \frac{1}{2} n \frac{\partial^2}{\partial x^2} D^m + rn.\end{aligned}$$

In order to prevent blow-up of terms in the limit $\Delta_x \to 0$, we require

$$v^m n + D^m \frac{\partial n}{\partial x} - n \frac{\partial}{\partial x} D^m = 0 \qquad \text{at} \quad x = X_{\min}.$$

As such, we have no flux of cells out of the physical space boundary at $x = X_{\min}$.

Boundary condition at $x = X_{\max}$. Revisiting Equation (5.1) to find the equation for the right-hand lattice site in physical space we multiply by $n^q_{N_x+1}$ and sum over all possible states $\mathbf{n}$ and $\mathbf{e}$ to give

$$\begin{aligned}&\Delta_t \sum_{\mathbf{n}} \sum_{\mathbf{e}} n^q_{N_x+1} \frac{\partial}{\partial t} p(\mathbf{n}, \mathbf{e}, t_h) = \\ &\sum_{\mathbf{n}} \sum_{\mathbf{e}} n^q_{N_x+1} \sum_{i=1}^{N_x+1} \sum_{j=1}^{N_y} \mu_-(j+1, N_i, e_i) \left\{ (n_i^{j+1}+1) p(U^{\mathrm{p}}_{i,j}\mathbf{n}, \mathbf{e}, t_h) - n_i^{j+1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\ &+ \sum_{\mathbf{n}} \sum_{\mathbf{e}} n^q_{N_x+1} \sum_{i=1}^{N_x+1} \sum_{j=2}^{N_y+1} \mu_+(j-1, N_i, e_i) \big\{ (n_i^{j-1}+1) p(D^{\mathrm{p}}_{i,j}\mathbf{n}, \mathbf{e}, t_h) \\ &\qquad\qquad - n_i^{j-1} p(\mathbf{n}, \mathbf{e}, t_h) \big\} \\ &+ \sum_{\mathbf{n}} \sum_{\mathbf{e}} n^q_{N_x+1} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y+1} \beta_-(j, N_i, e_i) \left\{ (n_{i+1}^{j}+1) p(R^{\mathrm{m}}_{i,j}\mathbf{n}, \mathbf{e}, t_h) - n_{i+1}^{j} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\ &+ \sum_{\mathbf{n}} \sum_{\mathbf{e}} n^q_{N_x+1} \sum_{i=2}^{N_x+1} \sum_{j=1}^{N_y+1} \beta_+(j, N_i, e_i) \left\{ (n_{i-1}^{j}+1) p(L^{\mathrm{m}}_{i,j}\mathbf{n}, \mathbf{e}, t_h) - n_{i-1}^{j} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\ &+ \sum_{\mathbf{n}} \sum_{\mathbf{e}} n^q_{N_x+1} \sum_{i=1}^{N_x+1} \sum_{j=1}^{N_y+1} \big\{ \gamma(j, N_i - 1, e_i)(n_i^j - 1) p(G_{i,j}\mathbf{n}, \mathbf{e}, t_h) \\ &\qquad\qquad - \gamma(j, N_i, e_i) n_i^j p(\mathbf{n}, \mathbf{e}, t_h) \big\}.\end{aligned} \tag{5.18}$$

Now we can repeat the analysis from the previous section, where we considered the action on the boundary $x = X_{\min}$, for $x = X_{\max}$. In the first two terms on the right-hand side of Equation (5.18), we are interested in the $i = N_x$ terms only. In the first

term, we need to consider the cases $j = q$ and $j = q - 1$. First, looking at $j = q$ and using the change of variables $\bar{n}^{q+1}_{N_x+1} = n^{q+1}_{N_x+1} + 1$ and $\bar{n}^{q}_{N_x+1} = n^{q}_{N_x+1} - 1$ gives

$$\begin{aligned}
&\sum_{\mathbf{n}}\sum_{\mathbf{e}} n^q_{N_x+1}\mu_-(q+1, N_{N_x+1}, e_{N_x+1})\big\{(n^{q+1}_{N_x+1}+1)p(U^{\mathrm{p}}_{N_x+1,q}\mathbf{n}, \mathbf{e}, t_h) \\
&\qquad\qquad - n^{q+1}_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h)\big\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} n^q_{N_x+1}\mu_-(q+1, N_{N_x+1}, e_{N_x+1})\times \\
&\qquad \big\{(n^{q+1}_{N_x+1}+1)p([n^1_{N_x+1}, \ldots, n^q_{N_x+1}-1, n^{q+1}_{N_x+1}+1, \ldots, n^{N_y+1}_{N_x+1}], \mathbf{e}, t_h) \\
&\qquad\qquad - n^{q+1}_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h)\big\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_-(q+1, N_{N_x+1}, e_{N_x+1})\times \\
&\qquad \big\{\bar{n}^{q+1}_{N_x+1}(\bar{n}^q_{N_x+1}+1)p([n^1_{N_x+1}, \ldots, \bar{n}^q_{N_x+1}, \bar{n}^{q+1}_{N_x+1}, \ldots, n^{N_y+1}_{N_x+1}], \mathbf{e}, t_h) \\
&\qquad\qquad - n^q_{N_x+1}n^{q+1}_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h)\big\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_-(q+1, N_{N_x+1}, e_{N_x+1})\big\{n^{q+1}_{N_x+1}(n^q_{N_x+1}+1)p(\mathbf{n}, \mathbf{e}, t_h) \\
&\qquad\qquad - n^q_{N_x+1}n^{q+1}_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h)\big\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_-(q+1, N_{N_x+1}, e_{N_x+1})n^{q+1}_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h) \\
&= \langle \mu_-(q+1, N_{N_x+1}, e_{N_x+1})n^{q+1}_{N_x+1}\rangle.
\end{aligned}$$

Using $j = q$ and $i = N_x + 1$, with the change of variables $\bar{n}^{q-1}_{N_x+1} = n^{q-1}_{N_x+1} - 1$ and $\bar{n}^{q}_{N_x+1} = n^{q}_{N_x+1} + 1$ in the first term, then we have

$$\begin{aligned}
&\sum_{\mathbf{n}}\sum_{\mathbf{e}} n^q_{N_x+1}\mu_-(q, N_{N_x+1}, e_{N_x+1})\times \\
&\qquad \big\{(n^q_{N_x+1}+1)p([n^1_{N_x+1}, \ldots, n^{q-1}_{N_x+1}-1, n^q_{N_x+1}+1, \ldots, n^{N_y+1}_{N_x+1}], \mathbf{e}, t_h) \\
&\qquad\qquad - n^q_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h)\big\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_-(q, N_{N_x+1}, e_{N_x+1})\times \\
&\qquad \big\{\bar{n}^q_{N_x+1}(\bar{n}^q_{N_x+1}-1)p([n^1_{N_x+1}, \ldots, \bar{n}^{q-1}_{N_x+1}, \bar{n}^q_{N_x+1}, \ldots, n^{N_y+1}_{N_x+1}], \mathbf{e}, t_h) \\
&\qquad\qquad - n^q_{N_x+1}n^q_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h)\big\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_-(q, N_{N_x+1}, e_{N_x+1})\big\{n^q_{N_x+1}(n^q_{N_x+1}-1)p(\mathbf{n}, \mathbf{e}, t_h) \\
&\qquad\qquad - n^q_{N_x+1}n^q_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h)\big\} \\
&= -\sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_-(q, N_{N_x+1}, e_{N_x+1})n^q_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h) \\
&= -\langle \mu_-(q, N_{N_x+1}, e_{N_x+1})n^q_{N_x+1}\rangle.
\end{aligned}$$

For the second term, we need to consider $j = q$ and $j = q + 1$ whilst maintaining $i = N_x + 1$. For $i = N_x + 1$ and $j = q$, whilst using the change of variables $\bar{n}^{q-1}_{N_x+1} = n^{q-1}_{N_x+1} + 1$ and $\bar{n}^{q}_{N_x+1} = n^{q}_{N_x+1} - 1$ we have

$$\begin{aligned}
&\sum_{\mathbf{n}}\sum_{\mathbf{e}} n^q_{N_x+1}\mu_+(q-1, N_{N_x+1}, e_{N_x+1})\big\{(n^{q-1}_{N_x+1}+1)p(D^{\mathrm{p}}_{N_x+1,q}\mathbf{n}, \mathbf{e}, t_h) \\
&\qquad\qquad - n^{q-1}_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h)\big\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} n^q_{N_x+1}\mu_+(q-1, N_{N_x+1}, e_{N_x+1})\times \\
&\qquad \big\{(n^{q-1}_{N_x+1}+1)p([n^1_{N_x+1}, \ldots, n^{q-1}_{N_x+1}+1, n^q_{N_x+1}-1, \ldots, n^{N_y+1}_{N_x+1}], \mathbf{e}, t_h) \\
&\qquad\qquad - n^{q-1}_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h)\big\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_+(q-1, N_{N_x+1}, e_{N_x+1})\times \\
&\qquad \big\{\bar{n}^{q-1}_{N_x+1}(\bar{n}^q_{N_x+1}+1)p([n^1_{N_x+1}, \ldots, \bar{n}^{q-1}_{N_x+1}, \bar{n}^q_{N_x+1}, \ldots, n^{N_y+1}_{N_x+1}], \mathbf{e}, t_h) \\
&\qquad\qquad - n^q_{N_x+1}n^{q-1}_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h)\big\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_+(q-1, N_{N_x+1}, e_{N_x+1})\big\{n^{q-1}_{N_x+1}(n^q_{N_x+1}+1)p(\mathbf{n}, \mathbf{e}, t_h) \\
&\qquad\qquad - n^q_{N_x+1}n^{q-1}_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h)\big\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_+(q-1, N_{N_x+1}, e_{N_x+1})n^{q-1}_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h) \\
&= \langle \mu_+(q-1, N_{N_x+1}, e_{N_x+1})n^{q-1}_{N_x+1}\rangle.
\end{aligned}$$

Now looking at $j = q+1$ and $i = N_x$, the change of variables $\bar{n}^{q+1}_{N_x+1} = n^{q+1}_{N_x+1} - 1$ and $\bar{n}^{q}_{N_x+1} = n^{q}_{N_x+1} + 1$ gives

$$
\begin{aligned}
&\sum_{\mathbf{n}}\sum_{\mathbf{e}} n^q_{N_x+1}\mu_+(q, N_{N_x+1}, e_{N_x+1})\left\{(n^q_{N_x+1}+1)p(D^{\mathrm{p}}_{N_x+1,q}\mathbf{n}, \mathbf{e}, t_h) - n^q_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad = \sum_{\mathbf{n}}\sum_{\mathbf{e}} n^q_{N_x+1}\mu_+(q, N_{N_x+1}, e_{N_x+1})\times \\
&\qquad \left\{(n^q_{N_x+1}+1)p([n^1_{N_x+1}, \ldots, n^q_{N_x+1}+1, n^{q+1}_{N_x+1}-1, \ldots, n^{N_y+1}_{N_x+1}], \mathbf{e}, t_h)\right. \\
&\qquad\qquad \left. - n^q_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad = \sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_+(q, N_{N_x+1}, e_{N_x+1})\times \\
&\qquad \left\{\bar{n}^q_{N_x+1}(\bar{n}^q_{N_x+1}-1)p([n^1_{N_x+1}, \ldots, \bar{n}^q_{N_x+1}, \bar{n}^{q+1}_{N_x+1}, \ldots, n^{N_y+1}_{N_x+1}], \mathbf{e}, t_h)\right. \\
&\qquad\qquad \left. - n^q_{N_x+1}n^q_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad = \sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_+(q, N_{N_x+1}, e_{N_x+1})\left\{n^q_{N_x+1}(n^q_{N_x+1}-1)p(\mathbf{n}, \mathbf{e}, t_h)\right. \\
&\qquad\qquad \left. - n^q_{N_x+1}n^q_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad = -\sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_+(q, N_{N_x+1}, e_{N_x+1})n^q_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h) \\
&\quad = -\langle \mu_+(q, N_{N_x+1}, e_{N_x+1})n^q_{N_x+1}\rangle.
\end{aligned}
$$

Next, we want to consider the movement terms in physical space. In these cases, we will only have non-zero contributions when $j = q$. Looking at the first term, we will have contributions only when $i = N_x$. Then, using the change of variables

$\bar{n}^q_{N_x+1} = n^q_{N_x+1} + 1$ and $\bar{n}^q_{N_x} = n^q_{N_x} - 1$ and dropping the bars, we arrive at

$$
\begin{aligned}
&\sum_{\mathbf{n}}\sum_{\mathbf{e}} n^q_{N_x+1}\beta_-(q, N_{N_x}, e_{N_x})\left\{(n^q_{N_x+1}+1)p(R^{\mathrm{m}}_{N_x,q}\mathbf{n}, \mathbf{e}, t_h) - n^q_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad= \sum_{\mathbf{n}}\sum_{\mathbf{e}} n^q_{N_x+1}\beta_-(q, N_{N_x}, e_{N_x})\times \\
&\qquad\left\{(n^q_{N_x+1}+1)p([n^q_1, \ldots, n^q_{N_x}-1, n^q_{N_x+1}+1], \mathbf{e}, t_h) - n^q_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad= \sum_{\mathbf{n}}\sum_{\mathbf{e}} n^q_{N_x}\beta_-(q, N_{N_x}, e_{N_x})\times \\
&\qquad\left\{(n^q_{N_x+1}+1)p([n^q_1, \ldots, n^q_{N_x}-1, n^q_{N_x+1}+1], \mathbf{e}, t_h) - n^q_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h)\right\}, \\
&\quad= \sum_{\mathbf{n}}\sum_{\mathbf{e}} \beta_-(q, N_{N_x}, e_{N_x})\times \\
&\qquad\left\{(\bar{n}^q_{N_x}-1)\bar{n}^q_{N_x+1}p([n^q_1, \ldots, \bar{n}^q_{N_x}, \bar{n}^q_{N_x+1}], \mathbf{e}, t_h) - n^q_{N_x}n^q_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad= \sum_{\mathbf{n}}\sum_{\mathbf{e}} \beta_-(q, N_{N_x}, e_{N_x})\times \\
&\qquad\left\{(n^q_{N_x}-1)n^q_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h) - n^q_{N_x}n^q_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad= -\sum_{\mathbf{n}}\sum_{\mathbf{e}} \beta_-(q, N_{N_x}, e_{N_x})n^q_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h) \\
&\quad= -\langle\beta_-(q, N_{N_x}, e_{N_x})n^q_{N_x+1}\rangle.
\end{aligned}
$$

We can repeat this for $j = q$ and $i = N_x + 1$ in the fourth term of Equation (5.18), using the change of variables $\bar{n}^q_{N_x+1} = n^q_{N_x+1} - 1$ and $\bar{n}^q_{N_x} = n^q_{N_x} + 1$ to give

$$
\begin{aligned}
&\sum_{\mathbf{n}}\sum_{\mathbf{e}} n^q_{N_x+1}\beta_+(q, N_{N_x+1}, e_{N_x+1})\left\{(n^q_{N_x}+1)p(L^{\mathrm{m}}_{N_x,q}\mathbf{n}, \mathbf{e}, t_h) - n^q_{N_x}p(\mathbf{n}, \mathbf{e}, t_h)\right\}, \\
&\quad= \sum_{\mathbf{n}}\sum_{\mathbf{e}} n^q_{N_x+1}\beta_+(q, N_{N_x+1}, e_{N_x+1})\times \\
&\qquad\left\{(n^q_{N_x}+1)p([n^q_1, \ldots, n^q_{N_x}+1, n^q_{N_x+1}-1], \mathbf{e}, t_h) - n^q_{N_x}p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad= \sum_{\mathbf{n}}\sum_{\mathbf{e}} \beta_+(q, N_{N_x+1}, e_{N_x+1})\times \\
&\qquad\left\{\bar{n}^q_{N_x}(\bar{n}^q_{N_x+1}+1)p([n^q_1, \ldots, \bar{n}^q_{N_x}, \bar{n}^q_{N_x+1}], \mathbf{e}, t_h) - n^q_{N_x+1}n^q_{N_x}p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad= \sum_{\mathbf{n}}\sum_{\mathbf{e}} \beta_+(q, N_{N_x+1}, e_{N_x+1})\left\{n^q_{N_x}(n^q_{N_x+1}+1)p(\mathbf{n}, \mathbf{e}, t_h) - n^q_{N_x}n^q_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad= \sum_{\mathbf{n}}\sum_{\mathbf{e}} \beta_+(q, N_{N_x+1}, e_{N_x+1})n^q_{N_x}p(\mathbf{n}, \mathbf{e}, t_h) \\
&\quad= \langle\beta_+(q, N_{N_x+1}, e_{N_x+1})n^q_{N_x}\rangle.
\end{aligned}
$$

The final term on right-hand side of Equation (5.18) has non-zero contributions when $j = q$ and $i = N_x + 1$ only. We employ the change of variable $\bar{n}^q_{N_x+1} = n^q_{N_x+1} - 1$ in

the second term to give

$$\begin{aligned}
&\sum_{\mathbf{n}}\sum_{\mathbf{e}} n^q_{N_x+1}\{\gamma(q, N_{N_x+1}-1, e_{N_x+1})(n^q_{N_x+1}-1)p(G_{N_x+1,q}\mathbf{n}, \mathbf{e}, t_h) \\
&\qquad\qquad - n^q_{N_x+1}\gamma(q, N_{N_x+1}, e_{N_x+1})p(\mathbf{n}, \mathbf{e}, t_h)\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} n^q_{N_x+1}\big\{\gamma(q, N_{N_x+1}-1, e_{N_x+1})(n^q_{N_x+1}-1)p([n^q_1, \ldots, n^q_{N_x+1}-1], \mathbf{e}, t_h) \\
&\qquad\qquad - n^q_{N_x+1}\gamma(q, N_{N_x+1}, e_{N_x+1})p(\mathbf{n}, \mathbf{e}, t_h)\big\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} \big\{\gamma(q, \bar{N_{N_x+1}}, e_{N_x+1})(\bar{n}^q_{N_x+1}-1)\bar{n}^q_{N_x+1}p([n^q_1, \ldots, \bar{n}^q_{N_x+1}], \mathbf{e}, t_h) \\
&\qquad\qquad - \gamma(q, N_{N_x+1}, e_{N_x+1})(n^q_{N_x+1})^2 p(\mathbf{n}, \mathbf{e}, t_h)\big\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} \gamma(q, N_{N_x+1}, e_{N_x+1})\left\{(n^q_{N_x+1}-1)n^q_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h) - (n^q_{N_x+1})^2 p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&= \sum_{\mathbf{n}}\sum_{\mathbf{e}} \gamma(q, N_{N_x+1}, e_{N_x+1})n^q_{N_x+1}p(\mathbf{n}, \mathbf{e}, t_h) \\
&= \langle\gamma(q, N_{N_x+1}, e_{N_x+1})n^q_{N_x+1}\rangle.
\end{aligned}$$

Putting these together, the equation for cell evolution on the right-hand boundary in physical space becomes

$$\begin{aligned}
\frac{\partial}{\partial t}\langle n^q_{N_x+1}\rangle &= \frac{1}{\Delta_t}\langle\mu_-(q+1, N_{N_x+1}, e_{N_x+1})n^{q+1}_{N_x+1}\rangle + \frac{1}{\Delta_t}\langle\mu_+(q-1, N_{N_x+1}, e_{N_x+1})n^{q-1}_{N_x+1}\rangle \\
&\quad - \frac{1}{\Delta_t}\langle\mu_-(q, N_{N_x+1}, e_{N_x+1})n^q_{N_x+1}\rangle - \frac{1}{\Delta_t}\langle\mu_+(q, N_{N_x+1}, e_{N_x+1})n^q_{N_x+1}\rangle \\
&\quad + \frac{1}{\Delta_t}\langle\beta_+(q, N_{N_x+1}, e_{N_x+1})n^q_{N_x}\rangle - \frac{1}{\Delta_t}\langle\beta_-(q, N_{N_x}, e_{N_x})n^q_{N_x+1}\rangle \\
&\quad + \frac{1}{\Delta_t}\langle\gamma(q, N_{N_x+1}, e_{N_x+1})n^q_{N_x+1}\rangle. \qquad (5.19)
\end{aligned}$$

Now we want to find the continuum limit of Equation (5.19). Recalling the continuum equivalents of the dependent variables whilst employing Taylor series expansions around $x = X_{\max}$, we can rewrite Equation (5.19) as (dropping the dependent variables for simplicity)

$$\begin{aligned}
\Delta_t\frac{\partial n}{\partial t} &= \left(\mu_- + \Delta_y\frac{\partial\mu_-}{\partial y} + \frac{\Delta_y^2}{2}\frac{\partial^2\mu_-}{\partial y^2} + \ldots\right)\left(n + \Delta_y\frac{\partial n}{\partial y} + \frac{\Delta_y^2}{2}\frac{\partial^2 n}{\partial y^2} + \ldots\right) \\
&\quad + \left(\mu_+ - \Delta_y\frac{\partial\mu_+}{\partial y} + \frac{\Delta_y^2}{2}\frac{\partial^2\mu_+}{\partial y^2} + \ldots\right)\left(n - \Delta_y\frac{\partial n}{\partial y} + \frac{\Delta_y^2}{2}\frac{\partial^2 n}{\partial y^2} + \ldots\right) \\
&\quad - \mu_+ n - \mu_- n + \gamma n + \beta_+\left(n - \Delta_x\frac{\partial n}{\partial x} + \frac{\Delta_x^2}{2}\frac{\partial^2 n}{\partial x^2} + \ldots\right) \\
&\quad - n\left(\beta_- - \Delta_x\frac{\partial\beta_-}{\partial x} + \frac{\Delta_x^2}{2}\frac{\partial^2\beta_-}{\partial x^2} + \ldots\right),
\end{aligned}$$

at $x = X_{\max}$ so that

$$\frac{\partial n}{\partial t} = \frac{\Delta_y}{\Delta_t}\frac{\partial}{\partial y}\left(\left(\mu_- - \mu_+\right)n\right) + \frac{\Delta_y^2}{2\Delta_t}\frac{\partial^2}{\partial y^2}\left(\left(\mu_- + \mu_+\right)n\right) - \frac{1}{\Delta_t}n\left(\beta_- - \beta_+\right) - \frac{\Delta_x}{\Delta_t}\left(\beta_+\frac{\partial n}{\partial x} - n\frac{\partial \beta_-}{\partial x}\right) + \frac{\Delta_x^2}{2\Delta_t}\left(\beta_+\frac{\partial^2 n}{\partial x^2} - n\frac{\partial^2 \beta_-}{\partial x^2}\right) + \frac{1}{\Delta_t}\gamma n. \tag{5.20}$$

Recalling Equations (5.5)–(5.9) and (5.17), then Equation (5.20) can be rewritten (ignoring higher order corrections) as

$$\frac{\partial n}{\partial t} = \frac{\partial}{\partial y}\left(v^p n\right) + \frac{\partial^2}{\partial y^2}\left(D^p n\right) - \frac{1}{\Delta_x}v^m n + \frac{1}{2}v^m\frac{\partial n}{\partial x} - \frac{1}{\Delta_x}D^m\frac{\partial n}{\partial x} + \frac{1}{\Delta_x}n\frac{\partial}{\partial x}D^m - \frac{1}{2}n\frac{\partial}{\partial x}v^m + \frac{1}{2}D^m\frac{\partial^2 n}{\partial x^2} - \frac{1}{2}n\frac{\partial^2}{\partial x^2}D^m + rn.$$

In order to prevent blow-up of terms in the limit $\Delta_x \to 0$, we require

$$-v^m n - D^m\frac{\partial n}{\partial x} + n\frac{\partial}{\partial x}D^m = 0 \qquad \text{at} \quad x = X_{\max}.$$

As such, we also have a no flux condition for the cells out of the physical space boundary at $x = X_{\max}$.

Boundary condition at $y = Y_{\min}$. Returning to Equation (5.1), we seek an equation for the evolution of the cell number on the left-most lattice site in phenotype space, *i.e.*, at y_1. To find this, we multiply Equation (5.1) by n_s^1 and sum over all

possible states $\mathbf{n}$ and $\mathbf{e}$ to give

$$\begin{aligned}
&\Delta_t \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^1 \frac{\partial}{\partial t} p(\mathbf{n}, \mathbf{e}, t_h) + \mathcal{O}(\Delta_t^2) \\
&\quad = \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^1 \sum_{i=1}^{N_x+1} \sum_{j=1}^{N_y} \mu_-(j+1, N_i, e_i) \times \\
&\qquad\qquad \left\{ (n_i^{j+1} + 1) p(U_{i,j}^{\mathrm{p}} \mathbf{n}, \mathbf{e}, t_h) - n_i^{j+1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&\qquad + \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^1 \sum_{i=1}^{N_x+1} \sum_{j=2}^{N_y+1} \mu_+(j-1, N_i, e_i) \times \\
&\qquad\qquad \left\{ (n_i^{j-1} + 1) p(D_{i,j}^{\mathrm{p}} \mathbf{n}, \mathbf{e}, t_h) - n_i^{j-1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&\qquad + \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^1 \sum_{i=1}^{N_x} \sum_{j=1}^{N_y+1} \beta_-(j, N_i, e_i) \left\{ (n_{i+1}^j + 1) p(R_{i,j}^{\mathrm{m}} \mathbf{n}, \mathbf{e}, t_h) - n_{i+1}^j p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&\qquad + \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^1 \sum_{i=2}^{N_x+1} \sum_{j=1}^{N_y+1} \beta_+(j, N_i, e_i) \left\{ (n_{i-1}^j + 1) p(L_{i,j}^{\mathrm{m}} \mathbf{n}, \mathbf{e}, t_h) - n_{i-1}^j p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&\qquad + \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^1 \sum_{i=1}^{N_x+1} \sum_{j=1}^{N_y+1} \Big\{ \gamma(j, N_i - 1, e_i)(n_i^j - 1) p(G_{i,j} \mathbf{n}, \mathbf{e}, t_h) \\
&\qquad\qquad - \gamma(j, N_i, e_i) n_i^j p(\mathbf{n}, \mathbf{e}, t_h) \Big\}.
\end{aligned} \tag{5.21}$$

Using the same methods as on the boundaries in physical space, we can change variables in each term to find an equation for the evolution of cell number. Consider the first term on the right-hand side. The non-zero contributions come from when $j = 1$ and $i = s$. In the second term, the non-zero terms are for $j = 2$ and $i = s$. The third term gives non-zero contributions when $j = 1$ and $i = s$ or $i = s - 1$. In the fourth term, there are non-zero contributions when $j = 1$ and $i = s$ or $i = s + 1$. The final term produces non-zero contributions only when $i = s$ and $j = 1$. Employing

this knowledge, Equation (5.21) becomes

$$
\begin{aligned}
&\Delta_t \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^1 \frac{\partial}{\partial t} p(\mathbf{n}, \mathbf{e}, t_h) + \mathcal{O}(\Delta_t^2) \\
&\quad = \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^1 \mu_-(2, N_s, e_s) \left\{ (n_s^2 + 1) p(U_{s,1}^{\mathrm{p}} \mathbf{n}, \mathbf{e}, t_h) - n_s^2 p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&\qquad + \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^1 \mu_+(1, N_s, e_s) \left\{ (n_s^1 + 1) p(D_{s,2}^{\mathrm{p}} \mathbf{n}, \mathbf{e}, t_h) - n_s^1 p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&\qquad + \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^1 \sum_{i=s,s-1} \beta_-(1, N_i, e_i) \left\{ (n_{i+1}^1 + 1) p(R_{i,1}^{\mathrm{m}} \mathbf{n}, \mathbf{e}, t_h) - n_{i+1}^1 p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&\qquad + \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^1 \sum_{i=s,s+1} \beta_+(1, N_i, e_i) \left\{ (n_{i-1}^1 + 1) p(L_{i,1}^{\mathrm{m}} \mathbf{n}, \mathbf{e}, t_h) - n_{i-1}^1 p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&\qquad + \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^1 \gamma(1, N_s, e_s) \left\{ n_s^1 p(\mathbf{n}, \mathbf{e}, t_h) - (n_s^1 + 1) p(G_{s,1} \mathbf{n}, \mathbf{e}, t_h) \right\}. \qquad (5.22)
\end{aligned}
$$

The first term can be rewritten using the change of variables $\bar{n}_s^2 = n_s^2 + 1$ and $\bar{n}_s^1 = n_s^1 - 1$ to give

$$
\begin{aligned}
&\sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^1 \mu_-(2, N_s, e_s) \left\{ (n_s^2 + 1) p(U_{s,1}^{\mathrm{p}} \mathbf{n}, \mathbf{e}, t_h) - n_s^{N_y+1} p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&\quad = \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^1 \mu_-(2, N_s, e_s) \times \\
&\qquad\qquad \left\{ (n_s^2 + 1) p([n_s^1 - 1, n_s^2 + 1, \ldots n_s^{N_y+1}], \mathbf{e}, t_h) - n_s^2 p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&\quad = \sum_{\mathbf{n}} \sum_{\mathbf{e}} \mu_-(2, N_s, e_s) \times \\
&\qquad\qquad \left\{ \bar{n}_s^2 (\bar{n}_s^1 + 1) p([\bar{n}_s^1, \bar{n}_s^2, \ldots, \bar{n}_s^{N_y+1}], \mathbf{e}, t_h) - n_s^2 n_s^1 p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&\quad = \sum_{\mathbf{n}} \sum_{\mathbf{e}} \mu_-(2, N_s, e_s) \left\{ n_s^2 (n_s^1 + 1) p(\mathbf{n}, \mathbf{e}, t_h) - n_s^1 n_s^2 p(\mathbf{n}, \mathbf{e}, t_h) \right\} \\
&\quad = \sum_{\mathbf{n}} \sum_{\mathbf{e}} \mu_-(2, N_s, e_s) n_s^2 p(\mathbf{n}, \mathbf{e}, t_h) \\
&\quad = \langle \mu_-(2, N_s, e_s) n_s^2 \rangle.
\end{aligned}
$$

The second term in the right-hand side of Equation (5.22) can then be rewritten as

the following (using $\bar{n}_s^2 = n_s^2 - 1$ and $\bar{n}_s^1 = n_s^1 + 1$):

$$\begin{aligned}
&\sum_{\mathbf{n}}\sum_{\mathbf{e}} n_s^1 \mu_+(1, N_s, e_s)\left\{(n_s^1+1)p(D_{s,1}^{\mathrm{p}}\mathbf{n}, \mathbf{e}, t_h) - n_s^1 p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad = \sum_{\mathbf{n}}\sum_{\mathbf{e}} n_s^1 \mu_+(1, N_s, e_s)\times \\
&\qquad\qquad \left\{(n_s^1+1)p([n_s^1+1, n_s^2-1, \ldots, n_s^{N_y+1}], \mathbf{e}, t_h) - n_s^1 p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad = \sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_+(1, N_s, e_s)\times \\
&\qquad\qquad \left\{\bar{n}_s^1(\bar{n}_s^1-1)p([\bar{n}_s^1, \bar{n}_s^2, \ldots, n_s^{N_y}], \mathbf{e}, t_h) - n_s^1 n_s^1 p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad = \sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_+(1, N_s, e_s)\left\{n_s^1(n_s^2+1)p(\mathbf{n}, \mathbf{e}, t_h) - n_s^1 n_s^1 p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\
&\quad = -\sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_+(1, N_s, e_s) n_s^1 p(\mathbf{n}, \mathbf{e}, t_h) \\
&\quad = -\langle \mu_+(1, N_s, e_s) n_s^1 \rangle.
\end{aligned}$$

The contributions from the terms describing movement in physical space are the same as in Equation (5.3) with $j = 1$ and the growth terms are also the same as those in Equation (5.3) with $i = s$ and $j = 1$. Putting these together gives

$$\begin{aligned}
\frac{\partial}{\partial t}\langle n_s^1 \rangle = &\frac{1}{\Delta_t}\langle \mu_-(2, N_s, e_s) n_s^2 \rangle - \frac{1}{\Delta_t}\langle \mu_+(1, N_s, e_s) n_s^1 \rangle \\
&+ \frac{1}{\Delta_t}\langle \beta_+(1, N_s, e_s) n_{s-1}^1 \rangle + \frac{1}{\Delta_t}\langle \beta_-(1, N_s, e_s) n_{s+1}^1 \rangle \\
&- \frac{1}{\Delta_t}\langle \beta_-(1, N_{s-1}, e_{s-1}) n_s^1 \rangle - \frac{1}{\Delta_t}\langle \beta_+(1, N_{s+1}, e_{s+1}) n_s^1 \rangle \\
&+ \frac{1}{\Delta_t}\langle \gamma(1, N_s, e_s) n_s^1 \rangle. \qquad (5.23)
\end{aligned}$$

Then, in the limit $\Delta_x, \Delta_y, \Delta_t \to 0$ Equation (5.23) can be rewritten at $y = Y_{\min}$

as follows:

$$\begin{aligned}\frac{\partial}{\partial t}n(x,y,t) = &\frac{1}{\Delta_t}\beta_+(y,\rho(x,t),e(x,t))n(x-\Delta_x,y,t)\\ &+\frac{1}{\Delta_t}\beta_-(y,\rho(x,t),e(x,t))n(x+\Delta_x,y,t)\\ &-\frac{1}{\Delta_t}\beta_-(y,\rho(x-\Delta_x,t),e(x-\Delta_x,t))n(x,y,t)\\ &-\frac{1}{\Delta_t}\beta_+(y,\rho(x+\Delta_x,t),e(x+\Delta_x,t))n(x,y,t)\\ &+\frac{1}{\Delta_t}\mu_-(y+\Delta_y,\rho(x,t),e(x,t))n(x,y+\Delta_y,t)\\ &-\frac{1}{\Delta_t}\mu_+(y,\rho(x,t),e(x,t))n(x,y,t)\\ &+\frac{1}{\Delta_t}\gamma(y,\rho(x,t),e(x,t))n(x,y,t).\end{aligned}$$

Employing the aforementioned Taylor expansions, we find

$$\begin{aligned}
\frac{\partial}{\partial t}n(x,y,t) = &\frac{1}{\Delta_t}\beta_+(y,\rho(x,t),e(x,t))\times \\
&\qquad\left[n(x,y,t) - \Delta_x\frac{\partial}{\partial x}n(x,y,t) + \frac{\Delta_x^2}{2}\frac{\partial^2}{\partial x^2}n(x,y,t)\right] \\
&+\frac{1}{\Delta_t}\beta_-(y,\rho(x,t),e(x,t))\times \\
&\qquad\left[n(x,y,t) + \Delta_x\frac{\partial}{\partial x}n(x,y,t) + \frac{\Delta_x^2}{2}\frac{\partial^2}{\partial x^2}n(x,y,t)\right] \\
&-\frac{1}{\Delta_t}n(x,y,t)\times \\
&\qquad\left[\beta_-(y,\rho(x,t),e(x,t)) - \Delta_x\frac{\partial}{\partial x}\beta_-(y,\rho(x,t),e(x,t)\right. \\
&\qquad\qquad\left.+\frac{\Delta_x^2}{2}\frac{\partial^2}{\partial x^2}\beta_-(y,\rho(x,t),e(x,t)\right] \\
&-\frac{1}{\Delta_t}n(x,y,t)\times \\
&\qquad\left[\beta_+(y,\rho(x,t),e(x,t)) + \Delta_x\frac{\partial}{\partial x}\beta_+(y,\rho(x,t),e(x,t)\right. \\
&\qquad\qquad\left.+\frac{\Delta_x^2}{2}\frac{\partial^2}{\partial x^2}\beta_+(y,\rho(x,t),e(x,t)\right] \\
&+\frac{1}{\Delta_t}\left[n(x,y,t) + \Delta_y\frac{\partial}{\partial y}n(x,y,t) + \frac{\Delta_y^2}{2}\frac{\partial^2}{\partial y^2}n(x,y,t)\right]\times \\
&\qquad\left[\mu_-(y,\rho(x,t),e(x,t)) + \Delta_y\frac{\partial}{\partial y}\mu_-(y,\rho(x,t),e(x,t)\right. \\
&\qquad\qquad\left.+\frac{\Delta_y^2}{2}\frac{\partial^2}{\partial y^2}\mu_-(y,\rho(x,t),e(x,t)\right] \\
&-\frac{1}{\Delta_t}\mu_+(y,\rho(x,t),e(x,t))n(x,y,t) \\
&+\frac{1}{\Delta_t}\gamma(y,\rho(x,t),e(x,t))n(x,y,t).
\end{aligned}$$

This can be rewritten as (dropping the dependent variables for simplicity)

$$\begin{aligned}
\frac{\partial n}{\partial t} = &\frac{1}{\Delta_t}\left(\mu_- - \mu_+\right)n + \frac{\Delta_y}{\Delta_t}\frac{\partial}{\partial y}\left(\mu_- n\right) + \frac{\Delta_y^2}{2\Delta_t}\frac{\partial^2}{\partial y^2}\left(\mu_- n\right) + \frac{\Delta_x}{\Delta_t}\frac{\partial}{\partial x}\Big(\left(\beta_- - \beta_+\right)n\Big) \\
&+\frac{\Delta_x^2}{2\Delta_t}\frac{\partial}{\partial x}\left((\beta_- + \beta_+)\frac{\partial n}{\partial x} - n\frac{\partial}{\partial x}(\beta_- + \beta_+)\right) + \frac{1}{\Delta_t}\gamma n, \qquad (5.24)
\end{aligned}$$

at $y = Y_{\min}$. Recalling Equations (5.5)–(5.9) such that

$$\mu_{\pm} = \frac{D^p \Delta_t}{\Delta_y^2} \mp \frac{v^p \Delta_t}{2\Delta_y}, \tag{5.25}$$

in the limit $\Delta_y, \Delta_t \to 0$. Then Equation (5.24) can be rewritten (ignoring higher order corrections) as

$$\begin{aligned}\frac{\partial n}{\partial t} = \frac{1}{\Delta_y} v^p n + \frac{\partial}{\partial y}\left(\frac{1}{2} v^p n + \frac{1}{\Delta_y} D^p n\right) + \frac{1}{2}\frac{\partial^2}{\partial y^2}\left(D^p n\right) + \frac{\partial}{\partial x}(v^m n) \\ + \frac{\partial}{\partial x}\left(D^m \frac{\partial n}{\partial x} - n \frac{\partial}{\partial x} D^m\right) + rn.\end{aligned}$$

Therefore, in order to prevent blow-up of terms at $y = Y_{\min}$, we require that

$$v^p n + \frac{\partial}{\partial y}(D^p n) = 0 \qquad \text{at} \quad y = Y_{\min}.$$

Boundary condition at $y = Y_{\max}$. Returning to Equation (5.1), we seek an equation for the evolution of the cell number on the upper most lattice site in phenotype space, corresponding to the site y_{N_y+1}. To find this, we multiply Equation (5.1) by $n_s^{N_y+1}$ and sum over all possible states $\mathbf{n}$ and $\mathbf{e}$ gives

$$\begin{aligned}\Delta_t \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^{N_y+1} \frac{\partial}{\partial t} p(\mathbf{n}, \mathbf{e}, t_h) = \\ \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^{N_y+1} \sum_{i=1}^{N_x+1} \sum_{j=1}^{N_y} \mu_-(j+1, N_i, e_i) \left\{(n_i^{j+1}+1) p(U_{i,j}^{\mathrm{p}} \mathbf{n}, \mathbf{e}, t_h) - n_i^{j+1} p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\ + \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^{N_y+1} \sum_{i=1}^{N_x+1} \sum_{j=2}^{N_y+1} \mu_+(j-1, N_i, e_i) \big\{(n_i^{j-1}+1) p(D_{i,j}^{\mathrm{p}} \mathbf{n}, \mathbf{e}, t_h) \\ - n_i^{j-1} p(\mathbf{n}, \mathbf{e}, t_h)\big\} \\ + \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^{N_y+1} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y+1} \beta_-(j, N_i, e_i) \left\{(n_{i+1}^{j}+1) p(R_{i,j}^{\mathrm{m}} \mathbf{n}, \mathbf{e}, t_h) - n_{i+1}^{j} p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\ + \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^{N_y+1} \sum_{i=2}^{N_x+1} \sum_{j=1}^{N_y+1} \beta_+(j, N_i, e_i) \left\{(n_{i-1}^{j}+1) p(L_{i,j}^{\mathrm{m}} \mathbf{n}, \mathbf{e}, t_h) - n_{i-1}^{j} p(\mathbf{n}, \mathbf{e}, t_h)\right\} \\ + \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^{N_y+1} \sum_{i=1}^{N_x+1} \sum_{j=1}^{N_y+1} \gamma(j, N_i, e_i) \{\gamma(j, N_i - 1, e_i)(n_i^j - 1) p(G_{i,j} \mathbf{n}, \mathbf{e}, t_h) \\ - \gamma(j, N_i, e_i) n_i^j p(\mathbf{n}, \mathbf{e}, t_h)\}.\end{aligned} \tag{5.26}$$

Using the same methods as on the boundaries in physical space, we can change variables in each term to find an equation for the evolution of cell number. Consider first the first term on the right-hand side. From previous analysis, we know that the only non-zero contributions come from when $j = N_y$ and $i = s$. In the second term, the non-zero terms are $j = N_y + 1$ and $i = s$. The third term gives contributions when $j = N_y + 1$ and $i = s$ or $i = s - 1$. In the fourth term, there are non-zero contributions when $j = N_y + 1$ and $i = s$ or $i = s + 1$. The final term produces non-zero contributions only when $i = s$ and $j = N_y + 1$. Thus, Equation (5.26) can be written as

$$\begin{aligned}
\Delta_t \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^{N_y+1} \frac{\partial}{\partial t} p(\mathbf{n}, \mathbf{e}, t_h) = \\
\sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^{N_y+1} \mu_-(N_y+1, N_s, e_s) \big\{ (n_s^{N_y+1} + 1) p(U^{\mathrm{p}}_{s,N_y} \mathbf{n}, \mathbf{e}, t_h) - n_s^{N_y+1} p(\mathbf{n}, \mathbf{e}, t_h) \big\} \\
+ \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^{N_y+1} \mu_+(N_y, N_s, e_s) \big\{ (n_s^{N_y} + 1) p(D^{\mathrm{p}}_{s,N_y+1} \mathbf{n}, \mathbf{e}, t_h) - n_s^{N_y} p(\mathbf{n}, \mathbf{e}, t_h) \big\} \\
+ \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^{N_y+1} \sum_{i=s,s-1} \beta_-(N_y+1, N_i, e_i) \big\{ (n_{i+1}^{N_y+1} + 1) p(R^{\mathrm{m}}_{i,N_y+1} \mathbf{n}, \mathbf{e}, t_h) \\
- n_{i+1}^{N_y+1} p(\mathbf{n}, \mathbf{e}, t_h) \big\} \\
+ \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^{N_y+1} \sum_{i=s,s+1} \beta_+(N_y+1, N_i, e_i) \big\{ (n_{i-1}^{N_y+1} + 1) p(L^{\mathrm{m}}_{i,N_y+1} \mathbf{n}, \mathbf{e}, t_h) \\
- n_{i-1}^{N_y+1} p(\mathbf{n}, \mathbf{e}, t_h) \big\} \\
+ \sum_{\mathbf{n}} \sum_{\mathbf{e}} n_s^{N_y+1} \{ \gamma(N_y+1, N_s-1, e_s)(n_s^{N_y+1} - 1) p(G_{s,N_y+1} \mathbf{n}, \mathbf{e}, t_h) \\
- \gamma(N_y+1, N_s, e_s) n_s^{N_y+1} p(\mathbf{n}, \mathbf{e}, t_h) \}. \qquad (5.27)
\end{aligned}$$

The first term in Equation (5.27) can be rewritten using the change of variables ($\bar{n}_s^{N_y+1} = n_s^{N_y+1} + 1$ and $\bar{n}_s^{N_y} = n_s^{N_y} - 1$) in the following way:

$$
\begin{aligned}
&\sum_{\mathbf{n}}\sum_{\mathbf{e}} n_s^{N_y+1}\mu_-(N_y+1,N_s,e_s)\big\{(n_s^{N_y+1}+1)p(U^{\mathrm{p}}_{s,N_y}\mathbf{n},\mathbf{e},t_h) - n_s^{N_y+1}p(\mathbf{n},\mathbf{e},t_h)\big\}\\
&\quad=\sum_{\mathbf{n}}\sum_{\mathbf{e}} n_s^{N_y+1}\mu_-(N_y+1,N_s,e_s)\times\\
&\qquad\big\{(n_s^{N_y+1}+1)p([n_s^1,\ldots,n_s^{N_y}-1,n_s^{N_y+1}-1],\mathbf{e},t_h) - n_s^{N_y+1}p(\mathbf{n},\mathbf{e},t_h)\big\}\\
&\quad=\sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_-(N_y+1,N_s,e_s)\times\\
&\qquad\big\{\bar{n}_s^{N_y+1}(\bar{n}_s^{N_y+1}-1)p([n_s^1,\ldots,\bar{n}_s^{N_y},\bar{n}_s^{N_y+1}],\mathbf{e},t_h) - n_s^{N_y+1}n_s^{N_y+1}p(\mathbf{n},\mathbf{e},t_h)\big\}\\
&\quad=\sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_-(N_y+1,N_s,e_s)\big\{n_s^{N_y+1}(n_s^{N_y+1}-1)p(\mathbf{n},\mathbf{e},t_h)\\
&\qquad\qquad - n_s^{N_y+1}n_s^{N_y+1}p(\mathbf{n},\mathbf{e},t_h)\big\}\\
&\quad=-\sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_-(N_y+1,N_s,e_s)n_s^{N_y+1}p(\mathbf{n},\mathbf{e},t_h)\\
&\quad=-\langle\mu_-(N_y+1,N_s,e_s)n_s^{N_y+1}\rangle.
\end{aligned}
$$

The second term in Equation (5.27) can then be rewritten as the following (using $\bar{n}_s^{N_y+1} = n_s^{N_y+1} - 1$ and $\bar{n}_s^{N_y} = n_s^{N_y} + 1$):

$$
\begin{aligned}
&\sum_{\mathbf{n}}\sum_{\mathbf{e}} n_s^{N_y+1}\mu_+(N_y,N_s,e_s)\big\{(n_s^{N_y}+1)p(D^{\mathrm{p}}_{s,N_y+1}\mathbf{n},\mathbf{e},t_h) - n_s^{N_y}p(\mathbf{n},\mathbf{e},t_h)\big\}\\
&\quad=\sum_{\mathbf{n}}\sum_{\mathbf{e}} n_s^{N_y+1}\mu_+(N_y,N_s,e_s)\times\\
&\qquad\big\{(n_s^{N_y}+1)p([n_s^1,\ldots,n_s^{N_y}+1,n_s^{N_y+1}-1],\mathbf{e},t_h) - n_s^{N_y}p(\mathbf{n},\mathbf{e},t_h)\big\}\\
&\quad=\sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_+(N_y,N_s,e_s)\times\\
&\qquad\big\{\bar{n}_s^{N_y}(\bar{n}_s^{N_y+1}+1)p([n_s^1,\ldots,\bar{n}_s^{N_y},\bar{n}_s^{N_y+1}],\mathbf{e},t_h) - n_s^{N_y}n_s^{N_y+1}p(\mathbf{n},\mathbf{e},t_h)\big\}\\
&\quad=\sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_+(N_y,N_s,e_s)\times\\
&\qquad\big\{n_s^{N_y}(n_s^{N_y+1}+1)p(\mathbf{n},\mathbf{e},t_h) - n_s^{N_y}n_s^{N_y+1}p(\mathbf{n},\mathbf{e},t_h)\big\}\\
&\quad=\sum_{\mathbf{n}}\sum_{\mathbf{e}} \mu_+(N_y,N_s,e_s)n_s^{N_y}p(\mathbf{n},\mathbf{e},t_h)\\
&\quad=\langle\mu_+(N_y,N_s,e_s)n_s^{N_y}\rangle.
\end{aligned}
$$

The contributions from the terms describing movement in physical space are the same as in the main body Equation (5.3) where $j = N_y$. This is also true for the growth

terms with $i = s$ and $j = N_y$. Putting these together, we have

$$\begin{aligned}\frac{\partial}{\partial t}\langle n_s^{N_y+1}\rangle &= \frac{1}{\Delta_t}\langle \mu_+(N_y, N_s, e_s) n_s^{N_y}\rangle - \frac{1}{\Delta_t}\langle \mu_-(N_y+1, N_s, e_s) n_s^{N_y+1}\rangle \\ &+ \frac{1}{\Delta_t}\langle \beta_+(N_y+1, N_s, e_s) n_{s-1}^{N_y+1}\rangle + \frac{1}{\Delta_t}\langle \beta_-(N_y+1, N_s, e_s) n_{s+1}^{N_y+1}\rangle \\ &- \frac{1}{\Delta_t}\langle \beta_-(N_y+1, N_{s-1}, e_{s-1}) n_s^{N_y+1}\rangle \\ &- \frac{1}{\Delta_t}\langle \beta_+(N_y+1, N_{s+1}, e_{s+1}) n_s^{N_y+1}\rangle \\ &+ \frac{1}{\Delta_t}\langle \gamma(N_y+1, N_s, e_s) n_s^{N_y+1}\rangle. \end{aligned} \tag{5.28}$$

We can take the limit $\Delta_x, \Delta_y, \Delta_t \to 0$, such that Equation (5.28) can be rewritten as the following at $y = Y_{\max}$:

$$\begin{aligned}\frac{\partial}{\partial t} n(x, y, t) &= \frac{1}{\Delta_t}\beta_+(y, \rho(x,t), e(x,t)) n(x - \Delta_x, y, t) \\ &+ \frac{1}{\Delta_t}\beta_-(y, \rho(x,t), e(x,t)) n(x + \Delta_x, y, t) \\ &- \frac{1}{\Delta_t}\beta_-(y, \rho(x - \Delta_x, t), e(x - \Delta_x, t)) n(x, y, t) \\ &- \frac{1}{\Delta_t}\beta_+(y, \rho(x + \Delta_x, t), e(x + \Delta_x, t)) n(x, y, t) \\ &+ \frac{1}{\Delta_t}\mu_+(y - \Delta_y, \rho(x,t), e(x,t)) n(x, y - \Delta_y, t) \\ &- \frac{1}{\Delta_t}\mu_-(y, \rho(x,t), e(x,t)) n(x, y, t) \\ &+ \frac{1}{\Delta_t}\gamma(y, \rho(x,t), e(x,t)) n(x, y, t). \end{aligned}$$

Using Taylor series expansions, we find

$$\begin{aligned}
\frac{\partial}{\partial t}n(x,y,t) = &\frac{1}{\Delta_t}\beta_+(y,\rho(x,t),e(x,t))\times \\
&\qquad\left[n(x,y,t) - \Delta_x\frac{\partial}{\partial x}n(x,y,t) + \frac{\Delta_x^2}{2}\frac{\partial^2}{\partial x^2}n(x,y,t)\right] \\
&+\frac{1}{\Delta_t}\beta_-(y,\rho(x,t),e(x,t))\times \\
&\qquad\left[n(x,y,t) + \Delta_x\frac{\partial}{\partial x}n(x,y,t) + \frac{\Delta_x^2}{2}\frac{\partial^2}{\partial x^2}n(x,y,t)\right] \\
&-\frac{1}{\Delta_t}n(x,y,t)\times \\
&\qquad\left[\beta_-(y,\rho(x,t),e(x,t)) - \Delta_x\frac{\partial}{\partial x}\beta_-(y,\rho(x,t),e(x,t)\right. \\
&\qquad\qquad\left.+\frac{\Delta_x^2}{2}\frac{\partial^2}{\partial x^2}\beta_-(y,\rho(x,t),e(x,t)\right] \\
&-\frac{1}{\Delta_t}n(x,y,t)\times \\
&\qquad\left[\beta_+(y,\rho(x,t),e(x,t)) + \Delta_x\frac{\partial}{\partial x}\beta_+(y,\rho(x,t),e(x,t)\right. \\
&\qquad\qquad\left.+\frac{\Delta_x^2}{2}\frac{\partial^2}{\partial x^2}\beta_+(y,\rho(x,t),e(x,t)\right] \\
&+\frac{1}{\Delta_t}\left[n(x,y,t) - \Delta_y\frac{\partial}{\partial y}n(x,y,t) + \frac{\Delta_y^2}{2}\frac{\partial^2}{\partial y^2}n(x,y,t)\right]\times \\
&\qquad\left[\mu_+(y,\rho(x,t),e(x,t)) - \Delta_y\frac{\partial}{\partial y}\mu_+(y,\rho(x,t),e(x,t)\right. \\
&\qquad\qquad\left.+\frac{\Delta_y^2}{2}\frac{\partial^2}{\partial y^2}\mu_+(y,\rho(x,t),e(x,t)\right] \\
&-\frac{1}{\Delta_t}\mu_-(y,\rho(x,t),e(x,t))n(x,y,t) \\
&+\frac{1}{\Delta_t}\gamma(y,\rho(x,t),e(x,t))n(x,y,t).
\end{aligned}$$

Now, dropping the dependent variables for simplicity, we find

$$\begin{aligned}
\frac{\partial n}{\partial t} = &-\frac{1}{\Delta_t}(\mu_- - \mu_+)n - \frac{\Delta_y}{\Delta_t}\frac{\partial}{\partial y}(\mu_+ n) + \frac{\Delta_y^2}{2\Delta_t}\frac{\partial^2}{\partial y^2}(\mu_+ n) + \frac{\Delta_x}{\Delta_t}\frac{\partial}{\partial x}\Big((\beta_- - \beta_+)\,n\Big) \\
&+\frac{\Delta_x^2}{2\Delta_t}\frac{\partial}{\partial x}\left((\beta_- + \beta_+)\frac{\partial n}{\partial x} - n\frac{\partial}{\partial x}(\beta_- + \beta_+)\right) + \frac{1}{\Delta_t}\gamma n, \qquad (5.29)
\end{aligned}$$

at $y = Y_{\max}$. Recalling Equations (5.5)–(5.9) and (5.25), then Equation (5.29) can be rewritten (ignoring higher order corrections) as

$$\frac{\partial n}{\partial t} = -\frac{1}{\Delta_y} v^p n + \frac{\partial}{\partial y}\left(\frac{1}{2} v^p n - \frac{1}{\Delta_y} D^p n\right) + \frac{1}{2}\frac{\partial^2}{\partial y^2}\left(D^p n\right) + \frac{\partial}{\partial x}(v^m n) + \frac{\partial}{\partial x}\left(D^m \frac{\partial n}{\partial x} - n\frac{\partial}{\partial x} D^m\right) + rn.$$

Therefore, in order to prevent blow-up of terms at $y = Y_{\max}$, we require

$$-v^p n - \frac{\partial}{\partial y}(D^p n) = 0 \qquad \text{at} \quad y = Y_{\max}.$$

Complete set of boundary conditions for Equation (5.10)**.** In summary, to complement the continuum equation (5.10) there are zero flux boundary conditions at $x = X_{\min}, X_{\max}$ and $y = Y_{\min}, Y_{\max}$, given by

$$v^m n + D^m \frac{\partial n}{\partial x} - n\frac{\partial D^m}{\partial x} = 0 \qquad \text{at} \quad x = X_{\min}, \tag{5.30}$$

$$-v^m n - D^m \frac{\partial n}{\partial x} + n\frac{\partial D^m}{\partial x} = 0 \qquad \text{at} \quad x = X_{\max}. \tag{5.31}$$

on the physical domain and

$$v^p n + \frac{\partial}{\partial y}(D^p n) = 0 \qquad \text{at} \quad y = Y_{\min}, \tag{5.32}$$

$$-v^p n - \frac{\partial}{\partial y}(D^p n) = 0 \qquad \text{at} \quad y = Y_{\max}, \tag{5.33}$$

at the ends of the phenotype domain. The differences in the boundary conditions in phenotype and physical space are observed as a result of the different assumptions underlying the movement probabilities in each space. In physical space, the probability of movement depends on the surrounding number of cells and elements of the local environment in the target site, whereas the probability of movement in phenotype space depends on the number of cells and elements of the local environment in the same site as the cell of interest (at the individual cell level).

5.3.2 Equation for the density of the local environment

Following the assumptions outlined in Section 5.2 and the methodology above, we can review the master equation (5.1) describing the evolution of the number of elements of the local environment at position x_i, denoted e_i, and multiply by e_s and sum over

all possible states $\mathbf{e}$ and $\mathbf{n}$ to give

$$\begin{aligned}
&\sum_{\mathbf{e}}\sum_{\mathbf{n}} e_s \Delta_t \frac{\partial}{\partial t} p(\mathbf{n},\mathbf{e},t_h) \\
&= \sum_{\mathbf{e}}\sum_{\mathbf{n}} e_s \sum_{i=1}^{N_x+1}\sum_{j=1}^{N_y+1} \lambda(j,n_i^j)\left\{(e_i+1)p(\mathbf{n},H_i\mathbf{e},t_h) - e_i p(\mathbf{n},\mathbf{e},t_h)\right\} \\
&= \sum_{\mathbf{e}}\sum_{\mathbf{n}} e_s \sum_{i=1}^{N_x+1}\sum_{j=1}^{N_y+1} \lambda(j,n_i^j)\big\{(e_i+1)p(\mathbf{n},[e_1,\ldots,e_i+1,\ldots,e_{N_x+1}],t_h) \\
&\qquad\qquad - e_i p(\mathbf{n},\mathbf{e},t_h)\big\},
\end{aligned}$$

recalling that contributions from the terms describing cell dynamics sum to zero. Now we consider two cases: $i = s$ and $i \neq s$. First, consider $i \neq s$ and use the change of variables $\bar{e}_i = e_i + 1$ in the second term, and then drop the bars to give

$$\begin{aligned}
&\sum_{\mathbf{e}}\sum_{\mathbf{n}} e_s \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1}\sum_{j=1}^{N_y+1} \lambda(j,n_i^j)\left\{(e_i+1)p(\mathbf{n},[e_1,\ldots,e_i+1,\ldots,e_{N_x+1}],t_h) - e_i p(\mathbf{n},\mathbf{e},t_h)\right\} \\
&= \sum_{\mathbf{e}}\sum_{\mathbf{n}} e_s \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1}\sum_{j=1}^{N_y+1} \lambda(j,n_i^j)\left\{\bar{e}_i p(\mathbf{n},[e_1,\ldots,\bar{e}_i,\ldots,e_{N_x+1}],t_h) - e_i p(\mathbf{n},\mathbf{e},t_h)\right\} \\
&= \sum_{\mathbf{e}}\sum_{\mathbf{n}} e_s \sum_{\substack{i=1,\\ i\neq s}}^{N_x+1}\sum_{j=1}^{N_y+1} \lambda(j,n_i^j)\left\{e_i p(\mathbf{n},\mathbf{e},t_h) - e_i p(\mathbf{n},\mathbf{e},t_h)\right\} \\
&= 0.
\end{aligned}$$

Now consider the case when $i = s$ and use the change of variables $\bar{e}_s = e_s + 1$ to give

$$\begin{aligned}
&\sum_{\mathbf{e}}\sum_{\mathbf{n}} \sum_{j=1}^{N_y+1} e_s \lambda(j,n_s^j)\left\{(e_s+1)p(\mathbf{n},[e_1,\ldots,e_s+1,\ldots,e_{N_x+1}],t_h) - e_s p(\mathbf{n},\mathbf{e},t_h)\right\} \\
&= \sum_{\mathbf{e}}\sum_{\mathbf{n}} \sum_{j=1}^{N_y+1} \lambda(j,n_s^j)\left\{\bar{e}_s(\bar{e}_s-1)p(\mathbf{n},[e_1,\ldots,\bar{e}_s,\ldots,e_{N_x+1}],\mathbf{e},t_h) - e_s^2 p(\mathbf{n},\mathbf{e},t_h)\right\} \\
&= \sum_{\mathbf{e}}\sum_{\mathbf{n}} \sum_{j=1}^{N_y+1} e_s\lambda(j,n_s^j)\left\{(e_s-1)p(\mathbf{n},\mathbf{e},t_h) - e_s p(\mathbf{n},\mathbf{e},t_h)\right\} \\
&= -\sum_{\mathbf{e}}\sum_{\mathbf{n}} \sum_{j=1}^{N_y+1} \lambda(j,n_s^j) e_s p(\mathbf{n},\mathbf{e},t_h) \\
&= -\sum_{j=1}^{N_y+1} \langle \lambda(j,n_s^j) e_s \rangle.
\end{aligned}$$

Putting this together, we arrive at

$$\Delta_t \frac{\partial}{\partial t}\langle e_s \rangle = -\sum_{j=1}^{N_y+1} \langle \lambda(j, n_s^j) e_s \rangle. \tag{5.34}$$

Defining

$$\lim_{\Delta_t \to 0} \frac{1}{\Delta_t} \lambda(y, n(x, y, t)) = \nu(y, n(x, y, t)),$$

which we can substitute into Equation (5.34), rearrange and take limits as $\Delta_x, \Delta_y, \Delta_t \to 0$, we find that the differential equation for the density of the local environment, $e(x, t)$, is given by

$$\frac{\partial}{\partial t} e(x, t) = -\int_{y=Y_{\min}}^{y=Y_{\max}} \nu(y, n(x, y, t)) e(x, t) \mathrm{d}y. \tag{5.35}$$

No boundary conditions are required for this equation, however the corresponding initial condition is

$$e(x, 0) = e_0(x). \tag{5.36}$$

Now that we have derived the coarse-grained model in full (Equations (5.10), (5.11), (5.30)–(5.35) and (5.36)), we present a series of applications that demonstrate the utility of this framework through the choice of specific functional forms for $v^m\left(y, \rho(x,t), e(x,t)\right)$, $D^m\left(y, \rho(x,t), e(x,t)\right)$, $v^p\left(y, \rho(x,t), e(x,t)\right)$, $D^p\left(y, \rho(x,t), e(x,t)\right)$, $r\left(y, \rho(x,t), e(x,t)\right)$ and $\nu(y, n(x, y, t))$. We will assume in this chapter that all movement in physical space is undirected, and therefore we take $\beta_+(y, \rho(x,t), e(x,t)) = \beta_-(y, \rho(x,t), e(x,t)$, such that $v^m = 0$ hereon in. Nevertheless, the general form of the governing equations is retained, so that the framework can be readily adapted and generalised to cases involving directed movement or other specific applications.

5.4 Broad spectrum applications in mathematical biology

In this section, we showcase the versatility of the PDE modelling framework given by Equations (5.10), (5.11), (5.30)–(5.35) and (5.36) by applying it to several exemplar biological scenarios. These applications demonstrate how the PDE framework effectively captures emergent population-level dynamics across diverse biological contexts. By considering a range of different underlying characteristics and interaction rules (prescribed in Appendix B.1), we showcase the ability of these models to encode complex behaviours while maintaining analytical and computational tractability.

5.4.1 Simulation methods

The deterministic, continuum counterpart of the IBM described in Section 5.2 is given by the PDEs in Equations (5.10) and (5.35), with boundary conditions given in Equations (5.30)–(5.33) and initial conditions given in Equations (5.11) and (5.36). To solve this system numerically, we use an advection–diffusion–reaction (A–DR) scheme that discretises the spatial variable x using a central finite difference stencil modified from Chapter 2, employing ghost points to enforce the zero flux boundary conditions. In the phenotypic axis, y, we use a finite volume scheme, which divides the axis into N_y+1 sites of equal width. The advective component is controlled using the Koren limiter [177]. The resulting system of ODEs takes the following form

$$\frac{\mathrm{d}n_i^j}{\mathrm{d}t} = \frac{((D^m)_{i+1}^j + (D^m)_i^j)(n_{i+1}^j - n_i^j) - ((D^m)_i^j + (D^m)_{i-1}^j)(n_i^j - n_{i-1}^j)}{2\Delta x^2} + \frac{n_i^{j+1} - 2n_i^j + n_i^{j-1}}{\Delta y^2}(D^p)_i^j + \mathcal{A}_{i,j} + n_i^j r(\bar{y}_j, N_i, e_i)$$

for $i = 1, \ldots, N_x + 1$, and $j = 1, \ldots, N_y + 1$, with appropriate substitutions at the boundary terms and with

$$\begin{aligned} N_i &= \sum_j n_i^j \Delta y, \\ (D^m)_i^j &= D^m(\bar{y}_j, N_i, e_i), \\ (D^p)_i^j &= D^p(\bar{y}_j, N_i, e_i), \\ \bar{y}_j &= \text{mean of } y_j \text{ and } y_{j+1}, \\ \mathcal{A}_{i,j} &= \text{flux-limited advection in } y. \end{aligned}$$

The advection term in the y-direction is discretised using a slope-limited upwind scheme and can be written as

$$\left(\frac{\partial}{\partial y}(v^p n)\right)_i^j \approx \frac{F_i^{j+1/2} - F_i^{j-1/2}}{\Delta y},$$

where $F_i^{j+1/2}$ is the numerical flux across the interface between phenotypic points y_j and y_{j+1}, given by

$$F_i^{j+1/2} = \begin{cases} (v^p)_i^{j+1/2} n_i^{j,+} & \text{if } (v^p)_i^{j+1/2} > 0, \\ (v^p)_i^{j+1/2} n_i^{j+1,-} & \text{if } (v^p)_i^{j+1/2} \leq 0, \end{cases}$$

which is calculated using

$$
\begin{aligned}
R_i^j &= \frac{n_i^{j+1} - n_i^j}{n_i^j - n_i^{j-1}}, \\
n_i^{j,+} &= n_i^j + \frac{1}{2}\phi(R_i^j)(n_i^{j+1} - n_i^j), \\
n_i^{j+1,-} &= n_i^{j+1} - \frac{1}{2}\phi(1/R_i^j)(n_i^{j+1} - n_i^j).
\end{aligned}
$$

Alongside this, to simulate the evolution of the density of the local environment, as described by Equation (5.35), we use a finite difference scheme, with a summation to approximate the integral, which can be written as

$$
\frac{de_i}{dt} = -e_i \sum_j \nu(\bar{y_j}, n_i^j)\Delta_y.
$$

The resulting system of ODEs is then integrated in time using python's in-built ODE solver `scipy.integrate.solve_ivp` with the explicit Runge-Kutta integration method of order 5 and time step $\Delta_t = 0.1$. The phenotype step is $\Delta_y = 0.02$ and the spatial step is $\Delta_x = 0.1$. Where constant speed, constant profile travelling waves are observed, the speed is estimated numerically by saving the solution at each time point, interpolating to find the critical spatial position x^* such that $\rho(x^*, t) = 0.1$ and then calculating the difference between two sequential critical spatial points and dividing by the time step between them. At each spatial position, the mean phenotype is obtained by computing the density-weighted sum of phenotypes and normalising by the total local density. Code is available for all computations at the following GitHub repository: `https://github.com/beckycrossley/cont_phen`.

5.4.2 Phenotypic structuring during population expansion

Understanding how cell populations expand and evolve is a fundamental question in biology, particularly in contexts such as tumour growth, microbial colony expansion, and tissue development. A key aspect of these processes is tracking cell lineages to uncover how phenotypic traits propagate and shape population dynamics over time. In this section, we demonstrate how the modelling framework developed in this chapter provides a convenient and effective approach for studying such lineage dynamics within an evolving population. Specifically, we consider a phenotypically structured population of homogeneous cells (*i.e.*, cells that share the same underlying behaviour but are distinguishable by a phenotypic marker) to gain deeper insights into how individual lineages contribute to the overall invasion process. By analysing

the spatio-temporal evolution of the phenotypic structure as the population spreads, we highlight how this approach enables the systematic tracking of cell lineages during population expansion.

Previous studies, such as those by [200, 340], have investigated similar population dynamics using coupled differential and stage-structured integrodifference equations [201], with further extensions incorporating trade-offs between reproductive and dispersal abilities [199]. While these approaches offer valuable insights into structured population dynamics, they rely on discrete phenotypic stages, which may limit the resolution of evolutionary and ecological interactions.

In contrast, this work provides a more nuanced perspective by modelling the evolution of a continuously structured phenotype, $y \in [0, 1]$, over space, $x \geq 0$, and time, $t \geq 0$. This allows for a finer representation of phenotypic variation and subsequent exploration of its role during population expansion, whilst, in this example, ignoring the impact of the environment. Specifically, we describe the spatio-temporal evolution of the cell population, $n(x, y, t)$, using the following governing equation:

$$\frac{\partial}{\partial t} n(x, y, t) = \frac{\partial^2}{\partial x^2} n(x, y, t) + r\left(\rho(x, t)\right) n(x, y, t). \tag{5.37}$$

As in [200], we study Equation (5.37) subject to two different functions describing net cell proliferation:

$$r_K(y, \rho(x, t)) = 1 - \rho(x, t), \tag{5.38}$$

for Fisher–KPP type invasion (pulled waves) and

$$r_A(y, \rho(x, t)) = (1 - \rho(x, t))(\rho(x, t) - p^*), \tag{5.39}$$

for the Allee effect, with $p^* \in (0, 1/2)$ (corresponding to pushed waves), where $\rho(x, t) = \int_{y=0}^{y=1} n(x, y, t) \mathrm{d}y$. The individual-based functions underlying these continuum equations can be found in Appendix B.1.1. We compliment this setup with an initial condition that ensures initial phenotypic structuring of the population, which can then be tracked over time. Specifically, we take

$$u_0(x, y) = \begin{cases} 1 & \text{if } x = 5y, \\ 0 & \text{otherwise.} \end{cases} \tag{5.40}$$

Recalling that cells in this case are homogeneous in phenotype and behaviour, and thus the phenotypic variable $y \in [0, 1]$ is used purely to label cells as they evolve, we can see that Figure 5.1 shows the phenotypic structure of the population of cells as

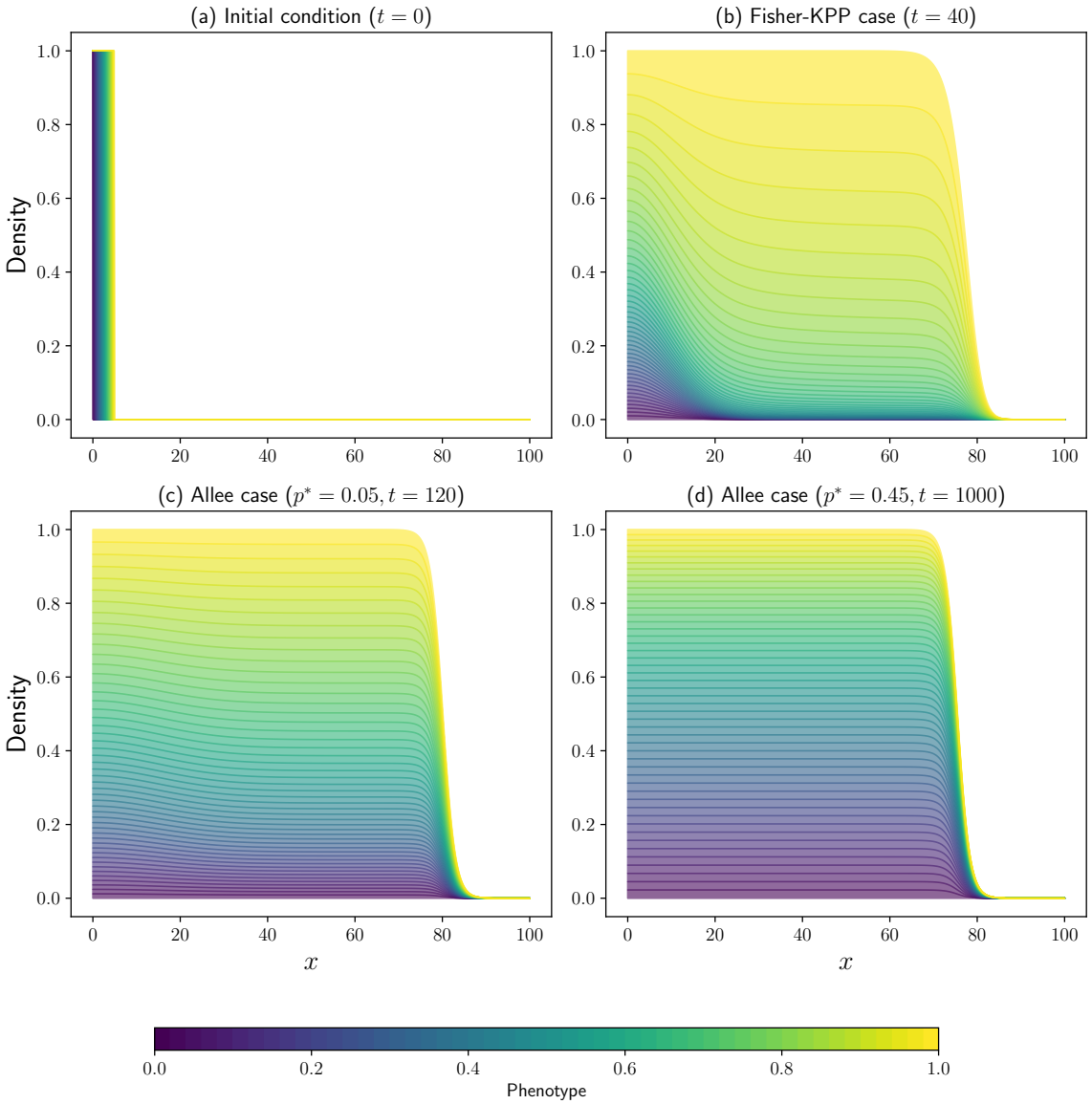

Figure 5.1: Evolution of the phenotypic structure of cells in Equation (5.37) subject to various growth terms. (a) The initial distribution of the cells with different phenotypes. (b) The spatial structure of the invading wave subject to the Fisher–KPP growth term (Equation (5.38)). Results in (c) and (d) show the spatial structure of the invading wave subject to the Allee effect (Equation (5.39)).

they invade, subject to the aforementioned growth terms (Equation (5.38) and Equation (5.39)). Specifically, Figure 5.1(b) shows the spatial structure of Equation (5.37) with the Fisher–KPP growth term (Equation (5.38)). As invasion progresses, the leading edge of the expanding population is dominated by cells that originated from the rightmost part of the initial distribution, with phenotypes closer to $y = 1$. Over time, the density of these dominant phenotypes increases and, due to diffusion, cells with these traits spread backward, integrating into the bulk of the population.

This is a form of what is known as *surfing* [172]—a phenomenon well-studied in the context of drifting genetic mutations in expanding populations. Surfing occurs when areas of low cell density allow space for increased growth, such as along the invading front (see Figure 5.1) [103].

The structure observed in the case of the Fisher–KPP type growth is often described as a 'vertical pattern' [200], which indicates that the total population is composed of different underlying phenotypes at different points in space, demonstrating a high level of spatial structuring. However, we see that even in a continuously structured cell population, it is those cells with the phenotypes that were nearest the front of the initial distribution of cells that dominate during population expansion. The lower phenotypes, which began at the rear of the invading population, primarily remain there and diminish in size over time and increasing space.

Alternatively, if we now consider the case of Allee growth (Equation (5.39)) in Equation (5.37), then the numerical results in Figure 5.1 display a different pattern of invasion. As the cells invade, a much larger proportion of all of the initially present phenotypes remain present in the travelling wave front at later times, and throughout the bulk of the invading population. As in the Fisher–KPP case, the cells in the right-most portion of the initial population, with phenotypes near unity, contribute the largest portions to the wave. However, with the addition of the Allee effect, there now exist contributions from all initially present phenotypic states throughout the wave. This is because the Allee effect introduces a dependency on the total local population density, which means cells at the leading edge, in areas of low density, have reduced proliferation, preventing them from rapidly outcompeting other phenotypes within the cell bulk. We note that by increasing the strength of the Allee effect, by taking p^* closer to $1/2$, the proportion of all of the phenotypes present in the wave becomes closer to one another (equalises). The spatial pattern observed in this case is described as 'horizontal' as it does not differ in space, but still shows high phenotypic variation in the front [283].

These structural differences at the wave front agree with those observed by [283] who consider a similar system but with discrete, rather than continuous, phenotypes. These results indicate that the lineage structure of an invading wave could potentially be used to distinguish between the underlying growth mechanisms of the cell populations. It is also notable that the speed of cell invasion subject to the Allee effect (Equation (5.39)) is significantly slower than the speed of invasion in the Fisher–KPP case (Equation (5.38)). As such, although the Allee effect maintains diversity across the travelling wave front, it also decreases the speed of invasion in doing so, and so a trade-off is observed between diversity maintenance and speed of population invasion, which favours faster invasion for a weaker Allee effect.

5.4.3 A go-or-grow model of cells invading into ECM

In this section, we apply the aforementioned system of Equations (5.10)–(5.33), (5.35) and (5.36) to a general model for collective cell migration into ECM, exploring how individual cell-level properties give rise to emergent population-level behaviours. By examining the interactions between cells and their environment at a population-level, we aim to uncover the underlying individual cell-ECM mechanisms that drive large-scale migration patterns and spatial organization, providing insight into how cell processes shape collective movement in biological systems.

In this work, we aim to model the fundamental trade-off in energetic costs associated with motility and proliferation, known as the 'go-or-grow' hypothesis [139], while also incorporating the role of ECM degradation by migrating cells. Our goal is to use this framework to bridge the gap between individual-cell behaviours and emergent population-level dynamics.

In earlier chapters, we captured this trade-off by considering only two discrete phenotypes [79]. However, biological evidence strongly supports the existence of a continuum of phenotypes rather than a simple dichotomy [25, 50]. By extending prior work to a continuously structured phenotypic model, we provide a more biologically realistic representation of go-or-grow dynamics, allowing us to explore how individual-level decision-making translates into large-scale invasion patterns.

We model the cells, denoted $n(x, y, t)$, as able to move, proliferate and degrade the ECM, $e(x, t)$. To model the trade-off between ECM degradation, motility and proliferation, we assume that cells with a larger value of the phenotype variable, $y \in [0, 1]$, have higher proliferative potential but lower motility and lower ECM degrading potential whilst, in comparison, cells with a lower value of the phenotypic variable y correspond to those with higher motility and higher ECM degrading potential, but

lower proliferative potential. Therefore, cells with phenotype $y = 0$ degrade ECM most rapidly and are the most motile, but they are unable to proliferate. On the other hand, cells with phenotype $y = 1$ are the most proliferative, but cannot degrade ECM or move. As in Chapter 4, we implement a linear relationship between the phenotype of the cells and the associated ability to degrade ECM, move and proliferate.

The initial functions describing the probabilities of transitions in the IBM underlying these continuum equations can be found in Appendix B.1.2. Hereon in, we focus on the corresponding continuum functions stated below.

Following the volume exclusion principles described in earlier chapters, we assume that the probability of a cell moving randomly in physical space linearly decreases as the occupancy of space increases. As described earlier, we also introduce the assumption that the probability of a cell moving randomly increases as the phenotype of the cell decreases. As such, the function describing movement in physical space can be written as

$$D^m(y, \rho(x,t), e(x,t)) = (1-y)\left(1 - \frac{\rho(x,t) + e(x,t)}{\kappa}\right),$$

where $\kappa > 0$ is the total available density for cells and ECM, known as the carrying capacity. Similarly, for all cells, the probability of cell proliferation linearly decreases as the space around the cell fills with other cells and ECM, but it also decreases as the phenotype decreases, such that

$$r(y, \rho(x,t), e(x,t)) = y\left(1 - \frac{\rho(x,t) + e(x,t)}{\kappa}\right).$$

Furthermore, we assume that the degradation rate of the ECM is proportional to the density of surrounding cells and decreases as the phenotype of the cells increases. We write this as

$$\nu(y, n(x,y,t)) = (1-y)n(x,y,t).$$

One major challenge in understanding the cell phenotype dynamics during invasion is the lack of direct experimental observations, as visualising phenotypic transitions in real time is extremely difficult. As a result, there is limited guidance in the literature on the appropriate mathematical forms for these transitions, leaving a key gap in the understanding of how phenotype-dependent behaviours shape collective migration.

This issue relates to the final undetermined functions in Equations (5.10) and (5.35), that describe the transitions between cell phenotypes. In this section, we systematically investigate the effects of different density-dependent phenotypic transition

rules, $\mu_{\pm}(y, \rho(x,t), e(x,t))$, along with their associated drift and diffusion terms, $v^p(y, \rho(x,t), e(x,t))$ and $D^p(y, \rho(x,t), e(x,t))$, respectively, as summarised in Table 5.1. By exploring how the invasion dynamics change under different assumptions, we aim to determine whether population-level (and therefore potentially more observable) behaviours can provide insight into the underlying, unobservable phenotypic structures, offering a potential approach for inferring hidden biological mechanisms from macroscopic invasion patterns.

Phenotypic drift	$\mu_-(y,\rho,e)$	$\mu_+(y,\rho,e)$	$v^p(y,\rho,e)$	$D^p(y,\rho,e)$
Cell-dependent	$\frac{\rho}{\kappa}$	$1-\frac{\rho}{\kappa}$	$2\frac{\rho}{\kappa}-1$	1
ECM-dependent	$\frac{e}{\kappa}$	$1-\frac{e}{\kappa}$	$2\frac{e}{\kappa}-1$	1
Space-dependent	$\frac{\rho+e}{\kappa}$	$1-\frac{\rho+e}{\kappa}$	$2\frac{\rho+e}{\kappa}-1$	1

Table 5.1: The functions employed in Equation (5.10) describing the probabilities of transitions up and down in phenotype space, resulting in the phenotypic drift, $v^p(y, \rho(x,t), e(x,t))$, and phenotypic diffusion, $D^p(y, \rho(x,t), e(x,t))$, functions shown.

We investigate three phenotypic drift mechanisms influencing cell transitions: firstly, cell-dependent drift, where cells shift toward more motile phenotypes at higher cell densities; next, ECM-dependent drift, where cells adopt more ECM-degrading phenotypes as ECM density increases; and finally, space-dependent drift, where cells become more proliferative as available space increases. Each drift term is bounded between zero and one, which is consistent with the constraints placed on the total cell and ECM density. By comparing these mechanisms for phenotype change, we assess how different environment- and density-dependent phenotypic transitions shape the structure of the migrating cell front.

We simulate this system of Equations (5.10) and (5.33)–(5.36) subject to the following non-compactly supported initial conditions for the cells and step initial conditions for the ECM:

$$n_0(x,y) = \frac{\exp\left(-100(x^2+y^2)\right)}{\max\left(\int_{X_{\min}}^{X_{\max}} \exp\left(-100(x^2+y^2)\right) \mathrm{d}x\right)}, \tag{5.41}$$

$$e_0(x) = \begin{cases} 0 & \text{if } \ n_0(x,y) > 0.001, \\ 0.5 & \text{otherwise.} \end{cases} \tag{5.42}$$

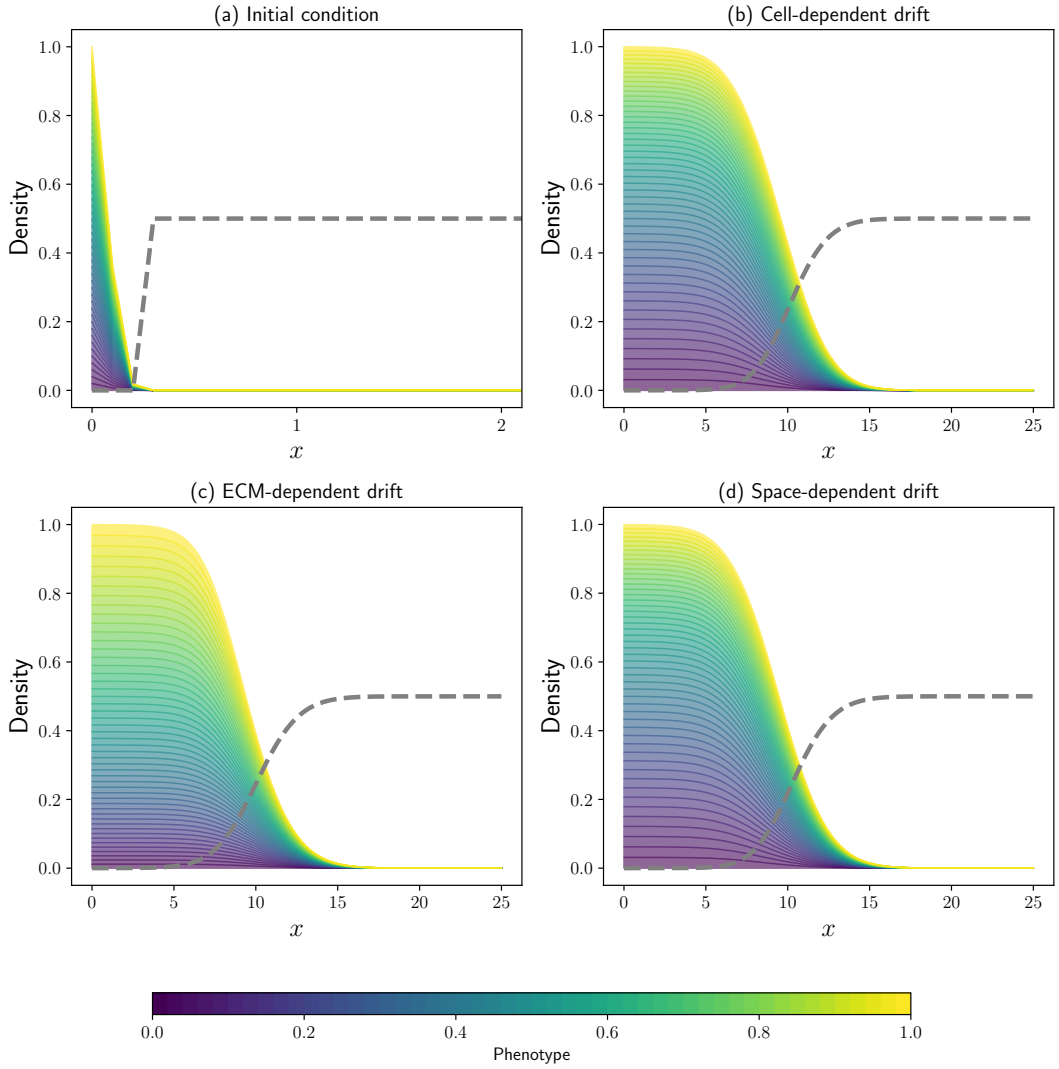

Figure 5.2: Evolution of the phenotypic structure of cells in Equations (5.10) and (5.35) subject to various phenotypic drift terms, with the corresponding ECM density shown as a dashed grey line. (a) The initial distribution of the ECM and the cells with different phenotypes. (b) The spatial structure of the invading wave subject to cell-dependent phenotypic drift. (c) The spatial structure of the invading wave subject to ECM-dependent phenotypic drift. (d) The spatial structure of the invading wave subject to space-dependent phenotypic drift. Results in (b), (c) and (d) are all plotted at time $t = 30$ and simulations are carried out with $\kappa = 1$. See Table 5.1 for explicit forms of the phenotypic drift terms.

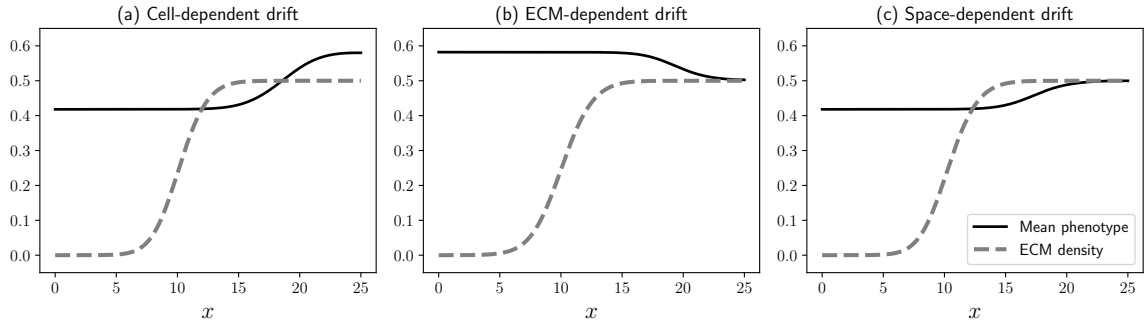

Figure 5.3: Evolution of the mean phenotype of the cells in Equations (5.10) and (5.35) subject to various phenotypic drift terms, with the corresponding ECM density shown as a dashed grey line. Results are all plotted at time $t = 30$ and simulations are carried out with $\kappa = 1$. See Table 5.1 for explicit forms of the phenotypic drift terms.

By examining the simulation results shown in Figure 5.2 we can see that, for a fully mixed initial population of cells, the phenotypic drift terms considered produce travelling wave solutions with similar, constant speeds of invasion.

In the case of cell-dependent phenotypic drift, we notice that there is a higher density of cells with lower phenotypes, which correspond to more motile and ECM degrading but less proliferative cells, in the bulk of the wave. At the front, we see a higher proportion of more proliferative and immobile cells, which are able to divide into the available space. As such, in the bulk of the wave the mean phenotype remains the same but, as x increases, we observe a gradual increase in mean phenotype at the wave front where extremely low cell densities induce cells to switch to a proliferative phenotype (see Figure 5.3). Very similar patterning is observed in the space-dependent phenotypic drift case, with a more gradual change in phenotype at the front of the travelling wave. In this context, this model predicts that the density of the ECM has minimal impact on the phenotypic structure of the invading wave.

However, when examining ECM-dependent phenotypic drift in Figure 5.2(c), we notice that the cells with higher phenotypes, corresponding to less motile and less degrading but more proliferative, now constitute a larger proportion of the population behind the invading front. Compared to cell- or space-dependent phenotypic drift, the mean phenotype of cells in the bulk is significantly higher. However, at the front of the invading wave in the ECM-dependent case, we see a larger proportion of cells with low phenotypes, namely, cells that have a higher ability to degrade the ECM and move into the subsequently available space.

As a result of the phenotypic structures developed when the system is subject to different phenotypic drift functions, the spatial structure of the individual phenotypes within the invading wave could be used to better understand the underlying mechanisms governing phenotypic transitions during collective cell migration into ECM.

5.4.4 T cell exhaustion

We now demonstrate the applicability of the framework to T cell exhaustion, starting from an individual-based description of the underlying dynamics. This approach systematically coarse-grains the underlying cell processes, ensuring that the resulting PDEs are not merely phenomenological but instead retain a direct mechanistic link to individual cell properties. Specifically, we incorporate a T cell population with varying levels of exhaustion invading into a tumour, and examine the role of phenotype-dependent drift in shaping its exhaustion dynamics. This application highlights how the framework developed in this chapter can be adapted to capture complex immune responses while preserving biologically meaningful connections between individual- and population-level behaviour.

T cells are a key component of the immune system, with an important role to play in locating and attacking tumour cells [355]. When space is available, T cells will infiltrate into the tumour where they kill malignant cells by releasing cytotoxic enzymes. During the sustained activation of T cells required to combat and restrict further growth of a tumour, T cells will differentiate and eventually "exhaust" [368]. This occurs as a result of continued exposure to the antigens of the tumour cells and, as T cells exhaust, their effector functions reduce [33, 67]. Exhaustion results in diminished cytokine production, impaired proliferation and reduced motility in the T cells [218]. Completely exhausted T cells are no longer able to move or grow [358], and have impaired toxicity, which reduces their ability to kill off tumour cells [154]. Furthermore, T cells will die much faster the more exhausted they are [358]. Understanding the dynamics of the T cells as they infiltrate a tumour and their exhaustion during this process is crucial to developing treatments for tumours, and so we develop a mathematical model to gain further insights.

As such, after applying the T cell-specific IBM functions described in the Appendix B.1.3 to the coarse-graining process, we investigate the spatio-temporal evolution of T cell density, denoted $T(x, y, t)$, invading into a population of tumour cells, with density denoted by $C(x, t)$. Similarly to before, we define $\rho(x, t)$ as

$$\rho(x,t) = \int_{y=Y_{\min}}^{y=Y_{\max}} T(x,y,t)\mathrm{d}y.$$

In this application, the phenotype of the cells, $y \in [0,1]$, represents the exhaustion of the T cells. As such, T cells with phenotype $y = 1$ are naïve, and are able to move freely and randomly in physical space, whilst also attacking nearby tumour cells and dividing [279, 363]. However, T cells with phenotype $y = 0$ are considered exhausted, or terminally-differentiated memory T cells, which are considered to be in a resting state [315, 335]. The resulting model takes the following form:

$$\begin{aligned}\frac{\partial}{\partial t}T(x,y,t) = \frac{\partial}{\partial x}\Bigg(& D^m\big(y,\rho(x,t),C(x,t)\big)\frac{\partial}{\partial x}T(x,y,t) \\ & - T(x,y,t)\frac{\partial}{\partial x}D^m\big(y,\rho(x,t),C(x,t)\big)\Bigg) \\ & + \frac{\partial}{\partial y}\Bigg(v^p\big(y,\rho(x,t),C(x,t)\big)T(x,y,t)\Bigg) \\ & + \frac{\partial^2}{\partial y^2}\Bigg(D^p\big(y,\rho(x,t),C(x,t)\big)T(x,y,t)\Bigg) \\ & + r\big(y,\rho(x,t),C(x,t)\big)T(x,y,t), \end{aligned} \tag{5.43}$$

where the net proliferation of the T cells, which depends on the exhaustion of the T cells and the available surrounding space for growth, can be described by

$$r(y,\rho(x,t),C(x,t)) = \gamma_1 y\left(1 - \frac{\rho(x,t)+C(x,t)}{\kappa}\right) - \gamma_0(1-y),$$

where $\gamma_1 \geq 0$ describes the growth rate, and $\gamma_0 \geq 0$ describes the death rate of the T cells. $\kappa > 0$ is the carrying capacity for the cells, as described in Section 5.4.3.

T cell movement in physical space is assumed to be random and undirected, but it is prevented by the presence of other T cells or tumour cells (in line with the volume-filling assumptions described in Section 5.4.3). As such, diffusion in physical space is given by

$$D^m(y,\rho(x,t),C(x,t)) = y\left(1 - \frac{\rho(x,t)+C(x,t)}{\kappa}\right),$$

which describes a higher diffusion rate for T cells with a higher phenotype.

Furthermore, we wish to model the phenotypic transitions of the T cells, or how exhausted the T cells become as they move, grow and interact with tumour cells. We have already assumed that the higher the phenotype of a T cell, the higher its probability of moving and growing, and this in turn will exhaust it. We also know that the tumour cells can be killed by the T cells, which we assume will further exhaust the T cells at a rate proportional to the number of interactions they have with the surrounding tumour cells. Therefore, we model the drift in phenotype space as

$$v^p(y,\rho(x,t),C(x,t)) = y(k_1 + k_2 C(x,t)),$$

where $k_1, k_2 \geq 0$ describe the exhaustion rate of the T cells as a result of movement and growth, and as a result of interactions with the tumour cells, respectively. We allow for some small randomness in the exhaustion levels of the T cells by including a diffusive term in phenotype space of the form

$$D^p(y, \rho(x,t), C(x,t)) = \varepsilon \ll 1.$$

It is well-known that tumour cells are also mobile and able to grow [324]. As such, we can derive an equation similar to Equation (5.35) for the dynamics of tumour cells which also includes terms describing random movement and proliferation terms, and are derived in the same manner as those in Equation (5.10). The resulting equation governing the evolution of the density of the tumour cells in space, $x \geq 0$, and time, $t \geq 0$, is therefore given by

$$\begin{aligned}\frac{\partial}{\partial t}C(x,t) = \frac{\partial}{\partial x}\Bigg(& D^C\big(\rho(x,t), C(x,t)\big)\frac{\partial}{\partial x}C(x,t) \\ & - C(x,t)\frac{\partial}{\partial x}D^C\big(\rho(x,t), C(x,t)\big)\Bigg) \\ & - \int_{y=0}^{y=1} \nu(y, T(x,y,t))C(x,t)\mathrm{d}y + g(\rho(x,t), C(x,t))C(x,t). \end{aligned} \tag{5.44}$$

We note here that the term $D^C(\rho(x,t), C(x,t))$ behaves similarly to $D^m(y, \rho(x,t), C(x,t))$ and that $g(\rho(x,t), C(x,t))$ behaves like the growth term in the previous application, but without the phenotype dependence. We have also assumed here that the tumour cells are homogeneous, which simplifies the equations and facilitates the analysis that follows, focusing on T cell phenotypic dependence/exhaustion.

Assuming that the motility and proliferation of tumour cells is restricted in regions of high cell density, we take

$$\begin{aligned} g(\rho(x,t), C(x,t)) &= r_C\left(1 - \frac{\rho(x,t) + C(x,t)}{\kappa}\right), \\ D^C(\rho(x,t), C(x,t)) &= 1 - \frac{\rho(x,t) + C(x,t)}{\kappa}, \end{aligned}$$

with $r_C \geq 0$ describing the growth rate of the tumour cells. Finally, T cells with a higher phenotype are less exhausted, and therefore kill tumour cells at a higher rate. As such, we write

$$\nu(y, T(x,y,t)) = \bar{\lambda} y T(x,y,t),$$

where $\bar{\lambda} \geq 0$ describes the rate of degradation of the tumour cells by the T cells, or the toxicity of the T cells on the tumour.

We simulate Equations (5.43) and (5.44) subject to the following initial conditions:

$$T_0(x,y) = \frac{\exp\left(-100(x^2+(y-1)^2)\right)}{\max\left(\int_{X_{\min}}^{X_{\max}} \exp\left(-100(x^2+(y-1)^2)\right)\,\mathrm{d}x\right)}, \tag{5.45}$$

$$C_0(x) = \begin{cases} 0 & \text{if } n_0(x,y) > 0.001, \\ 0.5 & \text{otherwise.} \end{cases} \tag{5.46}$$

Examining Figure 5.4 it is clear that two main invasion behaviours are exhibited, depending on the parameters of the system. The first, which can be observed in Figure 5.4(b) and Figure 5.4(c), shows T cells that attack the tumour cells and produce travelling wave type profiles that invade through the tumour cells into the domain, killing off the tumour as they do so.

In these simulations, altering the parameters of the system, namely the exhaustion rate of the T cells, leads to a range of different predicted behaviours. The initial population of T cells were naïve. As such, the phenotypic structure of the invading wave of T cells with a lower exhaustion rate consisted of a larger range of phenotypes of cells (see Figure 5.4(b)). Consequently, the invading wave retains cells with a high phenotype. Comparatively, by increasing the exhaustion rate of the T cells, a travelling wave of invasion can still be observed (see Figure 5.4(c), for example), but with a much more restricted set of phenotypes in the bulk of the wave. The very front of the wave contains T cells with higher phenotypes that are the least exhausted, but the total density of T cells invading into the tumour cells is decreased, and subsequently, so is the invasion speed.

In another scenario, whereby the growth of the tumour cells exceeds the degradation of the tumour cells by the T cells, results similar to those in Figure 5.4(d) are found, where the T cells exhaust extremely quickly as a result of a high number of interactions with tumour cells. This continued exposure quickly exhausts the T cells, which, in turn, increases death and creates available space for subsequent colonization by the tumour cells. In Figure 5.4(d), all of the T cells have died and we observe that the tumour cells (plotted as a dashed grey line) have grown to fill the entire domain, which they have already occupied in its entirety far ahead of the location of the initial population of T cells.

Complete exhaustion and death of the T cells can also be observed by decreasing the diffusion coefficient of the T cells, increasing (decreasing) the death (growth) rate of the T cells, decreasing the toxicity of the T cells or increasing the growth or motility rates of the tumour cells. In all of these cases, the tumour cells will eradicate the T cells and growth of the tumour to fill the available space will be observed.

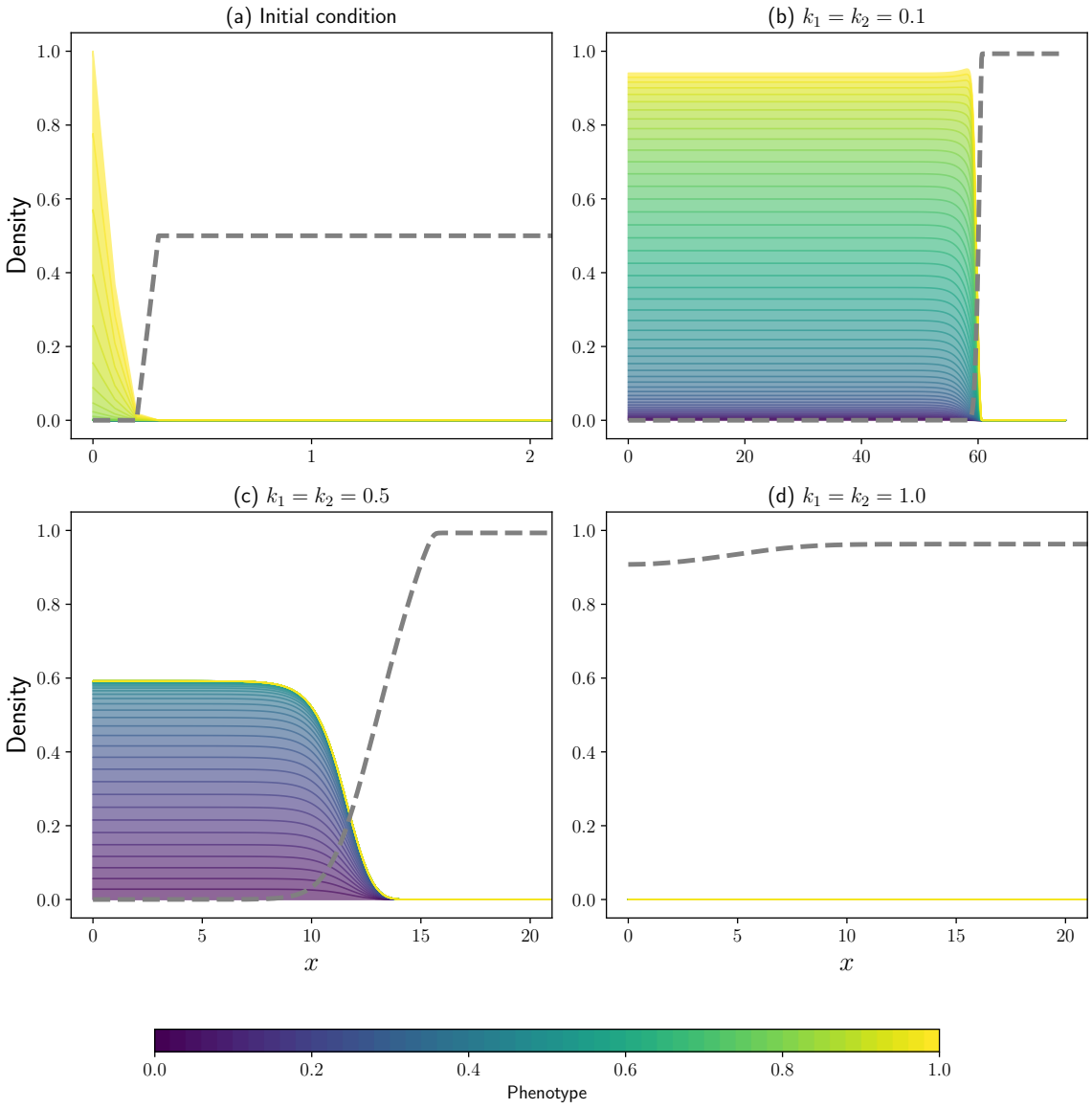

Figure 5.4: Evolution of the phenotypic structure of a population of T cells invading into a population of tumour cells, as described in Equations (5.43) and (5.44) subject to various exhaustion rates, with the corresponding tumour cell density shown as a dashed grey line. (a) The initial distribution of the cells. Results in (b), (c) and (d) show the spatial structure of the invading wave subject to exhaustion rates $k_1 = k_2 = 0.1, 0.5$ and 1.0, respectively. Solutions are plotted at $t = 50$, and simulations are carried out with $\varepsilon = 0.01$, $\gamma_0 = 1$, $\gamma_1 = 10$, $r_C = 0.1$, $\kappa = 1$, and $\bar{\lambda} = 10$.

This is an example of competitive exclusion [321]. The opposite effects on each of these parameters provides examples of when the T cells are able to either completely eradicate or prevent further growth of the tumour cells. These behaviours have been observed experimentally [296] and thus the modelling results shown in this work could provide useful insights into developing treatments for tumours in the future, once biologically realistic parameter sets are identified.

5.5 Discussion

In conclusion, numerous biological processes can be described in a simple way by considering phenotypically structured cell populations. Yet, general frameworks for modelling this, built from vastly adaptable underlying individual-based principles, are largely understudied [343]. In this chapter, we have demonstrated how to derive a general model for phenotypically structured cell migration from the underlying individual-based processes, that can be used to reproduce a variety of results in the literature, whilst modelling a continuum of phenotypes.

Whilst we have illustrated that the general modelling framework is easily adaptable and can be applied to a range of biological systems considering spatial invasion of structured populations, we highlight the need for adaptation of the underlying assumptions to the specific modelling question of interest. For instance, the continuum model derived in this work employed distinctly different boundary conditions in the physical and phenotypic domains. These differences arose from the varying assumptions about the processes governing movement in each domain. However, these conditions can be straightforwardly modified through analogous derivations. We therefore stress that this work represents an important first step towards generalising this modelling approach.

In changing the form of the phenotypic drift terms, care must also be taken to ensure that the limits taken in moving from the IBM to the continuum model are satisfied. For reaction–diffusion models with drift terms, the impact of these scalings and quantitative comparisons between the individual-based and resulting continuum models in such limits are well-studied in the literature [190, 195].

Overall, when simulation results from a particular biological scenario present vastly different phenotypically structured solutions, this general modelling framework could provide a useful tool for developing understanding of the mechanisms underlying heterogeneity during migration, such as the bet-hedging strategies observed as a result of environmental pressures in [7]. The framework presented in this work has

been derived from descriptions of interactions at an individual-based level and can be easily adapted to investigate a variety of biological applications. As a result, the versatility of this tool could help us understand the role of heterogeneity in a wide range of circumstances and its resulting insights could ultimately be used to help inform subsequent important treatment decisions [61].

Chapter 6

Discussion

The central aim of this thesis is to investigate how phenotypic heterogeneity—in particular, heterogeneous cell behaviour emerging from environmental cues, intrinsic stochasticity, and phenotypic switching—shapes the spatial structure of invading cell populations. Understanding the complicated interplay between these factors in determining the role and importance of phenotypic heterogeneity in a variety of contexts remains a cornerstone problem in mathematical biology, and has direct relevance to applications including wound healing, immunological invasion, and tumour growth. The work presented in this thesis contributes new analytical tools, original mathematical models, and novel mechanistic insights into how heterogeneous cell populations migrate and interact with their environment.

One key theme is the development of models in which population-level behaviour arises from phenotypic choice: cells may, for example, change motility, or modulate their interaction with the ECM, and these changes influence the emerging population dynamics in non-trivial ways. By combining individual-based descriptions, continuum analysis, numerical simulations and variational methods, the chapters in this thesis collectively demonstrate how differences at the microscopic or mesoscopic scale can be rigorously linked to macroscopic observations such as travelling wave speeds, wave profiles, and front stability.

In the following sections, an overview of the main contributions of each chapter is presented, highlighting how they build on one another, maintaining the same key drivers: continuum models should be robustly derived from underlying individual-based rules, and models should be compared to one another to establish the key differences between them and when mechanisms are superfluous. In Section 6.2, additional work beyond the scope of, but relevant to the thesis are briefly highlighted. Thereafter, Section 6.3 outlines several natural future research directions, including extensions of the modelling framework, additional biological mechanisms that could

be incorporated, and opportunities to validate the models with experimental data. This final chapter then concludes by reflecting on the broader implications of this work in future mathematical modelling of collective cell migration.

6.1 Summary of research in this thesis

Chapter 2 develops continuum limits for collective cell invasion into ECM by derivation from underlying individual-based rules. The resulting model and subsequent analysis reveals how nonlinear diffusion, substrate degradation, and environmental feedback collectively determine the minimal travelling wave speed. This work establishes that the ECM does not merely slow or redirect movement, it fundamentally alters the mechanism and ability for wave propagation. Depending on the relative timescales of ECM degradation and cell proliferation, the system either exhibits linearly selected or nonlinearly selected travelling waves, and by comparison to existing models in the literature, the work reveals where suitable model simplifications exist and when they can be used. Furthermore, this chapter also highlights how the initial ECM density can influence the selected wave speed, a feature that is absent from classical single species models, such as the Fisher–KPP model.

Chapter 3 introduces an optimal control formulation of wave speed selection, before extending the classical Benguria–Depassier variational method to systems of reaction–diffusion equations. The extension of the variational approach to a system of equations results in the establishment of new lower bounds for invasion speeds in two-species models, including some of the key systems involving ECM degradation that are studied in Chapter 2. The resulting framework clarifies when nonlinearities in either the diffusion or reaction terms are strong enough to drive nonlinear wave speed selection, and provides a new analytical structure to make progress in quantifying this, where previously only numerical exploration had been possible. In the setting describing collective cell invasion into ECM, the approach yielded new, tighter lower bounds that significantly outperform the predictions from linear theory. These new lower bounds match numerical simulations more closely than previous analytical results, particularly when ECM degradation rates are high.

Chapter 4 then turns to focus on understanding the role of cell phenotypes in mathematical models. The aim is to examine how switching between two states alters the collective invasion dynamics. Using on-lattice IBMs and continuum approximations, Chapter 4 shows that different switching mechanisms, including ECM-dependent, density-dependent, or space-dependent, produce qualitatively distinct

wave profiles and variations in the composition of the cell population at the wave front. From this it becomes clear that the phenotypic structure of the invading population may act as a fingerprint, indicating the underlying switching rules in the population. This observation suggests a potential inverse problem: in suitable experiments, it may be possible to infer the switching mechanisms in the population directly from the phenotypic distribution of cells throughout the wave front.

Finally, Chapter 5 extends the modelling framework in Chapter 4 to a phenotypically structured PDE system, where phenotype evolves continuously alongside space and time. This approach allows the description of more complex biological scenarios, such as go-or-grow behaviour or T cell exhaustion during tumour infiltration. Although simplified biological assumptions were adopted for tractability, the chapter demonstrates that a mechanistic, individual-based derivation can yield rich structured PDE systems that remain analytically and computationally tractable.

Collectively, these chapters show that phenotypic heterogeneity and environmental feedbacks can dramatically reshape invasion dynamics, and that combining variational analysis, continuum modelling, and individual-based simulations offers a powerful route to understanding these effects.

6.2 Additional work

During the course of developing the work in this thesis, several additional research projects and collaborations emerged. While not central to the main narrative of this thesis, these projects inform aspects of the modelling and analysis, and they demonstrate broader applicability of the ideas developed here. For completeness, the following sections briefly summarise a selection of these published related works.

6.2.1 Existence theory for models of cell invasion into ECM

After deriving the continuum model describing volume-filling migration of cells into ECM in Chapter 2, it became clear that no existence theory had been developed for models of this sort: coupling a nonlinear parabolic reaction–diffusion PDE with a monotonically increasing ODE solution in both the reaction and diffusion terms, despite existing work studying the solvability of models describing cancer invading into ECM [375] and inflammation during multiple sclerosis [104]. To address the before mentioned gap in the literature, we established an existence result for weak solutions under natural domain constraints $0 \leq u, m \leq 1$ and $0 \leq u + m \leq 1$.

To do this, we constructed a time-discrete regularised approximation by deriving entropy type and conditional regularity estimates that, when combined, yield compactness and H^1 control, despite the presence of the ODE. The approximation exploits the partial gradient flow structure of the governing equations in order to overcome the non-regularising nature of the ODE, and subsequently demonstrates the existence of weak solutions for the first model that this thesis develops [47].

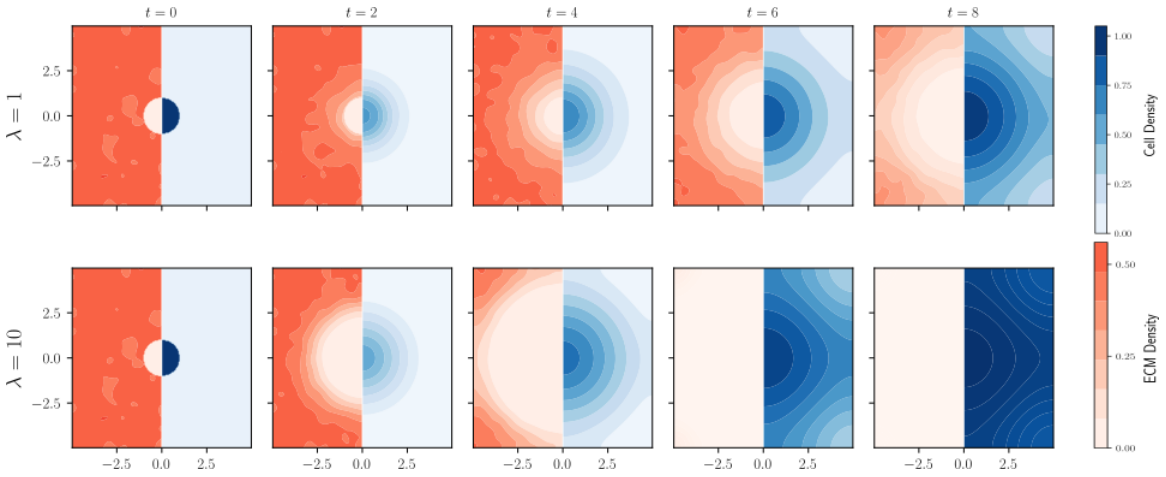

Figure 6.1: Plots of solutions to system (2.11) and (2.12) subject to zero flux boundary conditions and step initial conditions for the cells (blue, plotted only on the right half of the plots) and random initial data for the ECM, smoothed using a Gaussian filter with $\sigma = 5$ (red, plotted only on the left half of the plots). Results are symmetric, hence plotted in this way. Plots are shown subject to ECM degradation rates $\lambda = 1, 10$ at times $t = 0, 2, 4, 6, 8$.

Furthermore, in this work, travelling wave fronts were simulated and observed in both one and two dimensions, that were then used to explore the impact of heterogeneity in the surrounding ECM concentration. As in one dimension, the same travelling wave speeds were observed despite non-homogeneous ECM concentrations (see Figure 6.1 for example).

In summary, this work provides a rigorous existence theory for the class of cell–ECM models studied throughout this thesis and justifies the well-posedness of the continuum limits derived from the agent-based model, as well as developing a general analytical strategy for combining entropy methods with tailored time-discretisation to handle new PDE–ODE couplings.

This work was completed in collaboration with M. Schmidtchen and J.-F. Pietschmann and is published in the Journal of Differential Equations [80] as:
"Crossley RM, Pietschmann JF, Schmidtchen M. Existence of weak solutions for a

volume-filling model of cell invasion into extracellular matrix. Journal of Differential Equations. 428:721-46. 2025."

6.2.2 The impact of electrotaxis on cell-cell interactions

Besides physical constraints, chemical, magnetic and electric fields are all also known to impact cell migration properties. In another work, the role of external electric fields is investigated, in particular, considering their impact on cell–cell interactions. Using automated segmentation, tracking, and quantification of contact locations on the cell surface, this study analyses pairwise interactions of human corneal epithelial cells under control conditions and during electrotaxis [272].

The results show that while the frequency and duration of contacts are largely unchanged, the behavioural outcomes after a cell–cell collision, specifically the contact inhibition of locomotion and velocity alignment of the cells, are dramatically impacted by the electric field. Distinct zones on the cell membrane are mapped out, which under normal conditions produce characteristic repulsion or alignment responses. When an electric field is applied, these zones are spatially shifted or functionally altered. The outcomes demonstrate that electrotaxis affects not only single-cell polarity and directionality, which has previously been well-studied in the literature, but also fundamentally alters the outcomes of short-range physical interactions between cells, potentially leading to modified collective migration patterns. More broadly, the framework developed in this study provides a general method for quantifying and classifying cell–cell interaction rules in different environments, complementing the themes of this thesis on how external cues shape population-level invasion behaviours.

This work was completed in collaboration with S. F. Martina-Perez and is published in Biophysical Journal [78] as:
"Crossley RM, Martina-Perez SF. Electrotaxis disrupts patterns of cell-cell interactions of human corneal epithelial cells in vitro. Biophysical Journal. 124(8):1245-54. 2025"

6.2.3 Travelling waves in a minimal 'go-or-grow' model of cell invasion

As discussed in Chapter 4, the go-or-grow hypothesis suggests that certain cells undergo reversible phenotypic changes between a migratory and proliferative phenotypes. To make progress on the analysis of the travelling waves observed in these

scenarios, in this work we study the simplest possible model that captures this phenomenon. This model involves linear diffusion of the migrating population and logistic growth of the proliferative population, alongside a general density-dependent switching rate between the two states.

The study shows that, in the fast switching regime, the two population system formally reduces to a single reaction–diffusion equation with effective density-dependent diffusion and proliferation, thereby linking heterogeneous go-or-grow models with classic single-species nonlinear diffusion equations. Using this connection, we perform a detailed travelling wave analysis, deriving a dispersion relation that provides a lower bound on the invasion speed. A key result is that the minimum predicted travelling wave speed in the heterogeneous model is always bounded above by the minimum travelling wave speed of the Fisher–KPP model, and for biologically relevant switching functions (such as when switching into the proliferative state vanishes at zero density), the minimal wave speed cannot be attained, explaining discrepancies between previously observed analytical predictions and numerical simulations. Numerical results across a range of switching forms confirm these analytical insights and highlight regimes in which nonlinear switching causes deviations from linear theory. Overall, this work establishes a rigorous mathematical link between phenotype-structured invasion models and single-population PDEs, clarifying how switching rules shape macroscopic invasion speed, justifying the discrepancies observed between the linearly predicted and numerically observed wave speeds in Chapters 2 and 4, and motivating the work in Chapter 3.

This work was completed in collaboration with C. Falcó and R. E. Baker and is published in Applied Mathematics Letters [105] as:
"Falcó C, Crossley RM, Baker RE. Travelling waves in a minimal go-or-grow model of cell invasion. Applied Mathematics Letters. 158:109209. 2024."

6.2.4 Identifiability of phenotypic adaptation

In reality, the extent of phenotypic heterogeneity in a system may be identified directly from biological data. To investigate this, we examine the identifiability of phenotypic plasticity in cancer from low-cell-count proliferation assays. A stochastic IBM of cell adaptation is developed, whereby cells adapt along a continuous phenotypic axis between drug-sensitive and drug-resistant states. The dynamics of the phenotype change are governed by a deterministic drift term, which depends on the presence of treatment with a drug, and stochastic diffusion, representing spontaneous phenotype change. By exploiting the relationship between this stochastic model and

its associated Fokker–Planck PDE and chemical master equation, the study builds a likelihood framework that rigorously accounts for intrinsic noise. Using synthetic data under multiple experimental designs, we show that while proliferation rate and adaptation velocity are practically identifiable, the variance in the phenotype between cells is not identifiable from population-level observations, such as total cell counts or proliferation markers. Moreover, we demonstrate that population-level data alone cannot distinguish whether resistance arises from a binary switch between discrete states or from a continuum of phenotypes. Instead, both mechanisms generated indistinguishable cell count trajectories. Only finely grained data involving exact cell birth or death times (which is rarely available experimentally) contained sufficient information to recover the parameter governing random heterogeneity. These findings provide important guidance for the design of future experiments probing phenotypic plasticity, and show that simple homogeneous ODE models may be more appropriate for adequately describing population-level behaviour in many settings.

This work was completed in collaboration with A. P. Browning, C. Villa, P. K. Maini, A. L. Jenner, T. Cassidy and S. Hamis and is published in PLOS Computational Biology [42] as:
"Browning AP, Crossley RM, Villa C, Maini PK, Jenner AL, Cassidy T, Hamis S. Identifiability of phenotypic adaptation from low-cell-count experiments and a stochastic model. PLoS Computational Biology. 21(6):e1013202. 2025."

6.2.5 Optimal control of heterogeneous cell populations

In another piece of complementary work, a general optimal control framework for treating heterogeneous cell populations with combination drug therapies is developed and analysed. In this situation, the complexity arises as multiple drugs interact with one another synergistically and act on multiple phenotypically distinct cell types. This work focuses on a broad class of ODE models describing cell population growth and death and the actions of the drugs. Using Pontryagin's Minimum Principle, we derive explicit optimality conditions for these systems and show how nonlinear interactions modify classical regulatory solutions. The framework is then applied to three biologically grounded case studies: a two-population model of cervical cancer which is treated with cisplatin and paclitaxel; a neuroblastoma model featuring three cell types with complex differentiation pathways mediated by chemotherapy, signalling pathways and drug interactions; and a formulation for controlling proportions of cell types, rather than individual cell numbers, which is particularly relevant to engineered systems. In both biological cases, optimal dosing is shown to outperform constant

drug administration, particularly in low-toxicity regimes. We also reveal distinct regions of parameter space that generated qualitatively different optimal treatment strategies.

Overall, this work demonstrates how optimal control can aid in the design of time-varying, synergistic drug schedules, which are essential for biomedical treatments of systems containing multiple cell types. It also provides a flexible mathematical template for personalised treatment design in heterogeneous tissues for the future.

This work was completed in collaboration with S. F. Martina-Perez, S. W. S. Johnson, J. C. Kasemeier-Kulesa, P. M. Kulesa and R. E. Baker and is published in the Bulletin of Mathematical Biology [203] as:
"Martina-Perez SF, Johnson SW, Crossley RM, Kasemeier JC, Kulesa PM, Baker RE. Optimal control in combination therapy for heterogeneous cell populations with drug synergies. Bulletin of Mathematical Biology. 87:159. 2025"

6.3 Perspectives and future directions

During the development of this thesis, several open questions and avenues for further investigation, both mathematically and biologically, have been identified. While the models presented here capture a number of key features of collective cell migration including nonlinear diffusion, phenotypic switching, and cell–environment interactions, they also highlight mechanisms that remain unexplored or only partially understood. In what follows, several directions in which the present work could be extended, refined, or connected more closely to experimental systems are outlined. These perspectives are by no means exhaustive, but are chosen to represent the most promising opportunities for deepening our understanding of how heterogeneous cell populations migrate, interact with their surroundings, and give rise to the rich dynamics observed in biological tissues.

6.3.1 Model assumptions and biological extensions

Across Chapters 2, 4 and 5, the models developed all assume that cells experience purely diffusive movement. This movement is then considered to be modulated either by the presence of ECM or other cells around the cell, or the phenotype of the cell, without explicitly including any migratory support from the ECM fibres, other than in a volume-filling capacity. Haptotaxis, the movement of cells along a gradient in adhesive substrates, such as the ECM, is a well-studied biological mechanism. In future developments, haptotaxis could be included in the models to improve biological

realism, and, by explicitly modelling ECM fibres and their alignment, the migration of cells as a result of contact guidance can also be modelled.

In the IBM, several other assumptions could be relaxed. For example, processes such as proliferation and ECM degradation are currently restricted to impacting only the lattice site occupied by the cell. In reality, ECM degradation often also occurs when cells secrete matrix degrading enzymes that diffuse in the extracellular space, and may not be membrane-bound. Introducing non-local ECM degradation kernels could therefore offer a route to more realistic biological modelling and may likely result in a different continuum formulation and varied simulation outcomes, such as continued wave propagation even when $m_0 = 1$.

Employing a different modelling framework, such as a hybrid or force-based model, may allow for changes in cell proliferation, physical ECM cross-linking and remodelling, or mechanical interactions to be incorporated. Allowing ECM density to evolve through elastic deformation, or cross-linking [197] may reveal regimes where mechanical resistance dominates over biochemical cues. Further chemicals secreted by cells or present in the local environment could also be included as diffusible PDEs, in order to study the impact of chemotaxis on cells during their migration [289].

In the specialist population model developed in Chapter 4, it is often unclear how to compare between the models with different phenotypic switching functions. This suggests a further extension of this work to include energetic costs governing proliferation, movement, ECM degradation and phenotypic switching between states that allows for a more biologically conclusive investigation of which form of phenotypic switching function generates the highest rate of cell invasion.

6.3.2 Dimensionality and spatial structure

All of the models developed in Chapters 2, 4 and 5 were primarily analysed in one spatial dimension. Extending these results to two or three dimensions raises several new questions. In higher dimensions, invading fronts may display instabilities, that, in some scenarios, could lead to "fingering" or pattern formation, which has been observed in models of invasion into ECM [13, 55] or during tumour expansion when cells have heterogeneous motility [94, 189]. Analysing front stability in the ECM-driven invasion models derived here could reveal whether similar instabilities arise spontaneously or only under particular parameter regimes.

The phenotypically structured PDE of Chapter 5 also naturally accommodates additional phenotype dimensions, such as separating proliferation rate from motility, or introducing metabolic state, cell size, or receptor expression. Such extensions would

allow the modelling of richer experimental systems, including tumour spheroids where both spatial and phenotypic heterogeneity are essential features.

6.3.3 Travelling wave analysis

The travelling wave speeds observed in Chapters 2 and 4 reveal rich behaviours not present in classical Fisher–KPP type waves. The interplay between ECM degradation and cell movement suggests transitions between pulled, pushed, and semi-pushed waves, as described in recent ecological literature [32, 348]. This is likely due to the nonlinear cross-species dynamics that vary in strength for different parameter values. Quantifying these transitions formally would require studying how the ratio of observed wave speeds to linear predictions varies with parameters, and examining whether the wave profile develops characteristic signatures of pushed or semi-pushed behaviour.

In the model presented in Chapter 2, there is an opportunity to apply boundary layer theory and asymptotic analysis to arrive at an expression for the full travelling wave profile at long times when the rescaled ECM degradation rate $\lambda \to 0^+$. A further analytical challenge arises in determining the critical value λ_c in the before-mentioned models, where λ_c represents the critical value of the parameter that separates linear from nonlinear wave speed selection in models describing collective cell migration into ECM. Numerical simulations, together with the analysis in Chapter 3, strongly suggest that the transition depends sensitively on the initial ECM density. Future work could involve the use of basins of attraction, boundary layer theory or matched asymptotics near the leading edge to provide the tools to derive explicit formulae for the speed $c(\lambda, m_0)$ depending on the rescaled ECM degradation rate, λ, and the initial ECM density, m_0, and the critical value, λ_c. If possible, this knowledge could then further aid an investigation using perturbation methods into the shape of the wave front for intermediate values of λ.

The idea of a criterion for nonlinear wave speed selection could also be investigated further, in general. An analytical framework diagnosing parameter regimes whereby a PDE should be expected to travel at the speed predicted by linear analysis could be vital for future researchers validating their numerical models and provide genuinely valuable analytical insights. Furthermore, the optimal control approach to predicting the minimum travelling wave speed in taxis-based systems should be validated against numerical simulations and data, before further being expanded into systems of two or more species including other terms such as cross-dependent diffusion in both species.

6.3.4 Model validation

Tailoring the models developed in this thesis to specific biological systems presents a natural next step. Validation would involve estimating model parameters, comparing predicted wave speeds and profiles to experimental measurements, and determining which switching rules or ECM interactions most accurately match tissue-level observations. Histological data could be used to infer phenotypic distributions across the wave front, potentially allowing the inverse use of the results of Chapter 4: predicting switching mechanisms from observed population structure [308].

For the structured PDE model in Chapter 5, detailed biological data are needed to parameterise phenotype evolution, particularly in systems such as T cell exhaustion, where the biological meaning of "exhaustion level" remains debated [33]. The modelling framework offers a starting point, but substantial empirical work is required to identify appropriate switching, proliferation, and motility functions.

6.3.5 Learning models from data

Throughout this thesis, the continuum models studied were derived from their underlying individual-based equivalents. Whilst achievable, this process is not always straightforward and relies on sensible choices of limits, parameters and functions to ensure closure at the boundaries, amongst other technical requirements. Thankfully, over recent years, advancements in experimental biology have led to an increase in data, which is of higher resolution and quality than ever before. As a result, new techniques are being developed to learn the most appropriate mathematical models directly from the data [20, 309].

These techniques range from sparse regression [46, 215, 287], where the experimental observations are used to approximate derivatives in space and time, before being matched to the best fitting terms in a library of candidates to form a full PDE description of the data; to neural network based techniques including universal differential equations (UDEs) [88, 267, 266, 276], physics-informed neural networks (PINNs) [6, 23, 160, 255, 269, 270, 278, 323, 328, 370, 371], and, more recently, biologically-informed neural networks (BINNs) [102, 127, 125, 181, 182, 233, 232, 299]. In UDEs, a neural network component is trained within a differential equation, whereas in PINNs and BINNs, a chosen model structure is imposed as a soft constraint and multiple networks are used to learn the solution and the functions within the corresponding PDE. As such, all of the continuum models developed and studied during this thesis are amenable to these methods and could potentially be recovered directly from

data. This suggests the opportunity to learn functional forms of switching rates, diffusion coefficients, or phenotype evolution directly from complex spatio–temporal data, followed by the potential to resolve long-term modelling debates, such as those surrounding the go-or-grow hypothesis in cancer [71, 114, 347]. In BINNs, the variational and optimal control formulations of the wave speed could be integrated to provide structural constraints or priors for these learning approaches when travelling waves are exhibited, and therefore allow opportunities to bridge mechanistic modelling with modern inference techniques.

6.4 Closing remarks

This thesis explored the role of heterogeneity in reshaping and governing the collective movement of cells. By deriving new continuum limits, formulating a variational and optimal control framework for wave speed selection, and analysing both discrete and structured PDE models, this thesis has elucidated that seemingly subtle microscopic features can produce strikingly different macroscopic effects, and vice versa. Across the chapters, the recurring theme is that population-level behaviour relies heavily on the complex interplay between individual cells, their phenotypes and their interactions with the local environment. This thesis as a whole has demonstrated that mathematical models grounded in mechanistic principles provide a rigorous way to uncover this relationship and transparently compare models and methodologies from throughout the literature.

The models, tools and approaches developed in this thesis ultimately open pathways to richer biological modelling, deeper analytical understanding, and a more direct connection to experimental data across a variety of contexts. In doing so, the thesis lays conceptual and methodological foundations for future work on cell migration, phenotypic evolution, and the mathematics of invasion processes.

Appendix A

Appendix for Chapter 4

A.1 Spatially homogeneous steady states

In this section, we perform a steady state analysis for the system (4.11)-(4.13) subject to each of the phenotypic switching functions listed in Table 4.1. Since we are interested in travelling wave solutions, we seek spatially homogeneous steady states.

For all phenotypic switching mechanisms we consider, the spatially homogeneous steady states satisfy

$$u_2(1 - u_1 - u_2 - m) = 0, \tag{A.1}$$

$$mu_1 = 0. \tag{A.2}$$

Equation (A.2) implies that either $m = 0$, or $u_1 = 0$, and Equation (A.1) gives $u_2 = 0$ or $u_1 + u_2 + m = 1$.

A.1.1 Constant switching

For the case that $\gamma_{12}(u_1, u_2, m) = \gamma_{21}(u_1, u_2, m) = s$, the spatially homogeneous steady states of the system (4.11)-(4.13) also satisfy

$$-u_1 + u_2 = 0,$$

and the continuum of spatially homogeneous steady states is given by

$$\mathcal{A}_1 \coloneqq (u_1^*, u_2^*, m^*) = (0, 0, \bar{m}),$$
$$\mathcal{A}_2 \coloneqq (u_1^*, u_2^*, m^*) = \left(\bar{u_1}, \bar{u_2}, 0\right),$$

where $\bar{u_1}, \bar{u_2}, \bar{m} \in [0, 1]$, prescribed by initial conditions, and $\bar{u_1} = \bar{u_2} = 0.5$ (see Section A.2 for values of $\bar{u_1}$ and $\bar{u_2}$ when switching rates differ in either direction. The

steady state $\mathcal{A}_1$ describes the case with no cells present and only ECM. Alternatively, $\mathcal{A}_2$ describes a mixed population of cells, and no ECM, where the ratio between cells in phenotypic state 1 and 2 is determined by the phenotypic switching rates in either direction (see Figure A.1).

A.1.2 ECM-dependent switching

When we consider ECM-dependent switching, we find that the spatially homogeneous steady states must satisfy Equations (A.1) and (A.2) along with

$$-u_1(1-m) + u_2 m = 0.$$

By the same arguments as before, the resulting steady states are given by

$$\mathcal{B}_1 \coloneqq (u_1^*,\, u_2^*,\, m^*) = (0,\, 0,\, \bar{m}),$$
$$\mathcal{B}_2 \coloneqq (u_1^*,\, u_2^*,\, m^*) = (0,\, 1,\, 0).$$

Once again, the steady state described by no cells and only ECM (far ahead of the travelling wave) is given by $\mathcal{B}_1$, but now $\mathcal{B}_2$ describes a steady state consisting only of cells in phenotypic state 2.

A.1.3 Space-dependent switching

By considering phenotypic switching dependent on available space, the spatially homogeneous steady states satisfy

$$-su_1(1-u_1-u_2-m) + su_2(u_1+u_2+m) = 0,$$

such that the spatially homogeneous steady states are given by

$$\mathcal{C}_1 \coloneqq (u_1^*,\, u_2^*,\, m^*) = (0,\, 0,\, \bar{m}), \tag{A.3}$$
$$\mathcal{C}_2 \coloneqq (u_1^*,\, u_2^*,\, m^*) = (1,\, 0,\, 0). \tag{A.4}$$

$\mathcal{C}_1$ is the same steady state described for constant and ECM-dependent switching, where only ECM is present. $\mathcal{C}_2$ represents a steady state with only cells in phenotypic state 1, and no cells in phenotypic state 2 or ECM present.

A.1.4 Cell-dependent switching

Finally, the spatially homogeneous steady states under cell-dependent phenotypic switching satisfy Equations (A.1) and (A.2) and

$$-su_1(1-u_1-u_2)+su_2(u_1+u_2)=0,$$

to give the spatially homogeneous steady states, $\mathcal{C}_1$ and $\mathcal{C}_2$ (see Equations (A.3) and (A.4)), that are the same as those observed under space-dependent switching.

A.2 Distinct switching rates in either direction

In this section, we investigate the system (4.11)-(4.13) subject to the phenotypic switching functions listed in Table A.1 with different switching rates in either direction, such that cells in phenotypic state 1 switch to phenotypic state 2 at rate $s_{12} \in \mathbb{R}_+$ and, equivalently, cells in phenotypic state 2 switch to phenotypic state 1 at a rate given by $s_{21} \in \mathbb{R}_+$. Since there is limited evidence that distinctly different switching rates are biologically relevant [239], we only briefly discuss some of the main results for $s_{12} \neq s_{21}$ in this section to demonstrate the impact of these parameters on the model solutions.

Name	$\gamma_{12}(u_1,u_2,m)$	$\gamma_{21}(u_1,u_2,m)$
Constant switching	s_{12}	s_{21}
ECM-dependent switching	$s_{12}(1-m)$	$s_{21}m$
Space-dependent switching	$s_{12}(1-u_1-u_2-m)$	$s_{21}(u_1+u_2+m)$
Cell-dependent switching	$s_{12}(1-u_1-u_2)$	$s_{21}(u_1+u_2)$

Table A.1: Table listing the phenotypic switching functions.

Figure A.1 shows that, in the case of constant switching, the ratio between the phenotypic switching rates, s_{12} and s_{21}, directly determines the ratio between the volume fraction of cells in phenotypic states 1 and 2 in the bulk of the migrating cell population, as described in A.1.1. For s_{12}, $s_{21} \in [0,1]$ the relationship between the travelling wave speed and these parameters is symmetrical around $s_{12} = s_{21}$, which is the maximum speed observed numerically and predicted analytically in the fast phenotypic switching regime (see Section 4.4.3).

Furthermore, changes in the individual switching rates also impact the travelling wave profile and speed of migration. For example, when considering ECM-dependent

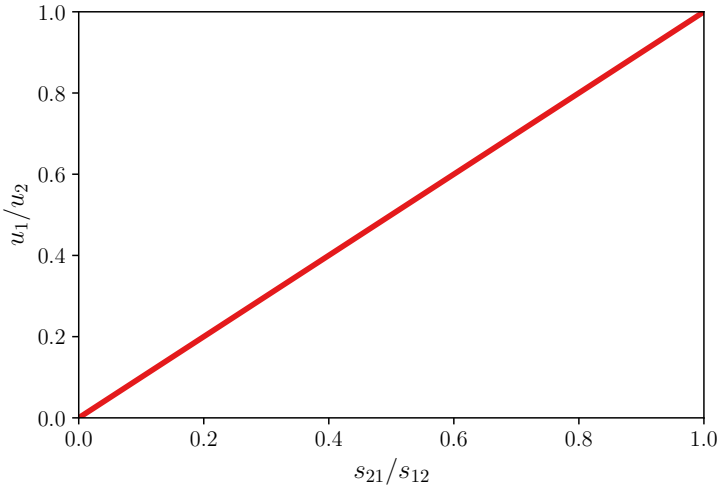

Figure A.1: Plot demonstrating the relationship between the ratio of the switching rates, s_{12} and s_{21}, and the ratio between the volume fraction of the two cell sub-populations behind the wave front, u_1/u_2 when simulating the system (4.11)-(4.13) subject to the initial conditions for the cells as in Equation (4.14), and for the ECM as in Equation (4.4), subject to constant phenotypic switching (see Table 4.1). This plot was produced by running simulations under a variety of switching rates and plotting the ratio between the resulting cell sub-population densities behind the wave front. The initial ECM volume fraction ahead of the cells is $m_0 = 0.5$, the ECM degradation rate is $\lambda = 1$, and the width of the region initially invaded by migrating cells is $\alpha = 1$ across all simulations. For more information regarding the numerical methods used see Section 4.3.1.

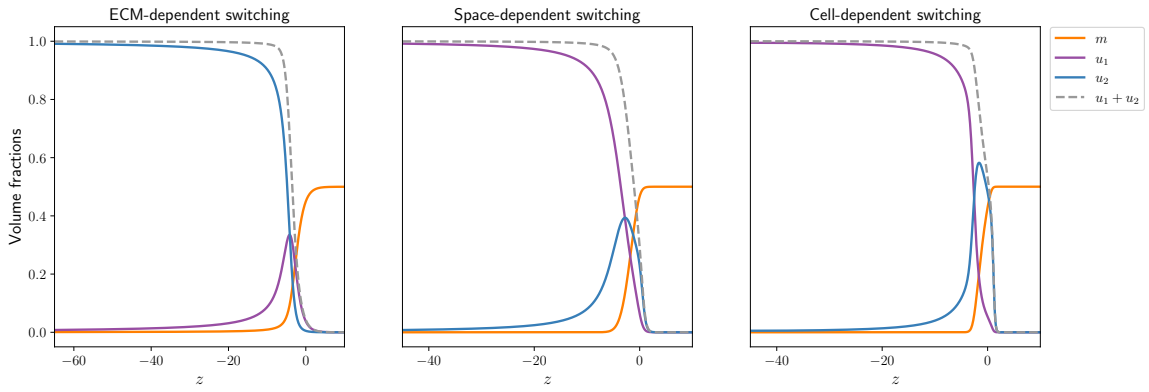

Figure A.2: Travelling wave profiles of the solutions of Equations (4.11)-(4.13) subject to the initial conditions for the cells as in Equation (4.14), and for the ECM as in Equation (4.4), plotted as a function of the travelling wave variable $z = x - ct$, where c is the numerically observed travelling wave speed. These solutions demonstrate that changing the switching rate in one direction leads to one sub-population dominating the migrating front. For ECM-dependent switching, the switching rate from phenotypic state 1 to 2 is $s_{12} = 1$ and the switching rate from phenotypic state 2 to 1 is $s_{21} = 10$. For space- and cell-dependent switching, the switching rate from phenotypic state 1 to 2 is $s_{12} = 10$ and the switching rate from phenotypic state 2 to 1 is $s_{21} = 1$. The initial ECM volume fraction ahead of the cells is $m_0 = 0.5$, the ECM degradation rate is $\lambda = 1$, the weighting of the specialists towards degrading ECM is $\theta_{S,D} = 0.5$ and the width of the region initially invaded by migrating cells is $\alpha = 1$ in all cases. For more information regarding the numerical methods used see Section 4.3.1.

switching, although increasing the switching rate from phenotypic state 2 to phenotypic state 1 decreases the travelling wave speed, it also changes the distribution of cells in the migrating front such that the front of the travelling wave is dominated by degrading cells in phenotypic state 1, ahead of a mixed region of cells in both states, and the bulk of proliferating cells remains in the rear (see Figure A.2).

A similar result can be observed in Figure A.2 for space-dependent switching and cell-dependent switching. In these cases, increasing the switching rate from phenotypic state 1 to phenotypic state 2 causes a leading population of cells in phenotypic state 2 at the front of the travelling wave. In all cases, the greater the difference between the switching rates, the larger and more concentrated this proportion of leader cells are.

Appendix B

Appendix for Chapter 5

B.1 Individual-based model functions

B.1.1 Phenotypic structuring during range expansion

As per the rules described in Section 5.2, we can write functions to describe the movement and growth of cells over time. In particular, noting that in this case we have homogeneous cells, with constant random movement in all directions, then we can write

$$\beta_{\pm}(j, N_{i\pm1}, e_{i\pm1}) = 1.$$

Furthermore, when considering Fisher–KPP-type invasion, we know that cells grow faster in areas with higher amounts of available space, which can be modelled as

$$\gamma_K(j, N_i) = 1 - \frac{N_i}{\kappa},$$

whereas, with the addition of the Allee effect, we instead have

$$\gamma_A(j, N_i) = \left(1 - \frac{N_i}{\kappa}\right)(N_i - p^*),$$

with $p^* \in (0, 1/2)$ and $\kappa > 0$ describing the maximum total number of cells that can fit in any single site. Implementing these functional forms during the coarse-graining process described in Section 5.3, absorbing constants in the continuum limit and rescaling as appropriate, the resulting continuum equation is given by Equation (5.37) with functions (5.38) and (5.39).

B.1.2 A go-or-grow model of cells invading into ECM

When describing cells moving into ECM, we know that volume filling constraints will affect the movement in physical space. In fact, as space decreases, cells have less space

in which to move. Furthermore, we implement a continuum of cell phenotypes in this case, such that cells in phenotypic state $j = Y_{\max}$ are the most proliferative, but least motile and degrading cells. As such, the individual-based functions describing movement in physical space can be written as

$$\beta_{\pm}(j, N_{i\pm1}, e_{i\pm1}) = (1-j)\left(1 - \frac{N_{i\pm1} + e_{i\pm1}}{\kappa}\right),$$

where $\kappa > 0$ is the total number of available sites for cells and ECM elements, known as the carrying capacity.

Cells are able to proliferate more rapidly when there is a larger amount of available space, and when they occupy a phenotypic state with higher values. As such, we have that

$$\gamma(j, N_i, e_i) = j\left(1 - \frac{N_i + e_i}{\kappa}\right).$$

Alternatively, it is cells in a lower phenotypic state, j, that degrade the surrounding ECM at a higher rate. The corresponding function to describe this is given by

$$\lambda(j, n_i^j) = (1-j)n_i^j.$$

In Section 5.4.3, we consider a number of different functions to describe movement in phenotypic space. The first phenotypic drift term we consider is cell-dependent drift, where cells transition into a phenotypic state with lower values at an increasing rate in regions with more cells present. The second phenotypic drift term we evaluate considers the role of the ECM in determining phenotypic transitions. In this case, cells transition to phenotypic states with lower values as the number of ECM elements in the same physical site increases. Finally we consider space-dependent phenotype transitions such that cells move into phenotypic sites with higher values at an increasing rate as the available space in the same physical site increases. These options are all described in Table B.1.

Implementing the individual-based functions described above during the coarse-graining process, absorbing parameters and rescaling gives the resulting continuum equations and functions as stated in Section 5.4.3 of the main text.

B.1.3 T cell exhaustion

In the case where we consider T cell exhaustion, we are simulating both T cells and tumour cells. Both the T cells and tumour cells can move and grow, whilst T cells have a further variable, exhaustion, attached to them, and they are able to kill off the tumour cells.

Phenotypic drift	$\mu_-(j, N_i, e_i)$	$\mu_+(j, N_i, e_i)$
Cell-dependent	$\frac{N_i}{\kappa}$	$1 - \frac{N_i}{\kappa}$
ECM-dependent	$\frac{e_i}{\kappa}$	$1 - \frac{e_i}{\kappa}$
Space-dependent	$\frac{N_i + e_i}{\kappa}$	$1 - \frac{N_i + e_i}{\kappa}$

Table B.1: Table listing the individual-based functions used during coarse-graining, that correspond to those continuum equivalents described in Table 5.1. The functions shown describe the probabilities of transitions up and down the phenotype space, $\mu_+(j, N_i, e_i)$ and $\mu_-(j, N_i, e_i)$, respectively.

The movement of both the tumour cells and T cells is subject to volume exclusion, but also depends on the exhaustion level of the cells for the T cells. As such, we can write that the individual-based functions describing the movement of the T cells are given by

$$\beta_\pm(j, N_{i\pm1}, e_{i\pm1}) = j\left(1 - \frac{N_{i\pm1} + e_{i\pm1}}{\kappa}\right),$$

where $\kappa > 0$ is the carrying capacity of each physical site, i. Alternatively, the movement of the tumour cells does not depend on the phenotype of the cells. For some rate $\bar{r}_\pm \geq 0$, describing the movement of the tumour cells, we can write

$$r_\pm(N_{i\pm1}, e_{i\pm1}) = \bar{r}_\pm\left(1 - \frac{N_{i\pm1} + e_{i\pm1}}{\kappa}\right),$$

which coarse grains to become the function given by $D^C(\rho(x,t), C(x,t))$ in the main text, using the same method as for $D^m(y, \rho(x,t), C(x,t))$.

T cells are able to divide and produce a daughter cell in the same phenotypic and physical site at a rate described by $\gamma_1 \geq 0$. However the rate of reproduction depends on available space and is also greater for less exhausted cells, in a phenotypic state with higher values. T cells can also die at a rate $\gamma_0 \geq 0$ which increases as they exhaust. As such, the function describing the net growth of the T cells at an individual-level is given by

$$\gamma(j, N_i, e_i) = \gamma_1 j\left(1 - \frac{N_i + e_i}{\kappa}\right) - \gamma_0(1 - j).$$

Concurrently, the probability of tumour cell proliferation increases with the available space. This probability behaves in a similar manner to $\gamma(j, N_i)$ in the previous application, but without phenotype dependence. As such, we write

$$b(N_i, e_i) = 1 - \frac{N_i + e_i}{\kappa}.$$

This term coarse grains to become the function given by $g(\rho(x,t), C(x,t))$ in the main text, using the same method as for $r(y, \rho(x,t), C(x,t))$.

T cells exhaust as a result of interactions with (being in the same site as) tumour cells. They also naturally exhaust. Both of these occur faster when a cell is less exhausted (in a higher phenotypic state). To implement this, we write that the probability of cells moving up or down in phenotype space can be written as

$$\mu_{\pm}(j, e_i) = \frac{1}{2}\left(1 \pm jk_1 \pm jk_2 e_i\right),$$

where $k_1, k_2 \geq 0$ describe the exhaustion rate of the T cells as a result of movement and growth, and as a result of interactions with the tumour cells, respectively.

Finally, tumour cells die (and are removed from their site) as a result of interactions with T cells in the same physical site at a rate $\tilde{\lambda} \geq 0$. The individual-based description of this is given by

$$\lambda(j, n_i^j) = \tilde{\lambda} j n_i^j.$$

Using these functions in a coarse-graining process similar to that in Section 5.3 and rescaling as appropriate, we find that the resulting system of equations is given by Equations (5.43) and (5.44), with continuum functions as described in Section 5.4.4.

Bibliography

[1] M. Acar, J. T. Mettetal, and A. Van Oudenaarden. Stochastic switching as a survival strategy in fluctuating environments. *Nature Genetics*, 40(4):471–475, 2008.

[2] E. Agliari, A. Fachechi, and C. Marullo. Nonlinear PDEs approach to statistical mechanics of dense associative memories. *Journal of Mathematical Physics*, 63(10), 2022.

[3] F. Ahmed and N. K. Haass. Microenvironment-driven dynamic heterogeneity and phenotypic plasticity as a mechanism of melanoma therapy resistance. *Frontiers in Oncology*, 8:173, 2018.

[4] B. Alberts. *Molecular Biology of the Cell*. Garland Science, 2017.

[5] R. Alert and X. Trepat. Physical models of collective cell migration. *Annual Review of Condensed Matter Physics*, 11:77–101, 2020.

[6] M. Almanstötter, R. Vetter, and D. Iber. PINNverse: Accurate parameter estimation in differential equations from noisy data with constrained physics-informed neural networks. *arXiv preprint arXiv:2504.05248*, 2025.

[7] L. Almeida, J. A. Denis, N. Ferrand, T. Lorenzi, A. Prunet, M. Sabbah, and C. Villa. Evolutionary dynamics of glucose-deprived cancer cells: insights from experimentally informed mathematical modelling. *Journal of the Royal Society Interface*, 21(210):20230587, 2024.

[8] R. Ananthakrishnan and A. Ehrlicher. The forces behind cell movement. *International Journal of Biological Sciences*, 3(5):303, 2007.

[9] A. Anderson and K. Rejniak. *Single-cell-based Models in Biology and Medicine*. Springer Science & Business Media, 2007.

[10] A. R. A. Anderson and M. A. J. Chaplain. Continuous and discrete mathematical models of tumor-induced angiogenesis. *Bulletin of Mathematical Biology*, 60(5):857–899, 1998.

[11] A. R. A. Anderson, M. A. J. Chaplain, E. L. Newman, R. J. C. Steele, and A. M. Thompson. Mathematical modelling of tumour invasion and metastasis. *Computational Mathematical Methods in Medicine*, 2:129–154, 2000.

[12] A. R. A. Anderson, K. A. Rejniak, P. Gerlee, and V. Quaranta. Microenvironment driven invasion: A multiscale multimodel investigation. *Journal of Mathematical Biology*, 58:579–624, 2009.

[13] A. R. A. Anderson, A. M. Weaver, P. T. Cummings, and V. Quaranta. Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment. *Cell*, 127(5):905–915, 2006.

[14] H. Anton, I. Bivens, and S. Davis. *Calculus: Multivariable Version*. Von Hoffmann Press, 2001.

[15] A. Ardaševa, A. R. A. Anderson, R. A. Gatenby, H. M. Byrne, P. K. Maini, and T. Lorenzi. Comparative study between discrete and continuum models for the evolution of competing phenotype-structured cell populations in dynamical environments. *Physical Review E*, 102(4):042404, 2020.

[16] O. Arino and M. Kimmel. Comparison of approaches to modeling of cell population dynamics. *SIAM Journal on Applied Mathematics*, 53(5):1480–1504, 1993.

[17] A. Arnold, L. Desvillettes, and C. Prévost. Existence of nontrivial steady states for populations structured with respect to space and a continuous trait. *Communications on Pure and Applied Analysis*, 11(1):83–96, 2012.

[18] D.G. Aronson and H.F. Weinberger. Nonlinear diffusion in population genetics, combustion, and nerve pulse propagation. In L.A. Peletier and J. Serrin, editors, *Partial Differential Equations and Related Topics*, volume 446 of *Lecture Notes in Mathematics*, pages 5–49. Springer, 1975.

[19] V. V. Artym, K. M. Yamada, and S. C. Mueller. ECM degradation assays for analyzing local cell invasion. *Extracellular Matrix Protocols: Second Edition*, pages 211–219, 2009.

[20] R. E. Baker, R. M. Crossley, C. Falcó, and S. F. Martina-Perez. Modelling collective cell migration in a data-rich age: challenges and opportunities for data-driven modelling. *arXiv preprint arXiv:2504.19974*, 2025.

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

[22] A. Barkovskaya, A. Buffone Jr, M. Žídek, and V. M. Weaver. Proteoglycans as mediators of cancer tissue mechanics. *Frontiers in Cell and Developmental Biology*, 8:569377, 2020.

[23] J.-H. Bastek, W. Sun, and D. M. Kochmann. Physics-informed diffusion models. *arXiv preprint arXiv:2403.14404*, 2024.

[24] H. J. E. Beaumont, J. Gallie, C. Kost, G. C. Ferguson, and P. B. Rainey. Experimental evolution of bet hedging. *Nature*, 462(7269):90–93, 2009.

[25] S. C. Bendall and G. P. Nolan. From single cells to deep phenotypes in cancer. *Nature Biotechnology*, 30(7):639–647, 2012.

[26] R. D. Benguria and M. C. Depassier. Validity of the linear speed selection mechanism for fronts of the nonlinear diffusion equation. *Physical Review Letters*, 73(16):2272, 1994.

[27] R. D. Benguria and M. C. Depassier. Speed of fronts of the reaction–diffusion equation. *Physical Review Letters*, 77(6):1171, 1996.

[28] R. D. Benguria and M. C. Depassier. Variational characterization of the speed of propagation of fronts for the nonlinear diffusion equation. *Communications in Mathematical Physics*, 175:221–227, 1996.

[29] R. D. Benguria, M. C. Depassier, and V. Méndez. Minimal speed of fronts of reaction–convection–diffusion equations. *Physical Review E*, 69(3):031106, 2004.

[30] R. D. Benguria, M. C. Depassier, and S. Rica. Variational estimates for the speed propagation of fronts in a nonlinear diffusive Fisher equation. *Chaos, Solitons & Fractals*, 164:112668, 2022.

[31] N. Berestycki, C. Mouhot, and G. Raoul. Existence of self-accelerating fronts for a non-local reaction–diffusion equations. *arXiv preprint arXiv:1512.00903*, 2015.

[32] G. Birzu, O. Hallatschek, and K. S. Korolev. Fluctuations uncover a distinct class of traveling waves. *Proceedings of the National Academy of Sciences*, 115(16):E3645–E3654, 2018.

[33] C. U. Blank, W. N. Haining, W. Held, P. G. Hogan, A. Kallies, E. Lugli, R. C. Lynn, M. Philip, A. Rao, N. P. Restifo, et al. Defining 'T cell exhaustion'. *Nature Reviews Immunology*, 19(11):665–674, 2019.

[34] A. B. Bloom and M. H. Zaman. Influence of the microenvironment on cell fate determination and migration. *Physiological Genomics*, 46(9):309–314, 2014.

[35] C. Borau, R. Chisholm, P. Richmond, and D. Walker. An agent-based model for cell microenvironment simulation using FLAMEGPU2. *Computers in Biology and Medicine*, 179:108831, 2024.

[36] E. Bouin and V. Calvez. Travelling waves for the cane toads equation with bounded traits. *Nonlinearity*, 27(9):2233, 2014.

[37] E. Bouin, V. Calvez, N. Meunier, S. Mirrahimi, B. Perthame, G. Raoul, and R. Voituriez. Invasion fronts with variable motility: phenotype selection, spatial sorting and wave acceleration. *Comptes Rendus: Mathématique*, 350(15-16):761–766, 2012.

[38] L. G. Bowden, M. J. Simpson, and R. E. Baker. Design and interpretation of cell trajectory assays. *Journal of the Royal Society Interface*, 10(88):20130630, 2013.

[39] S. L. K. Bowers, I. Banerjee, and T. A. Baudino. The extracellular matrix: At the center of it all. *Journal of Molecular and Cellular Cardiology*, 48(3):474–482, 2010.

[40] G. E. P. Box. Science and statistics. *Journal of the American Statistical Association*, 71(356):791–799, 1976.

[41] G. P. Brown, C. Shilton, B. L. Phillips, and R. Shine. Invasion, stress, and spinal arthritis in cane toads. *Proceedings of the National Academy of Sciences*, 104(45):17698–17700, 2007.

[42] A. P. Browning, R. M. Crossley, C. Villa, P. K. Maini, A. L. Jenner, T. Cassidy, and S. Hamis. Identifiability of heterogeneous phenotype adaptation from low-cell-count experiments and a stochastic model. *PLOS Computational Biology*, 21(6):e1013202, 2025.

[43] A. P. Browning, P. Haridas, and M. J. Simpson. A Bayesian sequential learning framework to parameterise continuum models of melanoma invasion into human skin. *Bulletin of Mathematical Biology*, 81(3):676–698, 2019.

[44] A. P. Browning, J. A. Sharp, T. Mapder, C. M. Baker, K. Burrage, and M. J. Simpson. Persistence as an optimal hedging strategy. *Biophysical Journal*, 120(1):133–142, 2021.

[45] M. Bruna and J. S. Chapman. Diffusion of multiple species with excluded-volume effects. *The Journal of Chemical Physics*, 137(20):204116, 2012.

[46] S. L. Brunton, J. L. Proctor, and J. N. Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15):3932–3937, 2016.

[47] M. Burger, M. Di Francesco, J.-F. Pietschmann, and B. Schlake. Nonlinear cross-diffusion with size exclusion. *SIAM Journal on Mathematical Analysis*, 42(6):2842–2871, 2010.

[48] H. Byrne and D. Drasdo. Individual-based and continuum models of growing cell populations: a comparison. *Journal of Mathematical Biology*, 58:657–687, 2009.

[49] G. A. Cabral-Pacheco, I. Garza-Velox, C. Castruita-De la Rosa, J. M. Ramirez-Acuña, B. A. Perez-Romero, J. F. Guerrero-Rodriguez, N. Martinez-Avila, and M. L. Martinez-Fierro. The roles of matrix metalloproteinases and their inhibitors in human diseases. *International Journal of Molecular Sciences*, 21(24):9739, 2020.

[50] N. R. Campbell, A. Rao, M. V. Hunter, M. K. Sznurkowska, L. Briker, M. Zhang, M. Baron, S. Heilmann, M. Deforet, C. Kenny, et al. Cooperation between melanoma cell states promotes metastasis through heterotypic cluster formation. *Developmental Cell*, 56(20):2808–2825, 2021.

[51] J. Canosa. On a nonlinear diffusion equation describing population growth. *IBM Journal of Research and Development*, 17(4):307–313, 1973.

[52] P. Carmeliet and R. K. Jain. Molecular mechanisms and clinical applications of angiogenesis. *Nature*, 473(7347):298–307, 2011.

[53] J. A. Carrillo, T. Lorenzi, and F. R. Macfarlane. Spatial segregation across travelling fronts in individual-based and continuum models for the growth of heterogeneous cell populations. *Bulletin of Mathematical Biology*, 87(6):77, 2025.

[54] J. A. Carrillo, M. Parisot, and Z. Szymańska. Mathematical modelling of collagen fibres rearrangement during the tendon healing process. *Kinetic and Related Models*, 14(2):283–301, 2021.

[55] P. Carter, A. Doelman, P. van Heijster, D. Levy, P. Maini, E. Okey, and P. Yeung. Deformations of acid-mediated invasive tumors in a model with allee effect. *Journal of Mathematical Biology*, 90(6):55, 2025.

[56] G. L. Celora, H. M. Byrne, and P. G. Kevrekidis. Spatio-temporal modelling of phenotypic heterogeneity in tumour tissues and its impact on radiotherapy treatment. *Journal of Theoretical Biology*, 556:111248, 2023.

[57] G. L. Celora, H. M. Byrne, and C. E. Zois. Phenotypic variation modulates the growth dynamics and response to radiotherapy of solid tumours under normoxia and hypoxia. *Journal of Theoretical Biology*, 527:110792, 2021.

[58] M. A. J. Chaplain, M. Lachowicz, Z. Szymanska, and D. Wrzosek. Mathematical modelling of cancer invasion: The importance of cell-cell adhesion and cell-matrix adhesion. *Mathematical Modelling Methods in Applied Sciences*, 21(4):719–743, 2011.

[59] M. A. J. Chaplain and G. Lolas. Mathematical modelling of cancer cell invasion of tissue: The role of the urokinase plasminogen activation system. *Mathematical Models and Methods in Applied Sciences*, 15(11):1685–1734, 2005.

[60] M. A. J. Chaplain and G. Lolas. Mathematical modelling of cancer invasion of tissue: dynamic heterogeneity. *Networks & Heterogeneous Media*, 1(3):399–439, 2006.

[61] M. A. J. Chaplain, T. Lorenzi, A. Lorz, and C. Venkataraman. Mathematical modelling of phenotypic selection within solid tumours. *Numerical Mathematics and Advanced Applications ENUMATH 2017 Editorial Board*, page 237, 2019.

[62] O. Chaudhuri, J. Cooper-White, P. A. Janmey, D. J. Mooney, and V. B. Shenoy. Effects of extracellular matrix viscoelasticity on cellular behaviour. *Nature*, 584:535–546, 2020.

[63] A. Chauviere, L. Preziosi, and H. Byrne. A model of cell migration within the extracellular matrix based on a phenotypic switching mechanism. *Mathematical Medicine and Biology: A Journal of the IMA*, 27(3):255–281, 2010.

[64] K. J. Cheung and S. Horne-Badovinac. Collective migration modes in development, tissue repair and cancer. *Nature Reviews Molecular Cell Biology*, 26(10):741, 2025.

[65] D. Z. Childs, C. J. E. Metcalf, and M. Rees. Evolutionary bet-hedging in the real world: empirical evidence and challenges revealed by plants. *Proceedings of the Royal Society B: Biological Sciences*, 277(1697):3055–3064, 2010.

[66] R. H. Chisholm, T. Lorenzi, L. Desvillettes, and B. D. Hughes. Evolutionary dynamics of phenotype-structured populations: from individual-level mechanisms to population-level consequences. *Journal of Applied Mathematics and Physics*, 67(4):100, 2016.

[67] A. Chow, K. Perica, C. A. Klebanoff, and J. D. Wolchok. Clinical implications of T cell exhaustion for cancer immunotherapy. *Nature Reviews Clinical Oncology*, 19(12):775–790, 2022.

[68] D. Cohen. Optimizing reproduction in a randomly varying environment. *Journal of Theoretical Biology*, 12(1):119–129, 1966.

[69] C. Colson, F. Sánchez-Garduño, H. M. Byrne, P. K. Maini, and T. Lorenzi. Travelling-wave analysis of a model of tumour invasion with degenerate, cross-dependent diffusion. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 477(2256):20210593, 2021.

[70] M. Conte and C. Surulescu. Mathematical modeling of glioma invasion: acid- and vasculature mediated go-or-grow dichotomy and the influence of tissue anisotropy. *Applied Mathematics and Computation*, 407:126305, 2021.

[71] A. Corcoran and R. F. Del Maestro. Testing the "go or grow" hypothesis in human medulloblastoma cell lines in two and three dimensions. *Neurosurgery*, 53(1):174–185, 2003.

[72] S. J. Cornell, Y. F. Suprunenko, D. Finkelshtein, P. Somervuo, and O. Ovaskainen. A unified framework for analysis of individual-based models in ecology and beyond. *Nature Communications*, 10(1):4716, 2019.

[73] B. D. Cox, M. Natarajan, M. R. Stettner, and C. L. Gladson. New concepts regarding focal adhesion kinase promotion of cell migration and proliferation. *Journal of Cellular Biochemistry*, 99(1):35–52, 2006.

[74] T. R. Cox and J. T. Erler. Remodeling and homeostasis of the extracellular matrix: Implications for fibrotic diseases and cancer. *Disease Models & Mechanisms*, 4(2):165–178, 2011.

[75] R. M. Crossley, S. Johnson, E. Tsingos, Z. Bell, M. Berardi, M. Botticelli, Q. J. S. Braat, J. Metzcar, M. Ruscone, Y. Yin, and R. Shuttleworth. Modeling the extracellular matrix in cell migration and morphogenesis: a guide for the curious biologist. *Frontiers in Cell and Developmental Biology*, 12:1354132, 2024.

[76] R. M. Crossley, P. K. Maini, and R. E. Baker. Modelling the impact of phenotypic heterogeneity on cell migration: A continuum framework derived from individual-based principles. *Bulletin of Mathematical Biology*, 87(9):123, 2025.

[77] R. M. Crossley, P. K. Maini, T. Lorenzi, and R. E. Baker. Traveling waves in a coarse-grained model of volume-filling cell invasion: Simulations and comparisons. *Studies in Applied Mathematics*, 151(4):1471–1497, 2023.

[78] R. M. Crossley and S. F. Martina-Perez. Electrotaxis disrupts patterns of cell-cell interactions of human corneal epithelial cells in vitro. *Biophysical Journal*, 124(8):1245–1254, 2025.

[79] R. M. Crossley, K. J. Painter, T. Lorenzi, P. K. Maini, and R. E. Baker. Phenotypic switching mechanisms determine the structure of cell migration into extracellular matrix under the 'go-or-grow' hypothesis. *Mathematical Biosciences*, page 109240, 2024.

[80] R. M. Crossley, J.-F. Pietschmann, and M. Schmidtchen. Existence of weak solutions for a volume-filling model of cell invasion into extracellular matrix. *Journal of Differential Equations*, 428:721–746, 2025.

[81] B. D. Cumming, D. L. S. McElwain, and Z. Upton. A mathematical model of wound healing and subsequent scarring. *Journal of the Royal Society Interface*, 7(42):19–34, 2010.

[82] L. Curtin, A. Hawkins-Daarud, K. G. Van Der Zee, K. R. Swanson, and M. R. Owen. Speed switch in glioblastoma growth rate due to enhanced hypoxia-induced migration. *Bulletin of Mathematical Biology*, 82(3):43, 2020.

[83] P. D. Dale, J. A. Sherratt, and P. K. Maini. The speed of corneal epithelial wound healing. *Applied Mathematics Letters*, 7(2):11–14, 1994.

[84] J. C. Dallon and J. A. Sherratt. A mathematical model for fibroblast and collagen orientation. *Bulletin of Mathematical Biology*, 60(1):101–129, 1998.

[85] J. C. Dallon, J. A. Sherratt, and P. K. Maini. Mathematical modelling of extracellular matrix dynamics using discrete cells: Fiber orientation and tissue regeneration. *Journal of Theoretical Biology*, 199(4):449–471, 1999.

[86] M. P. Dalwadi, I. M. Griffiths, and M. Bruna. Understanding how porosity gradients can make a better filter using homogenization theory. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 471(2182):20150464, 2015.

[87] A. De Pablo and A. Sánchez. Travelling wave behaviour for a porous-Fisher equation. *European Journal of Applied Mathematics*, 9(3):285–304, 1998.

[88] M. de Rooij, B. Erdős, N. A. W. van Riel, and S. D. O'Donovan. Physiology-informed regularisation enables training of universal differential equation systems for biological applications. *PLOS Computational Biology*, 21(1):e1012198, 2025.

[89] M. C. Deregibus, V. Cantaluppi, R. Calogero, M. Lo Iacono, C. Tetta, L. Biancone, S. Bruno, B. Bussolati, and G. Camussi. Endothelial progenitor cell–derived microvesicles activate an angiogenic program in endothelial cells by a horizontal transfer of mRNA. *Blood, The Journal of the American Society of Hematology*, 110(7):2440–2448, 2007.

[90] A. Deutsch, J. M. Nava-Sedeño, S. Syga, and H. Hatzikirou. BIO-LGCA: A ceelular automaton modelling class for analysing collective cell migration. *PLOS Computational Biology*, 17(6), 2021.

[91] T. J. DeWitt, A. Sih, and D. S. Wilson. Costs and limits of phenotypic plasticity. *Trends in Ecology & Evolution*, 13(2):77–81, 1998.

[92] R. B. Dickinson and R. T. Tranquillo. A stochastic model for adhesion-mediated cell random motility and haptotaxis. *Journal of Mathematical Biology*, 31(6):563–600, 1993.

[93] R. A. Dolbeer. Migration patterns for age and sex classes of blackbirds and starlings. *Journal of Field Ornithology*, 53(1):28–46, 1982.

[94] D. Drasdo and S. Hoehme. Modeling the impact of granular embedding media, and pulling versus pushing cells on growing cell clones. *New Journal of Physics*, 14(5):055025, 2012.

[95] O. Du Roure, A. Saez, A. Buguin, R. H. Austin, P. Chavrier, P. Siberzan, and B. Ladoux. Force mapping in epithelial cell migration. *Proceedings of the National Academy of Sciences*, 102(7):2390–2395, 2005.

[96] G. Dudziuk, M. Lachowicz, H. Leszczyński, and Z. Szymańska. A simple model of collagen remodeling. *Discrete and Continuous Dynamical Systems - B*, 24(5):2205–2217, 2019.

[97] G. A. Dunn and J. P. Heath. A new hypothesis of contact guidance in tissue cells. *Experimental Cell Research*, 101(1):1–14, 1976.

[98] M. H. Duong. Asymptotic equivalence of the discrete variational functional and a rate-large-deviation-like functional in the wasserstein gradient flow of the porous medium equation. *Asymptotic Analysis*, 92(1-2):85–106, 2015.

[99] M. Egeblad and Z. Werb. New functions for the matrix metalloproteinases in cancer progression. *Nature Reviews Cancer*, 2(3):161–174, 2002.

[100] M. El-Hachem, S. W. McCue, and M. J. Simpson. Travelling wave analysis of cellular invasion into surrounding tissues. *Physica D: Nonlinear Phenomena*, 428:133026, 2021.

[101] R. E. Ellis, J. Yuan, H. R. Horvitz, et al. Mechanisms and functions of cell death. *Annual Review of Cell Biology*, 7(1):663–698, 1991.

[102] H. A. Elmarakeby, J. Hwang, R. Arafeh, J. Crowdis, S. Gang, D. Liu, S. H. Al Dubayan, K. Salari, S. Kregel, C. Richter, et al. Biologically informed deep neural network for prostate cancer discovery. *Nature*, 598(7880):348–352, 2021.

[103] L. Excoffier, M. Foll, and R. J. Petit. Genetic consequences of range expansions. *Annual Review of Ecology, Evolution, and Systematics*, 40(1):481–501, 2009.

[104] S. Fagioli, E. Radici, and L. Romagnoli. On a chemotaxis model with nonlinear diffusion modelling multiple sclerosis. *Partial Differential Equations and Applications*, 6(1):3, 2025.

[105] C. Falcó, R. M. Crossley, and R. E. Baker. Travelling waves in a minimal go-or-grow model of cell invasion. *Applied Mathematics Letters*, 158:109209, 2024.

[106] P. Farrell, A. Hegarty, J. M. Miller, E. O'Riordan, and G. I. Shishkin. *Robust computational techniques for boundary layers*. Chapman and Hall/CRC, 1 edition, 2000.

[107] J. Ferrer, C. Prats, and D. López. Individual-based modelling: an essential tool for microbiology. *Journal of Biological Physics*, 34(1):19–37, 2008.

[108] E. C. Filipe, J. L. Chitty, and T. R. Cox. Charting the unexplored extracellular matrix in cancer. *International Journal of Experimental Pathology*, 99(2):58–76, 2018.

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

[110] I. C. Fortunato and R. Sunyer. The forces behind directed cell migration. *Biophysica*, 2(4):548–563, 2022.

[111] C. Frantz, K. M. Stewart, and V. M. Weaver. The extracellular matrix at a glance. *Journal of Cell Science*, 123(24):4195–4200, 2010.

[112] P. Friedl and E.-B. Bröcker. The biology of cell locomotion within three-dimensional extracellular matrix. *Cellular and Molecular Life Sciences*, 57:41–64, 2000.

[113] P. Friedl and D. Gilmour. Collective cell migration in morphogenesis, regeneration and cancer. *Nature Reviews Molecular Cell Biology*, 10(7):445–457, 2009.

[114] T. Garay, E. Juhász, E. Molnár, M. Eisenbauer, A. Czirók, B. Dekan, V. László, M. A. Hoda, B. Döme, J. Tímár, W. Klepetko, W. Berger, and B. Hegedűs. Cell migration or cytokinesis and proliferation? – revisiting the "go or grow" hypothesis in cancer cells in vitro. *Experimental Cell Research*, 319(20):3094–3103, 2013.

[115] R. A. Gatenby. Population ecology issues in tumor growth. *Cancer Research*, 51(10):2542–2547, 1991.

[116] R. A. Gatenby and E. T. Gawlinski. A reaction–diffusion model of cancer invasion. *Cancer Research*, 56(24):5745–5753, 1996.

[117] R. A. Gatenby, E. T. Gawlinski, A. F. Gmitro, B. Kaylor, and R. J. Gillies. Acid-mediated tumor invasion: A multidisciplinary study. *Cancer Research*, 66(10):5216–5223, 2006.

[118] A. Gerisch and M. A. J. Chaplain. Mathematical modelling of cancer cell invasion of tissue: Local and non-local models and the effect of adhesion. *Journal of Theoretical Biology*, 250(4):684–704, 2008.

[119] A. Gerisch and M. A. J. Chaplain. Mathematical modelling of cancer cell invasion of tissue: local and non-local models and the effect of adhesion. *Journal of Theoretical Biology*, 250(4):684–704, 2008.

[120] P. Gerlee and S. Nelander. The impact of phenotypic switching on glioblastoma growth and invasion. *PLOS Computational Biology*, 8(6):e1002556, 2012.

[121] P. Gerlee and S. Nelander. Travelling wave analysis of a mathematical model of glioblastoma growth. *Mathematical Biosciences*, 276:75–81, 2016.

[122] A. Giese, L. Kluwe, B. Laube, H. Meissner, M. E. Berens, and M. Westphal. Migration of human glioma cells on myelin. *Neurosurgery*, 38(4):755–764, 1996.

[123] A. Giese, M. A. Loo, N. Tran, D. Haskett, S. W. Coons, and M. E. Berens. Dichotomy of astrocytoma migration and proliferation. *International Journal of Cancer*, 67(2):275–282, 1996.

[124] R. Giniūnaitė, R. E. Baker, P. M. Kulesa, and P. K. Maini. Modelling collective cell migration: Neural crest as a model paradigm. *Journal of Mathematical Biology*, 80:481–504, 2020.

[125] T. Goo, C. Lee, S. Shin, H. Kim, and T. Park. Uncertainty quantification and statistical inference for biologically informed neural networks. In *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 4423–4428. IEEE, 2024.

[126] C. A. Gravenmier, M. Siddique, and R. A. Gatenby. Adaptation to stochastic temporal variations in intratumoral blood flow: the Warburg effect as a bet hedging strategy. *Bulletin of Mathematical Biology*, 80(5):954–970, 2018.

[127] C. S. Greene and J. C. Costello. Biologically informed neural networks predict drug responses. *Cancer Cell*, 38(5):613–615, 2020.

[128] M. G. Guenza. Everything you want to know about coarse-graining and never dared to ask: Macromolecules as a key example. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 15(2):e70022, 2025.

[129] Y. Guo, M. R. K. Mofrad, and A. B. Tepole. On modeling the multiscale mechanobiology of soft tissues: Challenges and progress. *Biophysical Review*, 3(3), 2022.

[130] P. B. Gupta, I. Pastushenko, A. Skibinski, C. Blanpain, and C. Kuperwasser. Phenotypic plasticity: driver of cancer initiation, progression, and therapy resistance. *Cell Stem Cell*, 24(1):65–78, 2019.

[131] M. E. Gurtin and R. C. MacCamy. On the diffusion of biological populations. *Mathematical Biosciences*, 33(1-2):35–49, 1977.

[132] N. K. Haass, K. A. Beaumont, D. S. Hill, A. Anfosso, P. Mrass, M. A. Munoz, I. Kinjyo, and W. Weninger. Real-time cell cycle imaging during melanoma growth, invasion, and drug response. *Pigment Cell & Melanoma Research*, 27(5):764–776, 2014.

[133] K.-P. Hadeler and F. Rothe. Travelling fronts in nonlinear diffusion equations. *Journal of Mathematical Biology*, 2(3):251–263, 1975.

[134] D. Hanahan and R. A. Weinberg. The hallmarks of cancer. *Cell*, 100(1):57–70, 2000.

[135] D. Hanahan and R. A. Weinberg. Hallmarks of cancer: the next generation. *Cell*, 144(5):646–674, 2011.

[136] A.M. Handorf, Y. Zhou, M.A. Halanski, and W. Li. Tissue stiffness dictates development, homeostasis, and disease progression. *Organogenesis*, 11(1), 2015.

[137] S. Hanna and M. El-Sibai. Signaling networks of Rho GTPases in cell motility. *Cellular Signalling*, 25(10):1955–1961, 2013.

[138] D. Harjanto and M. H. Zaman. Modeling extracellular matrix reorganization in 3D environments. *PLOS One*, 8(1):e52509, 2013.

[139] H. Hatzikirou, D. Basanta, M. Simon, K. Schaller, and A. Deutsch. 'Go or grow': the key to the emergence of invasion in tumour progression? *Mathematical Medicine and Biology: A Journal of the IMA*, 29(1):49–65, 2012.

[140] X. He and Y. Jiang. Substrate curvature regulates cell migration. *Physical Biology*, 14(3):P035006, 2017.

[141] X. He, B. Lee, and Y. Jiang. Extracellular matrix in cancer progression and therapy. *Medical Review*, 2(2):125–139, 2022.

[142] X. He, Y. Yang, Y. Han, C. Cao, Z. Zhang, L. Xi, C. Xiao, H. Guo, L. Wang, L. Han, Z. Qu, N. Liu, S. Han, and F. Xu. Extracellular matrix physical properties govern the diffusion of nanoparticles in tumor microenvironment. *Proceedings of the National Academy of Sciences*, 120(1):e2209260120, 2022.

[143] A. P. Hendry. Key questions on the role of phenotypic plasticity in eco-evolutionary dynamics. *Journal of Heredity*, 107(1):25–41, 2016.

[144] T. Hillen and K. Painter. Global existence for a parabolic chemotaxis model with prevention of overcrowding. *Advances in Applied Mathematics*, 26(4):280–301, 2001.

[145] T. Hillen and K. J. Painter. A user's guide to PDE models for chemotaxis. *Journal of Mathematical Biology*, 58(1-2):183–217, 2009.

[146] K. R. Hopper. Risk-spreading and bet-hedging in insect population biology. *Annual Review of Entomology*, 44(1):535–560, 1999.

[147] Y. Hosono. Traveling wave solutions for some density dependent diffusion equations. *Japan Journal of Applied Mathematics*, 3(1):163–196, 1986.

[148] R. S. Hotchkiss, A. Strasser, J. E. McDunn, and P. E. Swanson. Cell death. *New England Journal of Medicine*, 361(16):1570–1583, 2009.

[149] A. Ibrahim-Hashim, M. Robertson-Tessi, P. M. Enriquez-Navas, M. Damaghi, Y. Balagurunathan, J. W. Wojtkowiak, S. Russell, K. Yoonseok, M. C. Lloyd, M. M. Bui, et al. Defining cancer subpopulations by adaptive strategies rather than molecular properties provides novel insights into intratumoral evolution. *Cancer Research*, 77(9):2242–2254, 2017.

[150] O. Ilina, P. G. Gritsenko, S. Syga, J. Lippoldt, C. A. M. La Porta, O. Chepizhko, S. Grosser, M. Vullings, G. Bakker, J. Starruß, P. Bult, S. Zapperi, J. A. Käs, A. Deutsch, and P. Friedl. Cell-cell adhesion and 3D matrix confinement determine jamming transitions in breast cancer invasion. *Nature Cell Biology*, 22:1103–1115, 2020.

[151] S. Itani, J. Gray, and C. J. Tomlin. An ODE model for the HER2/3-AKT signaling pathway in cancers that overexpress HER2. In *Proceedings of the 2010 American Control Conference*, pages 1235–1241. IEEE, 2010.

[152] Y. Itoh. Membrane-type matrix metalloproteinases: Their functions and regulations. *Matrix Biology*, 44:207–223, 2015.

[153] A. Jabłońska-Trypuć, M. Matejczyk, and S. Rosochacki. Matrix metalloproteinases (MMPs), the main extracellular matrix (ECM) enzymes in collagen degradation, as a target for anticancer drugs. *Journal of Enzyme Inhibition and Medicinal Chemistry*, 31(sup1):177–183, 2016.

[154] Y. Jiang, Y. Li, and B. Zhu. T-cell exhaustion in the tumor microenvironment. *Cell Death & Disease*, 6(6):e1792–e1792, 2015.

[155] J. Jiménez-Sánchez, J. J. Bosque, G. A. Jiménez Londoño, D. Molina-García, Á. Martínez, J. Pérez-Beteta, C. Ortega-Sabater, A. F. Honguero Martínez, A. M. García Vicente, G. F. Calvo, et al. Evolutionary dynamics at the tumor edge reveal metabolic imaging biomarkers. *Proceedings of the National Academy of Sciences*, 118(6):e2018110118, 2021.

[156] M. K. Jolly, M. Boareto, B. Huang, D. Jia, M. Lu, E. Ben-Jacob, J. N. Onuchic, and H. Levine. Implications of the hybrid epithelial/mesenchymal phenotype in metastasis. *Frontiers in Oncology*, 5:155, 2015.

[157] M. K. Jolly, K. E. Ware, S. Gilja, J. A. Somarelli, and H. Levine. EMT and MET: necessary or permissive for metastasis? *Molecular Oncology*, 11(7):755–769, 2017.

[158] R. Jose and V. Singh. Swarming in bacteria: A tale of plasticity in motility behavior. *Journal of the Indian Institute of Science*, 100(3):515–524, 2020.

[159] N. K. Karamanos, A. D. Theocharis, Z. Piperigkou, D. Manou, A. Passi, S. S. Skandalis, D. H. Vynios, V. Orian-Rousseau, S. Ricard-Blum, C. E. H. Schmelzer, et al. A guide to the composition and functions of the extracellular matrix. *The Federation of European Biochemical Sciences Journal*, 288(24):6850–6912, 2021.

[160] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3(6):422–440, 2021.

[161] J. C. Kasemeier-Kulesa, J. M. Teddy, L. Postovit, E. A. Seftor, R. E. B. Seftor, M. J. C. Hendrix, and P. M. Kulesa. Reprogramming multipotent tumor cells with the embryonic neural crest microenvironment. *Developmental Dynamics*, 237(10):2657–2666, 2008.

[162] E. F. Keller and L. A. Segel. Initiation of slime mold aggregation viewed as an instability. *Journal of Theoretical Biology*, 26(3):399–415, 1970.

[163] E. F. Keller and L. A. Segel. Traveling bands of chemotactic bacteria: a theoretical analysis. *Journal of Theoretical Biology*, 30(2):235–248, 1971.

[164] J. B. Keller and R. M. Lewis. Asymptotic methods for partial differential equations: The reduced wave equation and Maxwell's equations. In *Surveys in Mathematics and its Applications*, pages 1–82. Springer, 1995.

[165] K. Kessenbrock, V. Plaks, and Z. Werb. Matrix metalloproteinases: regulators of the tumor microenvironment. *Cell*, 141(1):52–67, 2010.

[166] I. S. Kim, S. Heilmann, E. R. Kansler, Y. Zhang, M. Zimmer, K. Ratnakumar, R. L. Bowman, T. Simon-Vermot, M. Fennell, R. Garippa, et al. Microenvironment-derived factors driving metastatic plasticity in melanoma. *Nature Communications*, 8(1):14343, 2017.

[167] M.-C. Kim, R. Li, R. Abeyaratne, R. D. Kamm, and H. H. Asada. A computational modeling of invadopodia protrusion into an extracellular matrix fiber network. *Scientific Reports*, 12(1):1231, 2022.

[168] M.-C. Kim, Y. R. Silberberg, R. Abeyaratne, R. D. Kamm, and H. H. Asada. Computational modeling of three-dimensional ECM-rigidity sensing to guide directed cell migration. *Proceedings of the National Academy of Sciences*, 115(3):E390–E399, 2018.

[169] M.-C. Kim, J. Whisler, Y. R. Silberberg, R. D. Kamm, and H. H. Asada. Cell invasion dynamics into a three dimensional extracellular matrix fibre network. *PLOS Computational Biology*, 11(10):e1004535, 2015.

[170] Y. Kim and H. G. Othmer. Hybrid models of cell and tissue dynamics in tumor growth. *Mathematical Biosciences and Engineering*, 12(6):1141, 2015.

[171] J. R. King and P. McCabe. On the Fisher–KPP equation with nonlinear diffusion. *Nonlinearity*, 16(5):2091–2112, 2003.

[172] S. Klopfstein, M. Currat, and L. Excoffier. The fate of mutations surfing on the wave of a range expansion. *Molecular Biology and Evolution*, 23(3):482–490, 2006.

[173] N. Kolbe, N. Sfakianakis, C. Stinner, C. Surulescu, and J. Lenz. Modeling multiple taxis: tumor invasion with phenotypic heterogeneity, haptotaxis, and unilateral interspecies repellence. *Discrete and Continuous Dynamical Systems - B*, 26(1):443–481, 2021.

[174] A. N. Kolmogorov, I. Petrovskii, and N. S. Piskunov. A study of the equation of diffusion with increase in the quantity of matter, and its application to a biological problem. *Moscow University Biological Sciences Bulletin*, 1(6):1–25, 1937.

[175] J. Konen, E. Summerbell, B. Dwivedi, K. Galior, Y. Hou, L. Rusnak, A. Chen, J. Saltz, W. Zhou, L. H. Boise, et al. Image-guided genomics of phenotypically heterogeneous populations reveals vascular signalling during symbiotic collective cancer invasion. *Nature Communications*, 8(1):15078, 2017.

[176] R. E. Kopp. Pontryagin maximum principle. In *Mathematics in Science and Engineering*, volume 5, pages 255–279. Elsevier, 1962.

[177] B. Koren. *A Robust Upwind Discretization Method for Advection, Diffusion and Source Terms*, volume 45. Centrum Wiskunde & Informatica, 1993.

[178] M. Kot. *Elements of Mathematical Ecology*. Cambridge University Press, 2001.

[179] J. H. Kristensen and M. A. Karsdal. *Biochemistry of Collagens, Laminins and Elastin*, chapter 30 - Elastin, pages 197–201. Academic Press, 2016.

[180] T. Kurata, K. Tamura, H. Kaneda, T. Nogami, H. Uejima, G. Asai, K. Nakagawa, and M. Fukuoka. Effect of re-treatment with gefitinib ('Iressa', ZD1839) after acquisition of resistance. *Annals of Oncology*, 15(1):173–174, 2004.

[181] J. H. Lagergren, J. T. Nardini, R. E. Baker, M. J. Simpson, and K. B. Flores. Biologically-informed neural networks guide mechanistic modeling from sparse experimental data. *PLOS Computational Biology*, 16(12):e1008462, 2020.

[182] J. H. Lagergren, J. T. Nardini, G. M. Lavigne, E. M. Rutter, and K. B. Flores. Learning partial differential equations for biological transport models from noisy spatio-temporal data. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 476(2234):20190800, 2020/07/06 2020.

[183] K.-Y. Lam and Y. Lou. *Introduction to Reaction–Diffusion Equations: Theory and Applications to Spatial Ecology and Evolutionary Biology*. Springer, 2022.

[184] L. Lamalice, F. Le Boeuf, and J. Huot. Endothelial cell migration during angiogenesis. *Circulation Research*, 100(6):782–794, 2007.

[185] J. R. Lange and B. Fabry. Cell and tissue mechanics in cell migration. *Experimental Cell Research*, 319(16):2418–2423, 2013.

[186] A. Lazarides. Coarse-graining a restricted solid-on-solid model. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 73(4):041605, 2006.

[187] B. Lee, J. Konen, S. Wilkinson, A. I. Marcus, and Y. Jiang. Local alignment vectors reveal cancer cell-induced ECM fiber remodeling dynamics. *Scientific Reports*, 7(1):39498, 2017.

[188] M. C. Lloyd, J. J. Cunningham, M. M. Bui, R. J. Gillies, J. S. Brown, and R. A. Gatenby. Darwinian dynamics of intratumoral heterogeneity: not solely random mutations but also variable environmental selection forces. *Cancer Research*, 76(11):3136–3144, 2016.

[189] T. Lorenzi, A. Lorz, and B. Perthame. On interfaces between cell populations with different mobilities. *Kinetic and Related Models*, 10(1):299–311, 2017.

[190] T. Lorenzi, F. R. Macfarlane, and C. Villa. Discrete and continuum models for the evolutionary and spatial dynamics of cancer: a very short introduction through two case studies. *Trends in Biomathematics: Modeling Cells, Flows, Epidemics, and the Environment*, pages 359–380, 2020.

[191] T. Lorenzi and K. J. Painter. Trade-offs between chemotaxis and proliferation shape the phenotypic structuring of invading waves. *International Journal of Non-Linear Mechanics*, 139:103885, 2022.

[192] T. Lorenzi, B. Perthame, and X. Ruan. Invasion fronts and adaptive dynamics in a model for the growth of cell populations with heterogeneous mobility. *European Journal of Applied Mathematics*, 33(4):766–783, 2022.

[193] P. Lu and Y. Lu. Born to run? Diverse modes of epithelial migration. *Frontiers in Cell and Developmental Biology*, page 2187, 2021.

[194] F. R. Macfarlane, T. Lorenzi, and K. J. Painter. The impact of phenotypic heterogeneity on chemotactic self-organisation. *Bulletin of Mathematical Biology*, 84(12):143, 2022.

[195] F. R. Macfarlane, X. Ruan, and T. Lorenzi. Individual-based and continuum models of phenotypically heterogeneous growing cell populations. *AIMS Bioengineering*, 9(1):68–92, 2022.

[196] P. K. Maini, D. L. S. McElwain, and D. I. Leavesley. Traveling wave model to interpret a wound-healing cell migration assay for human peritoneal mesothelial cells. *Tissue Engineering*, 10(3-4):475–482, 2004.

[197] A. A. Malik, B. Wennberg, and P. Gerlee. The impact of elastic deformations of the extracellular matrix on cell migration. *Bulletin of Mathematical Biology*, 82:1–19, 2020.

[198] B. P. Marchant, J. Norbury, and J. A. Sherratt. Travelling wave solutions to a haptotaxis-dominated model of malignant invasion. *Nonlinearity*, 14(6):1653, 2001.

[199] N. G. Marculis, M. L. Evenden, and M. A Lewis. Modeling the dispersal–reproduction trade-off in an expanding population. *Theoretical Population Biology*, 134:147–159, 2020.

[200] N. G. Marculis, J. Garnier, R. Lui, and M. A. Lewis. Inside dynamics for stage-structured integrodifference equations. *Journal of Mathematical Biology*, 80:157–187, 2020.

[201] N. G. Marculis and M. A. Lewis. Inside dynamics of integrodifference equations with mutations. *Bulletin of Mathematical Biology*, 82(1):7, 2020.

[202] N. K. Martin, E. A. Gaffney, R. A. Gatenby, and P. K. Maini. Tumour–stromal interactions in acid-mediated invasion: a mathematical model. *Journal of Theoretical Biology*, 267(3):461–470, 2010.

[203] S. F. Martina-Perez, S. W. S. Johnson, R. M. Crossley, J. C. Kasemeier, P. M. Kulesa, and R. E. Baker. Optimal control in combination therapy for heterogeneous cell populations with drug synergies. *Bulletin of Mathematical Biology*, 87(11):159, 2025.

[204] W. D. Martinson, R. McLennan, J. M. Teddy, M. C. McKinney, L. A. Davidson, R. E. Baker, H. M. Byrne, P. M. Kulesa, and P. K. Maini. Dynamic fibronectin assembly and remodeling by leader neural crest cells prevents jamming in collective cell migration. *eLife*, 12:e83792, 2023.

[205] D. Matthes, E.-M. Rott, and A. Schlichting. Diffusive transport on the real line: semi-contractive gradient flows and their discretization. *Nonlinearity*, 38(11):115005, 2025.

[206] S. W. McCue, M. El-Hachem, and M. J. Simpson. Exact sharp-fronted travelling wave solutions of the Fisher–KPP equation. *Applied Mathematics Letters*, 114:106918, 2021.

[207] J. B. McGillen. *Mathematical Modelling of Metabolism and Acidity in Cancer*. PhD thesis.

[208] J. B. McGillen, E. A. Gaffney, N. K. Martin, and P. K. Maini. A general reaction–diffusion model of acidity in cancer invasion. *Journal of Mathematical Biology*, 68(5):1199–1224, 2014.

[209] R. McLennan, L. Dyson, K. W. Prather, J. A. Morrison, R. E. Baker, P. K. Maini, and P. M. Kulesa. Multiscale mechanisms of cell migration during development: theory and experiment. *Development*, 139(16):2935–2944, 2012.

[210] R. McLennan, L. J. Schumacher, J. A. Morrison, J. M. Teddy, D. A. Ridenour, A. C. Box, C. L. Semerad, H. Li, W. McDowell, D. Kay, et al. Neural crest migration is driven by a few trailblazer cells with a unique molecular signature narrowly confined to the invasive front. *Development*, 142(11):2014–2025, 2015.

[211] R. McLennan, L. J. Schumacher, J. A. Morrison, J. M. Teddy, D. A. Ridenour, A. C. Box, C. L. Semerad, H. Li, W. McDowell, D. Kay, P. K. Maini, R. E. Baker, and P. M. Kulesa. Neural crest migration is driven by a few trailblazer cells with a unique molecular signature narrowly confined to the invasive front. *Development*, 142(11):2014–2025, 2015.

[212] F. Merino-Casallo, M. J. Gomez-Benito, S. Hervas-Raluy, and J. M. Garcia-Aznar. Unravelling cell migration: defining movement from the cell surface. *Cell Adhesion & Migration*, 16(1):25–64, 2022.

[213] J. H. Merkin and D. J. Needham. Propagating reaction–diffusion waves in a simple isothermal quadratic autocatalytic chemical system. *Journal of Engineering Mathematics*, 23(4):343–356, 1989.

[214] J. H. Merkin and D. J. Needham. The development of travelling waves in a simple isothermal chemical system ii. cubic autocatalysis with quadratic and linear decay. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 430(1879):315–345, 1990.

[215] D. A. Messenger and D. M. Bortz. Weak SINDy for partial differential equations. *Journal of Computational Physics*, 443:110525, 2021.

[216] J. Metzcar, B. S. Duggan, B. Fischer, M. Murphy, R. Heiland, and P. Macklin. A simple framework for agent-based modeling with extracellular matrix. *Bulletin of Mathematical Biology*, 87(3):43, 2025.

[217] J. Metzcar, Y. Wang, R. Heiland, and P. Macklin. A review of cell-based computational modeling in cancer biology. *Journal of Clinical Oncology Clinical Cancer Informatics*, 2:1–13, 2019.

[218] B. C. Miller, D. R. Sen, R. Al Abosy, K. Bi, Y. V. Virkud, M. W. LaFleur, K. B. Yates, A. Lako, K. Felt, G. S. Naik, et al. Subsets of exhausted CD8+ T cells differentially mediate tumor control and respond to checkpoint blockade. *Nature Immunology*, 20(3):326–336, 2019.

[219] J. A. Mitchel, A. Das, M. J. O'Sullivan, I. T. Stancil, S. J. DeCamp, S. Koehler, O. H. Ocaña, J. P. Butler, J. J. Fredberg, M. A. Nieto, et al. In primary airway epithelial cells, the unjamming transition is distinct from the epithelial-to-mesenchymal transition. *Nature Communications*, 11(1):5053, 2020.

[220] B. Morris, L. Curtin, A. Hawkins-Daarud, M. E. Hubbard, R. Rahman, S. J. Smith, D. Auer, N. L. Tran, L. S. Hu, J. M. Eschbacher, K. A. Smith, A. Stokes, K. R. Swanson, and M. R. Owen. Identifying the spatial and temporal dynamics of molecularly-distinct glioblastoma sub-populations. *Mathematical Biosciences and Engineering*, 17(5):4905, 2020.

[221] K. W. Morton and D. F. Mayers. *Numerical Solution of Partial Differential Equations: An Introduction*. Cambridge University Press, Cambridge, 2 edition, 2005.

[222] J. M. Muncie and V. M. Weaver. The physical and biochemical properties of the extracellular matrix regulate cell fate. *Current Topics in Developmental Biology*, 130:1–37, 2018.

[223] R. J. Murphy, P. R. Buenzli, R. E. Baker, and M. J. Simpson. A one-dimensional individual-based mechanical model of cell movement in heterogeneous tissues and its coarse-grained approximation. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 475(2227):20180838, 2019.

[224] J. D. Murray. *Mathematical Biology*. Springer, Berlin / Heidelberg, Germany, 1993.

[225] J. D. Murray. *Mathematical Biology II: Spatial models and biomedical applications*, volume 3. Springer New York, 2001.

[226] J. D. Murray. *Mathematical Biology I: An Introduction*. Springer, 2002.

[227] J. D. Murray, E. A. Stanley, and D. L. Brown. On the spatial spread of rabies among foxes. *Proceedings of the Royal society of London. Series B. Biological Sciences*, 229(1255):111–150, 1986.

[228] P. J. Murray, C. M. Edwards, M. J. Tindall, and P. K. Maini. From a discrete to a continuum model of cell dynamics in one dimension. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 80(3):031912, 2009.

[229] P. J. Murray, A. Walter, A. G. Fletcher, C. M. Edwards, M. J. Tindall, and P. K. Maini. Comparing a discrete and continuum model of the intestinal crypt. *Physical Biology*, 8(2):026011, 2011.

[230] H. Nagase, R. Visse, and G. Murphy. Structure and function of matrix metalloproteinases and TIMPs. *Cardiovascular Research*, 69(3):562–573, 2006.

[231] H. Nagase and J. F. Woessner. Matrix metalloproteinases. *Journal of Biological Chemistry*, 274(31):21491–21494, 1999.

[232] J. T. Nardini. Forecasting and predicting stochastic agent-based model data with biologically-informed neural networks. *Bulletin of Mathematical Biology*, 86(11):130, 2024.

[233] J. T. Nardini, J. H. Lagergren, A. Hawkins-Daarud, L. Curtin, B. Morris, E. M. Rutter, K. R. Swanson, and K. B. Flores. Learning equations from biological data with limited time samples. *Bulletin of Mathematical Biology*, 82(9):119, 2020.

[234] D. J. Needham and J. H. Merkin. The development of travelling waves in a simple isothermal chemical system with general orders of autocatalysis and decay. *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences*, 337(1646):261–274, 1991.

[235] D. F. Newgreen, G. J. Pettet, and K. A. Landman. Chemotactic cellular migration: smooth and discontinuous travelling wave solutions. *SIAM Journal on Applied Mathematics*, 63(5):1666–1681, 2003.

[236] K. Newton, A. Strasser, N. Kayagaki, and V. M. Dixit. Cell death. *Cell*, 187(2):235–256, 2024.

[237] Y. T. Nguyen Edalgo and A. N. Ford Versypt. Mathematical modeling of metastatic cancer migration through a remodeling extracellular matrix. *Processes*, 6(5):58, 2018.

[238] D. Nichol, M. Robertson-Tessi, P. Jeavons, and A. R. A. Anderson. Stochasticity in the genotype-phenotype map: implications for the robustness and persistence of bet-hedging. *Genetics*, 204(4):1523–1539, 2016.

[239] G. L. Nicolson. Generation of phenotypic diversity and progression in metastatic tumor cells. *Cancer and Metastasis Reviews*, 3:25–42, 1984.

[240] P. C. Nowell. The clonal evolution of tumor cell populations: Acquired genetic lability permits stepwise selection of variant sublines and underlies tumor progression. *Science*, 194(4260):23–28, 1976.

[241] B. Nusgens, C. Merrill, C. Lapiere, and E. Bell. Collagen biosynthesis by cells in a tissue equivalent matrix *in vitro*. *Collagen and Related Research*, 4(5):351–363, 1984.

[242] A. Okubo. Introduction: The mathematics of ecological diffusion. In Akira Okubo and Simon A. Levin, editors, *Diffusion and Ecological Problems: Modern Perspectives*, Interdisciplinary Applied Mathematics, pages 1–9. Springer, New York, 2001.

[243] A. Okubo and P. Kareiva. Some examples of animal diffusion. In Akira Okubo and Simon A. Levin, editors, *Diffusion and Ecological Problems: Modern Perspectives*, Interdisciplinary Applied Mathematics, pages 170–196. Springer, New York, 2001.

[244] A. Okubo, P. K. Maini, M. H. Williamson, and J. D. Murray. On the spatial spread of the grey squirrel in Britain. *Proceedings of the Royal Society of London. B. Biological Sciences*, 238(1291):113–125, 1989.

[245] L. Olsen, P. K. Maini, J. A. Sherratt, and J. Dallon. Mathematical modelling of anisotropy in fibrous connective tissue. *Mathematical Biosciences*, 158(2):145–170, 1999.

[246] L. Olsen, P. K. Maini, J. A. Sherratt, and B. Marchant. Simple modelling of extracellular matrix alignment in dermal wound healing I. Cell flux induced alignment. *Computational and Mathematical Methods in Medicine*, 1(3):175–192, 1998.

[247] L. Olsen, J. A. Sherratt, P. K. Maini, and F. Arnold. A mathematical model for the capillary endothelial cell-extracellular matrix interactions in wound-healing

angiogenesis. *IMA Journal of Mathematics Applied in Medicine and Biology*, 14(4):261–281, 1997.

[248] H. A. Orr. Fitness and its role in evolutionary genetics. *Nature Reviews Genetics*, 10(8):531–539, 2009.

[249] J. M. Osborne, A. G. Fletcher, J. M. Pitt-Francis, P. K. Maini, and D. J. Gavaghan. Comparing individual-based approaches to modelling the self-organization of multicellular tissues. *PLOS Computational Biology*, 13(2):e1005387, 2017.

[250] L. Oswald, S. Grosser, D. M. Smith, and J. A. Käs. Jamming transitions in cancer. *Journal of Physics D: Applied Physics*, 50(48):483001, 2017.

[251] K. J. Painter. Modelling cell migration strategies in the extracellular matrix. *Journal of Mathematical Biology*, 58(4):511 – 543, 2009.

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

[253] K. J. Painter and J. A. Sherratt. Modelling the movement of interacting cell populations. *Journal of Theoretical Biology*, 225(3):327–339, 2003.

[254] D. Pally and A. Naba. Extracellular matrix dynamics: A key regulator of cell migration across length-scales and systems. *Current Opinion in Cell Biology*, 86, 2024.

[255] G. Pang and G. E. Karniadakis. Physics-informed learning machines for partial differential equations: Gaussian processes versus neural networks. In *Emerging Frontiers in Nonlinear Science*, pages 323–343. Springer, 2020.

[256] P. Patra and S. Klumpp. Population dynamics of bacterial persistence. *PLOS One*, 8(5):e62814, 2013.

[257] C. J. Penington, B. D. Hughes, and K. A. Landman. Building macroscale models from microscale probabilistic models: A general probabilistic approach for nonlinear diffusion and multispecies phenomena. *Physical Review E*, 84(4):041120, 2011.

[258] M. Pensalfini and A. B. Tepole. Mechano-biological and bio-mechanical pathways in cutaneous wound healing. *PLOS Computational Biology*, 19(3):e1010902, 2023.

[259] B. Perthame. PDE models for chemotactic movements: parabolic, hyperbolic and kinetic. *Applications of Mathematics*, 49(6):539–564, 2004.

[260] B. Perthame, N. Vauchelet, and Z. Wang. The flux limited Keller–Segel system; properties and derivation from kinetic equations. *Revista Matemática Iberoamericana*, 36(2):357–386, 2019.

[261] A. J. Perumpanani and H. M. Byrne. Extracellular matrix concentration exerts selection pressure on invasive cells. *European Journal of Cancer*, 35(8):1274–1280, 1999.

[262] S. V. Petrovskii and B.-L. Li. *Exactly Solvable Models of Biological Invasion*. CRC Press, 2005.

[263] G. J. Pettet, D. L. McElwain, and J. Norbury. Lotka-Volterra equations with chemotaxis: walls, barriers and travelling waves. *IMA Journal of Mathematics Applied in Medicine and Biology*, 17(4):395–413, 2000.

[264] K. Pham, A. Chauviere, H. Hatzikirou, X. Li, H. M. Byrne, V. Cristini, and J. Lowengrub. Density-dependent quiescence in glioma invasion: instability in a simple reaction–diffusion model for the migration/proliferation dichotomy. *Journal of Biological Dynamics*, 6(sup1):54–71, 2012.

[265] T. Philippi and J. Seger. Hedging one's evolutionary bets, revisited. *Trends in Ecology & Evolution*, 4(2):41–44, 1989.

[266] M. Philipps, A. Körner, J. Vanhoefer, D. Pathirana, and J. Hasenauer. Non-negative universal differential equations with applications in systems biology. *IFAC-PapersOnLine*, 58(23):25–30, 2024.

[267] M. Philipps, N. Schmid, and J. Hasenauer. Universal differential equations for systems biology: Current state and open problems. *bioRxiv*, pages 2024–11, 2024.

[268] J. Pleyer and C. Fleck. Agent-based models in cellular systems. *Frontiers in Physics*, 10:968409, 2023.

[269] L. Podina, B. Eastman, and M. Kohandel. Universal physics-informed neural networks: Symbolic differential operator discovery with sparse data. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 27948–27956. PMLR, 2023.

[270] L. Podina, A. Ghodsi, and M. Kohandel. Learning chemotherapy drug action via universal physics-informed neural networks. *arXiv preprint arXiv:2404.08019*, 2024.

[271] R. Poincloux, F. Lizárraga, and P. Chavrier. Matrix invasion by tumour cells: a focus on MT1-MMP trafficking to invadopodia. *Journal of Cell Science*, 122(17):3015–3024, 2009.

[272] T. P. Prescott, K. Zhu, M. Zhao, and R. E. Baker. Quantifying the impact of electric fields on single-cell motility. *Biophysical Journal*, 120(16):3363–3373, 2021.

[273] L. Qin, D. Yang, W. Yi, H. Cao, and G. Xiao. Roles of leader and follower cells in collective cell migration. *Molecular Biology of the Cell*, 32(14):1267–1272, 2021.

[274] P. J. Quesenberry and J. M. Aliotta. Cellular phenotype switching and microvesicles. *Advanced Drug Delivery Reviews*, 62(12):1141–1148, 2010.

[275] E. M. Rabie, S. X. Zhang, A. P. Kourouklis, A. N. Kilinc, A. K. Simi, D. C. Radisky, J. Tien, and C. M. Nelson. Matrix degradation and cell proliferation are coupled to promote invasion and escape from an engineered human breast microtumor. *Integrative Biology*, 13(1):17–29, 2021.

[276] C. Rackauckas, Y. Ma, J. Martensen, C. Warner, K. Zubov, R. Supekar, D. Skinner, A. Ramadhan, and A. Edelman. Universal differential equations for scientific machine learning. *arXiv preprint arXiv:2001.04385*, 2020.

[277] N. Rahman, E. O'Neill, M. Irnaten, D. Wallace, and C. O'Brien. Corneal stiffness and collagen cross-linking proteins in glaucoma: Potential for novel therapeautic strategy. *Journal of Ocular Pharmacology and Therapeutics*, 36(8), 2020.

[278] M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems

involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.

[279] M. Reina-Campos and A. W. Goldrath. Antitumour T cells stand the test of time. *Nature*, 576(7787):392–394, 2019.

[280] E. G. Rens and R. M. H. Merks. Cell shape and durotaxis explained from cell-extracellular matrix forces and focal adhesion dynamics. *iScience*, 23(9):101488, 2020.

[281] S. G. Ricoult, T. E. Kennedy, and D. Juncker. Substrate-bound protein gradients to study haptotaxis. *Frontiers in Bioengineering and Biotechnology*, 3, 2015.

[282] M. Robertson-Tessi, R. J. Gillies, R. A. Gatenby, and A. R. A. Anderson. Impact of metabolic heterogeneity on tumor growth, invasion, and treatment outcomes. *Cancer Research*, 75(8):1567–1579, 2015.

[283] L. Roques, J. Garnier, F. Hamel, and E. K. Klein. Allee effect promotes diversity in traveling waves of colonization. *Proceedings of the National Academy of Sciences*, 109(23):8828–8833, 2012.

[284] P. Rørth. Collective cell migration. *Annual Review of Cell and Developmental Biology*, 25(1):407–429, 2009.

[285] R. J. H. Ross, C. A. Yates, and R. E. Baker. Inference of cell–cell interactions from population density characteristics and cell trajectories on static and growing domains. *Mathematical Biosciences*, 264:108–118, 2015.

[286] P. Royer. Advection–diffusion in porous media with low scale separation: modelling via higher-order asymptotic homogenisation. *Transport in Porous Media*, 128(2):511–551, 2019.

[287] S. H. Rudy, S. L. Brunton, J. L. Proctor, and J. N. Kutz. Data-driven discovery of partial differential equations. *Science Advances*, 3(4):e1602614, 2017.

[288] F. Sabeh, R. Shimizu-Hirota, and S. J. Weiss. Protease-dependent versus-independent cancer cell invasion programs: three-dimensional amoeboid movement revisited. *Journal of Cell Biology*, 185(1):11–19, 2009.

[289] M. M. Salek, F. Carrara, V. Fernandez, J. S. Guasto, and R. Stocker. Bacterial chemotaxis in a microfluidic t-maze reveals strong phenotypic heterogeneity in chemotactic sensitivity. *Nature Communications*, 10(1):1877, 2019.

[290] F. Sánchez-Garduño and P. K. Maini. Travelling wave phenomena in some degenerate reaction–diffusion equations. *Journal of Differential Equations*, 117(2):281–319, 1995.

[291] A. Saraswathibhatla, D. Indana, and O. Chaudhuri. Cell–extracellular matrix mechanotransduction in 3D. *Nature Reviews Molecular Cell Biology*, 24(77):495–516, 2023.

[292] O. Saut, J.-B. Lagaert, T. Colin, and H. M. Fathallah-Shaykh. A multilayer grow-or-go model for GBM: effects of invasive cells and anti-angiogenesis on growth. *Bulletin of Mathematical Biology*, 76:2306–2333, 2014.

[293] G. Schaller and M. Meyer-Hermann. Continuum versus discrete model: a comparison for multicellular tumour spheroids. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 364(1843):1443–1464, 2006.

[294] D. K. Schlüter, I. Ramis-Conde, and M. A. J. Chaplain. Computational modeling of single-cell migration: The leading role of extracellular matrix fibers. *Biophysical Journal*, 103(6):1141–1151, 2012.

[295] F. Schreiber, S. Littmann, G. Lavik, S. Escrig, A. Meibom, M. M. M. Kuypers, and M. Ackermann. Phenotypic heterogeneity driven by nutrient limitation promotes growth in fluctuating environments. *Nature Microbiology*, 1(6):1–7, 2016.

[296] R. D. Schreiber, L. J. Old, and M. J. Smyth. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. *Science*, 331(6024):1565–1570, 2011.

[297] M. Scianna and L. Preziosi. A cellular potts model for analyzing cell migration across constraining pillar arrays. *Axioms*, 10(1):32, 2021.

[298] M. Scianna, L. Preziosi, and K. Wolf. A cellular potts model simulating cell migration on and in matrix environments. *Mathematical Biosciences and Engineering*, 10(1):235–261, 2013.

[299] D. A. Selby, M. Sprang, J. Ewald, and S. J. Vollmer. Beyond the black box with biologically informed neural networks. *Nature Reviews Genetics*, pages 1–2, 2025.

[300] B. G. Sengers, C. P. Please, and R. O. C. Oreffo. Experimental characterization and computational modelling of two-dimensional cell spreading for skeletal regeneration. *Journal of the Royal Society Interface*, 4(17):1107–1117, 2007.

[301] X. Shao, I. N. Taha, K. R. Clauser, Y. Gao, and A. Naba. MatrisomeDB: the ECM-protein knowledge database. *Nucleic Acids Research*, 48(D1):D1136–D1144, 2020.

[302] S. V. Sharma, D. Y. Lee, B. Li, M. P. Quinlan, F. Takahashi, S. Maheswaran, U. McDermott, N. Azizian, L. Zou, M. A. Fischbach, et al. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell*, 141(1):69–80, 2010.

[303] A. Shellard and R. Mayor. Rules of collective migration: from the wildebeest to the neural crest. *Philosophical Transactions of the Royal Society B*, 375(1807):20190387, 2020.

[304] J. A. Sherratt. Chemotaxis and chemokinesis in eukaryotic cells: The Keller-Segel equations as an approximation to a detailed model. *Bulletin of Mathematical Biology*, 56(1):129–146, 1994.

[305] J. A. Sherratt and J. C. Dallon. Theoretical models of wound healing: past successes and future challenges. *Comptes Rendus Biologies*, 325(5):557–564, 2002.

[306] J. A. Sherratt and J. D. Murray. Models of epidermal wound healing. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 241(1300):29–36, 1990.

[307] J. A. Sherratt, E. H. Sage, and J. D. Murray. Chemical control of eukaryotic cell movement: A new model. *Journal of Theoretical Biology*, 162(1):23–40, 1993.

[308] M. J. Simpson and R. E. Baker. Parameter identifiability, parameter estimation and model prediction for differential equation models. *arXiv preprint arXiv:2405.08177*, 2024.

[309] M. J. Simpson, A. P. Browning, D. J. Warne, O. J. Maclaren, and R. E. Baker. Parameter identifiability and model selection for sigmoid population growth models. *Journal of Theoretical Biology*, 535:110998, 2022.

[310] M. J. Simpson, B. D. Hughes, and K. A. Landman. Diffusing populations: Ghosts or folks? *Australasian Journal of Engineering Education*, 15(2):59–68, 2009.

[311] M. J. Simpson, K. A. Landman, and B. D. Hughes. Multi-species simple exclusion processes. *Physica A: Statistical Mechanics and its Applications*, 388(4):399–406, 2009.

[312] M. J. Simpson, K. A. Landman, and B. D. Hughes. Pathlines in exclusion processes. *Physical Review E*, 79(3):031920, 2009.

[313] M. Slatkin. Hedging one's evolutionary bets. *Nature*, 250(5469):704–705, 1974.

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

[315] J. Sprent and C. D. Surh. Normal T cell homeostasis: the conversion of naive cells into memory-phenotype cells. *Nature Immunology*, 12(6):478–484, 2011.

[316] R. E. A. Stace, T. Stiehl, M. A. J. Chaplain, A. Marciniak-Czochra, and T. Lorenzi. Discrete and continuum phenotype-structured models for the evolution of cancer cell populations under chemotherapy. *Mathematical Modelling of Natural Phenomena*, 15:14, 2020.

[317] P. A. Stephens, W. J. Sutherland, and R. P. Freckleton. What is the Allee effect? *Oikos*, pages 185–190, 1999.

[318] T. L. Stepien, E. M. Rutter, and Y. Kuang. Traveling waves of a go-or-grow model of glioma growth. *SIAM Journal on Applied Mathematics*, 78(3):1778–1801, 2018.

[319] W. G. Stetler-Stevenson, S. Aznavoorian, and L. A. Liotta. Tumor cell interactions with the extracellular matrix during invasion and metastasis. *Annual Review of Cell Biology*, 9(1):541–573, 1993.

[320] B. M. Stokes, T. Rogers, and R. James. Speed and shape of population fronts with density-dependent diffusion. *Bulletin of Mathematical Biology*, 86(12):1–18, 2024.

[321] M. A. R. Strobl, A. L. Krause, M. Damaghi, R. Gillies, A. R. A. Anderson, and P. K. Maini. Mix and match: phenotypic coexistence as a key facilitator of cancer invasion. *Bulletin of Mathematical Biology*, 82:1–26, 2020.

[322] E. R. Summerbell, J. K. Mouw, J. S. K. Bell, C. M. Knippler, B. Pedro, J. L. Arnst, T. O. Khatib, R. Commander, B. G. Barwick, J. Konen, et al. Epigenetically heterogeneous tumor cells direct collective invasion through filopodia-driven fibronectin micropatterning. *Science Advances*, 6(30):eaaz6197, 2020.

[323] L. Sun, H. Gao, S. Pan, and J.-X. Wang. Surrogate modeling for PDEs with random inputs using physics-informed neural networks. *Journal of Computational Physics*, 408:109770, 2020.

[324] S. Suresh. Biomechanics and biophysics of cancer cells. *Acta Biomaterialia*, 3(4):413–438, 2007.

[325] S. Syga, H. P. Jain, M. Krellner, H. Hatzikirou, and A. Deutsch. Evolution of phenotypic plasticity leads to tumor heterogeneity with implications for therapy. *PLOS Computational Biology*, 20(8):e1012003, 2024.

[326] A. Szabó, K. Varga, Tams Garay, Balázs Hegedűs, and András Czirók. Invasion from a cell aggregate—the roles of active cell motion and mechanical equilibrium. *Physical Biology*, 9(1):016010, 2012.

[327] A. Tam, J. E. F. Green, S. Balasuriya, E. L. Tek, J. M. Gardner, J. F. Sundstrom, V. Jiranek, and B. J. Binder. Nutrient-limited growth with non-linear cell diffusion as a mechanism for floral pattern formation in yeast biofilms. *Journal of Theoretical Biology*, 448:122–141, 2018.

[328] A. M. Tartakovsky, C. O. Marrero, P. Perdikaris, G. D. Tartakovsky, and D. Barajas-Solano. Physics-informed deep neural networks for learning parameters and constitutive relationships in subsurface flow problems. *Advances in Water Resources*, 134:103548, 2020.

[329] P. R. Taylor, R. E. Baker, M. J. Simpson, and C. A. Yates. Coupling volume-excluding compartment-based models of diffusion at different scales: Voronoi

and pseudo-compartment approaches. *Journal of the Royal Society Interface*, 13(120):20160336, 2016.

[330] E. Theveneau and R. Mayor. Collective cell migration of epithelial and mesenchymal cells. *Cellular and Molecular Life Sciences*, 70(19):3481–3492, 2013.

[331] J. P. Thiery. Mechanisms of cell migration in the vertebrate embryo. *Cell Differentiation*, 15(1):1–15, 1984.

[332] D. Tosh and J. M. W. Slack. How cells change their phenotype. *Nature Reviews Molecular Cell Biology*, 3(3):187–194, 2002.

[333] X. Trepat, Z. Chen, and K. Jacobson. Cell migration. *Comprehensive Physiology*, 2(4):2369, 2012.

[334] A. Tursynkozha, A. Kashkynbayev, B. Shupeyeva, E. M. Rutter, and Y. Kuang. Traveling wave speed and profile of a "go or grow" glioblastoma multiforme model. *Communications in Nonlinear Science and Numerical Simulation*, 118:107008, 2023.

[335] T. van den Broek, J. A. M. Borghans, and F. Van Wijk. The full spectrum of human naive T cells. *Nature Reviews Immunology*, 18(6):363–373, 2018.

[336] P. Van Liedekerke, M. M. Palm, N. Jagiella, and D. Drasdo. Simulating tissue mechanics with agent-based models: concepts, perspectives and some novel results. *Computational Particle Mechanics*, 2:401–444, 2015.

[337] W. Van Saarloos. Front propagation into unstable states. ii. linear versus nonlinear marginal stability and rate of convergence. *Physical Review A*, 39(12):6367, 1989.

[338] S. Vedel, S. Tay, D. M. Johnston, H. Bruus, and S. R. Quake. Migration of cells in a social context. *Proceedings of the National Academy of Sciences*, 110(1):129–134, 2013.

[339] J.-W. Veening, W. K. Smits, and O. P. Kuipers. Bistability, epigenetics, and bet-hedging in bacteria. *Annual Reviews Microbiology*, 62(1):193–210, 2008.

[340] J. Venegas-Ortiz, R. J. Allen, and M. R. Evans. Speed of invasion of an expanding population by a horizontally transmitted trait. *Genetics*, 196(2):497–507, 2014.

[341] S. A. Vilchez Mercedes, F. Bocci, H. Levine, J. N. Onuchic, M. K. Jolly, and P. K. Wong. Decoding leader cells in collective cancer invasion. *Nature Reviews Cancer*, 21(9):592–604, 2021.

[342] C. Villa, M. A. J. Chaplain, A. Gerisch, and T. Lorenzi. Mechanical models of pattern and form in biological tissues: The role of stress–strain constitutive equations. *Bulletin of Mathematical Biology*, 83(7):80, 2021.

[343] C. Villa, P. K. Maini, A. P. Browning, A. L. Jenner, S. Hamis, and T. Cassidy. Reducing phenotype-structured partial differential equations models of cancer evolution to systems of ordinary differential equations: a generalised moment dynamics approach. *Journal of Mathematical Biology*, 91(2):22, 2025.

[344] R. B. Vinter. *Optimal Control*, volume 2. Springer, 2010.

[345] M. Vishwakarma, J. Di Russo, D. Probst, U. S. Schwarz, T. Das, and J. P. Spatz. Mechanical interactions among followers determine the emergence of leaders in migrating epithelial cell collectives. *Nature Communications*, 9(1):3469, 2018.

[346] R. Visse and H. Nagase. Matrix metalloproteinases and tissue inhibitors of metalloproteinases: structure, function, and biochemistry. *Circulation Research*, 92(8):827–839, 2003.

[347] S. T. Vittadello, S. W. McCue, G. Gunasingh, N. K. Haass, and M. J. Simpson. Examining go-or-grow using fluorescent cell-cycle indicators and cell-cycle-inhibiting drugs. *Biophysical Journal*, 118(6):1243–1247, 2020.

[348] C.-H. Wang, S. Matin, A. B. George, and K. S. Korolev. Pinned, locked, pushed, and pulled traveling waves in structured environments. *Theoretical Population Biology*, 127:102–119, 2019.

[349] K. Wang, D. Wen, X. Xu, R. Zhou, F. Jiang, S. Yuan, Y. Zhang, Y. Gao, and Q. Li. Extracellular matrix stiffness - The central cue for skin fibrosis. *Frontiers in Molecular Biosciences*, 10, 2023.

[350] Z. Wang, J. D. Butner, R. Kerketta, V. Cristini, and T. S. Deisboeck. Simulating cancer growth with multiscale agent-based modeling. In *Seminars in Cancer Biology*, volume 30, pages 70–78. Elsevier, 2015.

[351] P. J. Wangersky. Lotka-Volterra population models. *Annual Review of Ecology and Systematics*, 9:189–218, 1978.

[352] A. M. Weaver. Invadopodia: specialized cell structures for cancer invasion. *Clinical & Experimental Metastasis*, 23:97–105, 2006.

[353] S. D. Webb, J. A. Sherratt, and R. G. Fish. Alterations in proteolytic activity at low pH and its association with invasion: a theoretical model. *Clinical & Experimental Metastasis*, 17:397–407, 1999.

[354] C. B. Weinberg and E. Bell. Regulation of proliferation of bovine aortic endothelial cells, smooth muscle cells, and adventitial fibroblasts in collagen lattices. *Journal of Cellular Physiology*, 122(3):410–414, 1985.

[355] W. Weninger, M. A. Crowley, N. Manjunath, and U. H. von Andrian. Migratory properties of naive, effector, and memory CD8+ T cells. *The Journal of Experimental Medicine*, 194(7):953–966, 2001.

[356] Z. Werb. ECM and cell surface proteolysis: regulating cellular ecology. *Cell*, 91(4):439–442, 1997.

[357] J. West, M. Robertson-Tessi, and A. R. A. Anderson. Agent-based methods facilitate integrative science in cancer. *Trends in Cell Biology*, 33(4):300–311, 2023.

[358] E. J. Wherry and R. Ahmed. Memory CD8 T-cell differentiation during viral infection. *Journal of Virology*, 78(11):5535–5545, 2004.

[359] S. Wiggins and M. Golubitsky. *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, volume 2. Springer, 2003.

[360] D. B. Wilson, H. Byrne, and M. Bruna. Reactions, diffusion, and volume exclusion in a conserved system of interacting particles. *Physical Review E*, 97(6):062137, 2018.

[361] J. Winkler, A. Abisoye-Ogunniyan, K. J. Metcalf, and Z. Werb. Concepts of extracellular matrix remodelling in tumour progression and metastasis. *Nature Communications*, 11(1):5120, 2020.

[362] T. P. Witelski. Merging traveling waves for the porous–Fisher's equation. *Applied Mathematics Letters*, 8(4):57–62, 1995.

[363] T. Worbs and R. Förster. T cell migration dynamics within lymph nodes during steady state: an overview of extracellular and intracellular factors influencing the basal intranodal T cell motility. *Visualizing Immunity*, pages 71–105, 2009.

[364] J.-S. Wu, J. Jiang, B.-J. Chen, K. Wang, Y.-L. Tang, and X.-H. Liang. Plasticity of cancer cell invasion: Patterns and mechanisms. *Translational Oncology*, 14(1):100899, 2021.

[365] A. H. Wyllie. Cell death. *Cytology and Cell Physiology*, pages 755–785, 1987.

[366] B. Xue and S. Leibler. Benefits of phenotypic plasticity for population growth in varying environments. *Proceedings of the National Academy of Sciences*, 115(50):12745–12750, 2018.

[367] K. M. Yamada, J. W. Collins, D. A. Cruz Walma, A. D. Doyle, S. G. Morales, J. Lu, K. Matsumoto, S. S. Nazari, R. Sekiguchi, Y. Shinsato, and S. Wang. Extracellular matrix dynamics in cell migration, invasion and tissue morphogenesis. *International Journal of Experimental Pathology*, 100(3):144–152, 2019.

[368] J. S. Yi, M. A. Cox, and A. J. Zajac. T-cell exhaustion: characteristics, causes and conversion. *Immunology*, 129(4):474–481, 2010.

[369] K. Yoshizato, T. Taira, and N. Yamamoto. Growth inhibition of human fibroblasts by reconstituted collagen fibrils. *Biomedical Research*, 6(2):61–71, 1985.

[370] R. Yu and R. Wang. Learning dynamical systems from data: An introduction to physics-guided deep learning. *Proceedings of the National Academy of Sciences*, 121(27):27948–27956, 2024.

[371] L. Yuan, Y.-Q. Ni, X.-Y. Deng, and S. Hao. A-PINN: Auxiliary physics informed neural networks for forward and inverse problems of nonlinear integro-differential equations. *Journal of Computational Physics*, 462:111260, 2022.

[372] Y. Yuan. Spatial heterogeneity in the tumor microenvironment. *Cold Spring Harbor Perspectives in Medicine*, 6(8), 2016.

[373] B. Yue. Biology of the extracellular matrix: an overview. *Journal of Glaucoma*, page S20, 2014.

[374] Y. Yue and C. Ou. Traveling wavefronts to a model of precursor and differentiated cells: Impacting parameter-structure transition from monostable to bistable, and from monotone to non-monotone. *Journal of Dynamics and Differential Equations*, pages 1–36, 2024.

[375] D. Zhang and C. Jin. Global solvability to a cancer invasion model with remodeling of ECM and porous medium diffusion. *Communications in Mathematical Sciences*, 20(6):1493–1516, 2022.

[376] S. Zhang, A. Chong, and B. D. Hughes. Persistent exclusion processes: Inertia, drift, mixing, and correlation. *Physical Review E*, 100(4):042415, 2019.

[377] E. L. Zoeller, B. Pedro, J. Konen, B. Dwivedi, M. Rupji, N. Sundararaman, L. Wang, J. R. Horton, C. Zhong, B. G. Barwick, et al. Genetic heterogeneity within collective invasion packs drives leader and follower cell phenotypes. *Journal of Cell Science*, 132(19):jcs231514, 2019.