Interactions and dynamics in collective cell behaviour

Carles Falcó i Gandia

St John's College

University of Oxford

A thesis submitted for the degree

Doctor of Philosophy

Trinity 2025

Acknowledgements

I feel incredibly privileged to have spent the last four years working alongside two brilliant mathematicians and extraordinary scientists. Ruth and José, thank you for making this journey so amazing, overflowing with exciting ideas, opportunities, and discussions that were as challenging as they were fun. You've shown me just how incredible the academic life can be, and I look forward to many more conversations and collaborations. Through the tougher moments, your help has been invaluable — your support these past years has meant the world to me, and I will be forever be grateful!

I've also had the pleasure of having fantastic collaborators. Amanda Sierra, Marta Pereira, and the group of Ian Wong at Brown University — your spirit of collaboration is truly inspiring. Thank you for believing in bold research directions that few others would trust. On the theoretical side, Alex Browning, Giulia Celora, Antonio Esposito, Duncan Martinson, and Markus Schmidtchen have truly been role models: offering invaluable advice, teaching me patiently analysis, and reminding me to enjoy life both at Oxford and across countless conferences. Some of these collaborations won't appear in this thesis, yet they've profoundly shaped who I am as a researcher.

The Mathematical Institute has been a truly marvellous place to work. To my officemates and friends, Sam, Simon, and Sergi: thank you for countless stimulating discussions, continuous support, and endless afternoons dissecting whatever drama Sergi brought to the table. My gratitude also goes to all the fellow students who shared this journey with me: Arianna, for our invaluable therapy lunches at St Hugh's; Jun, for always being willing to listen to wild ideas and offering incredibly useful feedback; Zoe for countless bouldering sessions, and for the soothing, invaluable chats; and Becky, for growing alongside me these past four years and being the best academic sister I

could've asked for. Lastly, to Alejandro and Francesco: I hope our dinners, spontaneous trips to the pub (or to Mallorca!), and coffee-therapy sessions continue for many years to come.

To my friends Antoinette, Eve, and Hugh: thank you for keeping me both sane and insane, for our countless visits to BBuona, and for all those once beloved (now dreaded) rowing escapades. To my long-time friends Gal·la, Marc, Maria, Kim, Sergi[1], and Teo: sharing the highs and lows of our journey in Barcelona, Oxford, and beyond has been an absolute gift. And finally, to Adrià and Lucía: I can't put into words how grateful I am to share my everyday life with you. You've turned Oxford into home and made it so much fun! I hope our long research walks, random late night shenanigans, and football-watching evenings never end.

Hi ha molts altres noms que no cabran en aquests agraïments. Han estat quatre anys intensos, i moltes altres persones m'han acompanyat — ja sabeu qui sou. Hi ha amics que t'acompanyen tota la vida i que, tot i estar lluny, sempre es fan presents. A Borre, Fran, Lurbe, Jordi, Jorge, Juan Tomás, Marti i Nacho: gràcies per compartir aquest viatge amb mi.

Finalment, vull agrair tota la meua família. Fa poc vaig llegir que el nom *Helena* signfica *llum* o *torxa*. Tinc la sort de tindre'n dues, i n'estic enormement agraït. Gràcies a mon pare José, a Rafel, a Ana, a Vicent, a David i a Paula; i als meus iaios, Ana i Toni. Oxford no és tant *càlid* com Ontinyent (*o California!*), potser per això us he trobat tant a faltar.

This thesis was made possible by the financial support of the *la Caixa* Foundation (ID 100010434, code LCF/BQ/EU21/11890128) and by funding from the UK Engineering and Physical Sciences Research Council through the Mathematical Institute.

[1] Yes, Sergi has indeed managed to appear twice in here...

Abstract

Collective cell behaviours, from tissue growth and collision to adhesion-driven sorting, underpin key processes in development, regeneration and disease. Yet building models that are both mechanistically accurate and quantitatively calibrated to high-resolution data remains a major challenge. In this thesis, we tackle three interlocking challenges: deriving minimal but predictive continuum and discrete models from basic principles; integrating these models with diverse experimental assays via rigorous inference; and analysing their emergent dynamics to uncover the mechanisms driving collective phenomena.

We begin with the classical Fisher–KPP and Porous–Fisher equations for tissue spreading, which we combine with a Bayesian inference framework to infer motility and proliferation parameters from epithelial monolayer assays. We then extend this picture by incorporating cell cycle data into a two-stage continuum model, revealing how crowding delays cell cycle progression and shapes proliferation patterns in expanding tissues. Extending to tissue–tissue interactions, we show that only models accounting for population pressure can reproduce the sharp interfaces observed in monolayer collision experiments. We then move to the study of adhesive interactions between cells. To bring classical nonlocal adhesion models closer to quantitative data, we derive a local fourth-order model for cell–cell adhesion in the short-range interaction limit, and show that it preserves the full phenomenology of sorting patterns in the context of differential adhesion. Finally, we build and analyse an agent-based spheroid model to explain emergent collective orbiting driven by boundary curvature, cell–cell, and cell–matrix adhesion. Together, these advances pave the way for the next generation of quantitative biology, where mechanistic modelling and data-driven methods combine to yield predictive, experimentally grounded frameworks.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

During development, embryos, which initially consist of a single cell, grow and self-organise to form complex and functional multicellular organisms. As embryos grow, a wide range of developmental processes are regulated by the coordinated activity of groups of cells. These processes include rapid proliferation phases, such as those occurring shortly after zygote formation [25, 168], phases of increased motility and extreme mechanical deformation during gastrulation [65], and phases driven by cellular self-organisation and differentiation during tissue and organ formation [158]. Whether it is the coordination of cell proliferation, migration, differentiation, or a combination thereof, all these processes are examples of cells interacting with each other and behaving collectively, making the right decisions at the right time and in the right place [134].

At later stages, cells must maintain existing tissues and regenerate them when necessary. After injury, cells migrate toward damaged regions and align their proliferation with movement to restore tissue structure [146]. Similarly, tissue homeostasis depends on continuous intercellular communication [39, 211]. Any disruption in the precise control of cell behaviour, whether during homeostatic maintenance or active growth, can lead to diseases like cancer, in which collective cell migration plays a key role [2].

Current advances in experimental techniques, imaging and data processing methodologies allow us to study collective cell behaviour, both *in vivo* and *in vitro*, at an unprecedented temporal and spatial resolution. Having access to such a vast amount of data presents new opportunities to use mechanistic models to test, validate, and refine biological hypotheses. To do this, however, it is essential to consider how we can bet-

ter integrate models with quantitative data, by simplifying current models to enable model calibration and inference while retaining sufficient mechanistic complexity to describe the biological system of interest, or by extending them to include new fundamental interactions. In line with these ideas, this thesis aims to: (i) use and adapt existing mechanistic models to describe and better understand the collective cell behaviour of migrating and proliferating cells; (ii) integrate mathematical models with experimental data through quantitative comparison of models and data, model calibration, and parameter inference; and (iii) explore the extent to which simple models, with a small number of parameters, can accurately describe collective cell behaviour under controlled experimental conditions and predict new behaviour when these conditions are altered. All these objectives share a common key goal: to better understand the *emergence and regulation of collective cell behaviours.*

Focusing on these goals, in the rest of the introduction, we will elaborate on the modelling framework that serves as the mathematical foundation for the studies in this thesis. Based on the idea that the correct coordination of migration and proliferation can give place to collective behaviours, we organise the discussion into different subsections that address how interactions regulate cell motility and cell division rules, respectively.

1.1 Interactions and models of cell movement

We briefly review here a general framework to model the dynamics of interacting cell populations. Given a system of N cells each with mass m, we label the positions of the cell nuclei as $\mathbf{x}_i \in \mathbb{R}^d$, and their velocities by $\mathbf{v}_i \in \mathbb{R}^d$, where typically $d = 1, 2, 3$. In the abscence of noise, Newton's second law for this system reads

$$\begin{cases} \dfrac{\mathrm{d}\mathbf{x}_i}{\mathrm{d}t} & = \mathbf{v_i}\,, \\ m\dfrac{\mathrm{d}\mathbf{v}_i}{\mathrm{d}t} & = \mathbf{f}_i + \mathbf{F}_i\,, \quad \text{for } i = 1, \ldots, N\,, \end{cases} \tag{1.1}$$

where we represent friction and self-propulsion forces by $\mathbf{f_i}$, and cell–cell interactions and other physical interactions with the environment by $\mathbf{F}_i$.

1.1.1 Attraction–repulsion models

For most cases in this thesis, the force $\mathbf{F}_i$ arises from an interaction potential W^N which models soft repulsion between cells at short distances and adhesion at larger distances opposing cell–cell detachment [2, 46]. We would typically write

$$\mathbf{F}_i = -\sum_{j \neq i} \nabla W^N(\mathbf{x}_j - \mathbf{x}_i)\,.$$

Active motion and friction enter through a velocity-dependent term $\mathbf{f}_i$, for example,

$$\mathbf{f}_i = \alpha \mathbf{v}_i - \beta |\mathbf{v_i}|^2 \mathbf{v_i},$$

where α and β parameterise the strength of self-propulsion and friction, respectively. Variants of this model have been studied extensively: it was first introduced in [140], followed by numerical explorations of milling and flocking patterns in [79], derivations of the corresponding kinetic and hydrodynamic descriptions in [48, 142], and analyses of pattern stability in [1].

More generally, agent-based models based on attraction–repulsion are ubiquitous across the sciences. From human crowd dynamics [9] to cell sorting in developmental biology [215], a wide range of systems can be described using long-range attraction — either towards a target or between agents — coupled with short-range repulsion to prevent overcrowding or to model volume exclusion. Originally developed for studying many-body systems, such as gravitational and electrostatic interactions, this framework has found applications in modern fields like biology, social sciences, machine learning, and optimisation. Attraction–repulsion models are typically formulated as systems of ordinary differential equations, and they become computationally expensive when the number of agents is very large, necessitating the use of macroscopic models, which we discuss below. The study of aggregation–diffusion phenomena remains an active field, with ongoing advancements in the analysis, development, and application of mathematical models [10, 105].

Very often, and following the idea that cell inertia is negligible, an overdamped version of Eq. (1.1) is used instead. This well-known approximation can be obtained by considering, for instance, $\mathbf{f}_i = -\zeta \mathbf{v}_i$ for $m/\zeta \ll 1$, and simplifies the above model

to the first-order system

$$\frac{\mathrm{d}\mathbf{x}_i}{\mathrm{d}t} = -\sum_{j\neq i} \nabla W^N(\mathbf{x}_j - \mathbf{x}_i)\,, \quad \text{for } i = 1, \dots, N\,, \tag{1.2}$$

where we absorbed the friction coefficient ζ in the definition of the interaction potential W^N. This approximation assumes that the inertia timescale m/ζ is much shorter than the experimental timescales of interest.

While the assumption of negligible inertia might be accurate for long timescales, and this is indeed what many mathematical models incorporate [2, 46], certain experimental studies of collective cell migration show that cell migration dynamics are underdamped [31, 161, 172]. This is due to the effects of polarisation on cells, which introduces persistence or an effective inertia. This has been noted recently by *Brückner and Broedersz* [30], who discuss on the mathematical similarities between first-order models with cell-polarity interaction mechanisms [24] and second-order models of the form of Eq. (1.1).

1.1.2 Continuum models of interacting cell populations

Although some collective cell behaviours are best described at the single-cell level, cell-based models nevertheless pose several challenges. Namely, they have high computational costs and lack analytic insights for large numbers of cells. Furthermore, advances in experimental techniques and imaging methodologies now allow us to study cell migration in large tissues, where tissue sizes can reach several millimeters, and the trajectories of tens of thousands of cells can be tracked [112]. Although this wealth of data is available, such large numbers also suggest using a more concise description, which is possible through continuum models. Spatially continuous models have the added advantage of involving fewer parameters, facilitating model and parameter inference.

In the limit of large number of cells N, cell dynamics can be described by a density $\rho(\mathbf{x}, t)$, for $\mathbf{x} \in \mathbb{R}^d$ and $t > 0$. Next, we explain how, given a system of cells with positions $\{\mathbf{x}_i(t)\}_{i=1}^N$, we can find a suitable approximation of the empirical cell density,

$$\rho^N(\mathbf{x}, t) = \frac{1}{N}\sum_{i=1}^{N} \delta_{\mathbf{x}_i(t)}\,,$$

where $\delta_{\mathbf{x}_i(t)}$ denotes a Dirac delta centred at $\mathbf{x}_i(t)$. Note that large tissues grow over timescales which are much larger than timescales associated with changes in the speed and direction of individual cells, and hence for long times we may assume overdamped dynamics, as given by Eq. (1.2).

We now take into account the specific shape of the potential W^N and how it scales with the number of cells. The simplest way to to represent cell–cell repulsion at short distances is to assume that, for small distances, W^N approaches a Dirac delta, δ_0, at the origin as $N \to +\infty$. This scaling has been studied rigorously in [166], where the following form of the potential is considered

$$W^N(\mathbf{x}) = \epsilon N^{\beta-1}\psi\left(N^{\beta/d}\mathbf{x}\right) + \frac{1}{N}W(\mathbf{x}),$$

with ψ is a typical repulsive potential with unit volume, and W is an attractive potential which models adhesive forces by opposing cell–cell detachment. The parameter $\epsilon > 0$ measures the relative strength of repulsion to attraction. Under this scaling, and for any $\beta \in (0,1)$, the empirical measure in the limit $N \to +\infty$ can be identified via a mean-field approximation with the solution, $\rho(\mathbf{x},t)$, of the aggregation–diffusion equation

$$\partial_t \rho = \nabla \cdot \left(\rho \nabla \left(\epsilon\rho + W * \rho\right)\right), \tag{1.3}$$

with $(W*\rho)(\mathbf{x},t) = \int W\left(\mathbf{x}-\mathbf{y}\right)\rho(\mathbf{y},t)\,\mathrm{d}\mathbf{y}$.

In the case of two interacting species [56], described by two densities, ρ and η, we can follow the same ideas to obtain

$$\begin{cases} \partial_t \rho = \nabla \cdot \left(\rho \nabla \left(W_{11} * \rho + W_{12} * \eta + \epsilon(\rho+\eta)\right)\right), \\ \partial_t \eta = \nabla \cdot \left(\eta \nabla \left(W_{21} * \rho + W_{22} * \eta + \epsilon(\rho+\eta)\right)\right), \end{cases} \tag{1.4}$$

where W_{11}, W_{22} model homotypic cell–cell adhesion forces, and W_{12}, W_{21} represent the heterotypic or cross-adhesion interactions. Again, the parameter $\epsilon > 0$ gives the strength of cell–cell repulsion and is a measure of contact inhibition of locomotion. Existence of solutions for this system is proven in [97] — see also [73] for the case without cross-diffusion.

1.2 Modelling cell proliferation regulation

We have presented a detailed framework to describe how interactions within cell populations determine cell motility; yet, cell proliferation is also determined by fundamental interactions and specific cell division rules. In particular, the correct regulation of cell proliferation is crucial for the emergence of collective cell behaviour during tissue morphogenesis, homeostasis, and regeneration, and uncontrolled cell division often leads to tumor formation. From a mathematical perspective, these ideas are typically represented in phenomenological models. Such models assume exponential growth at early stages or low cell densities, with saturation effects at high densities due to cell competition and mechanical regulation. Examples include the logistic growth model, where the effective growth rate decreases linearly with cell density, and the Gompertz and Richards' models, which employ different functional forms for the growth rate but exhibit similar qualitative behavior [144, 189]. More generally, for spatiotemporal models of cell migration, we could extend Eq. (1.3) to include cell proliferation,

$$\partial_t \rho = \nabla \cdot (\rho \nabla (\epsilon \rho + W * \rho)) + \rho R(\rho),$$

where the density-dependent growth rate $R(\rho)$ is a non-increasing function of the total cell density ρ. A rigorous particle derivation of the aggregation–diffusion equation coupled with a proliferation term of logistic type is studied in [94].

At a more mechanistic level, experimental studies of cell proliferation often explore how spatial and mechanical constraints within tissues affect different stages of the cell cycle [75, 199, 210], which can be visualised via the widely used FUCCI cell cycle marker [13, 182], which consists of red and green fluorescent proteins that are fused to proteins that are differentially expressed in different cell cycle stages.

Mathematical models of cell cycle dynamics often build on the framework developed by *Vittadello et al.* [213]. In this basic model, a population of cells is divided into different subpopulations, each one corresponding to one or the combination of multiple cell cycle stages, and transition rates between successive cell cycle stages depend on cell–cell interactions and cell cycle regulation mechanisms [90]. The model of *Vittadello et al.* has inspired several studies extending the model to account for realistic cell cycle duration distributions [100, 221]. However, a quantitative exploration of whether

these models can accurately describe and quantify cell cycle dynamics in collectively migrating cells is currently lacking.

1.3 Structure of the thesis and outline

The overall goal of this thesis is to extend and use existing frameworks to model and quantify interactions between different cell populations, and apply this methodology to better understand the emergence and regulation of collective cell behaviours, including applications to tissue engineering and developmental biology. We begin in Chapter 2 by presenting two classical continuum descriptions of epithelial tissue spreading — the Fisher–KPP and Porous–Fisher models — and outlining a Bayesian inference framework for calibrating these models to experimental data. By applying this approach to epithelial monolayer expansion assays [112], we obtain good estimates for all model parameters, and show that both models can be reliably tuned to reproduce the observed tissue growth dynamics.

Following this framework, in Chapter 3 we move to study interactions that directly regulate proliferation. We revisit the same epithelial monolayer expansion experiments of *Heinrich et al.* [112], using spatiotemporal measurements of cellular densities at different cell cycle stages to inform a minimal two-stage continuum model. In this model, cell cycle progression rates are explicit functions of local crowding, and by fitting these density-dependent rates via Bayesian inference [90], we quantify how the durations of each cell-cycle phase lengthen with increasing tissue density — in excellent agreement with recent mechanical-checkpoint studies [75, 199]. Furthermore we provide, using a travelling wave analysis, a quantitative exploration into how tissue crowding regulates cell proliferation patterns in growing epithelia.

Building on the tunable single-tissue and two-stage proliferation models of Chapters 2 and 3, in Chapter 4 we turn to the mechanics of tissue–tissue interactions. Revisiting *Heinrich et al.*'s collision assays [114], we formulate two-species versions of both Fisher–KPP and Porous–Fisher models, and show that only the latter, with its crowding–driven population pressure, reproduces the experimentally observed sharp interfaces and interface kinetics across a wide array of geometries. After verifying that the parameters inferred from isolated expansions qualitatively capture early collision

dynamics, we fully calibrate the Porous–Fisher system to rectangular collision data, accurately predicting boundary displacement under width or density mismatches. Finally, by analysing the resulting pressure gradients and their dependence on the pressure–density exponent, we derive explicit scaling laws for interface motion both without and with proliferation, thus providing a unified, minimal continuum framework for designing and interpreting tissue–tissue collisions and self-assembly experiments

We continue by examining continuum models of cell movement incorporating cell–cell adhesion, which is fundamental for the emergence of collective and cohesive migration. Following the idea that differential adhesion serves as a mechanism for cellular self-organisation during morphogenesis [198, 208], similar mathematical models have been used to understand cell sorting phenomena involved in developmental processes. These include, for instance, stripe formation in zebrafish [216, 217], gangliogenesis in the enteric system [109], the formation of mosaic-like cellular patterns in the olfactory [56, 125] and auditory epithelium [204], and the appearance of columnar structures in the *Drosophila* brain [207]. Although these models have been successful in understanding the underlying biology, they are mathematically complex, reducing analytical insight and limiting the extent to which they can be calibrated and used to describe experimental data. In Chapter 5 we build on a widely used nonlocal model of cell movement and derive, in the limit of short-range interactions, a local continuum model of cell–cell adhesion [87, 88]. The obtained model resembles a system of thin-film equations, where the different model parameters play the role of relative surface tensions between cell populations, allowing us to understand differential adhesion in terms of differential surface tension, as originally phrased by Malcolm Steinberg in the *differential adhesion hypothesis* [198]. Furthermore, we show that a suitable interpretation of the different parameters predicts adhesion-based pattern formation consistent with the differential adhesion hypothesis, which is confirmed via numerical simulations of the model.

While continuum models, combined with experimental data, can provide mechanistic explanations to large-scale tissue dynamics, other collective cell behaviours characterised by smaller spatial scales can be more easily addressed by following the dynamics of each cell. Chapter 6 addresses collective cell migration in confined en-

vironments, studying the effect of curvature and cell-matrix interactions on collective motion. By using agent-based models, similar to the previously mentioned ones (Eq. (1.1)), we explain the emergence of orbiting migration states in confined epithelial spheroids, suggesting that a suitable combination of cell–cell and cell-matrix interactions are needed for cellular self-organisation in confined environments.

The main body of this thesis is based on four research papers (three published and one currently under review). Chapters 2-4 are based on a collaboration with the group of Daniel Cohen (Princeton University). The experimental data have been previously published by *Heinrich et al.* [112, 114]. Chapter 5 is based on theoretical research, and Chapter 6 is based on a collaboration with Ian Wong, Jiwon Kim, Hyuntae Jeong (Brown University) and collaborators. My contribution lies in the development of methodology, modelling, inference, and in assistance with the analysis of the experimental data.

Chapter 2

Continuum models of tissue growth and parameter identifiability

The work in this chapter was published as

"C. Falcó, D. J. Cohen, J. A. Carrillo & R. E. Baker. *Quantifying tissue growth, shape and collision via continuum models and Bayesian inference.* Journal of the Royal Society Interface, 20(204):20230184, 2023" [89],

using data originally published by *Heinrich et al.* [112]. My contributions include data curation, model development and analysis, computational implementation, and writing of the publication. Experimental data used to calibrate the models (Figure 2.1) is available on [113]. Code used to perform the parameter estimation and to solve numerically the models is available on Github.

2.1 Introduction

At the tissue scale, mathematical models are usually based on a continuum description, where the cell density evolves according to a partial differential equation (PDE). Arguably the most famous continuum model of tissue spreading is the reaction-diffusion Fisher–KPP equation [163], which is based on the assumption that cell movement is essentially random, and that cells proliferate according to a logistic growth law. This model, and variants of it, have been used to describe a variety of tissue formation experiments [146, 185, 186].

From a biological perspective, however, the random motion assumption is not very realistic, as cells are able to sense the pressure exerted by neighbouring cells and direct

their movement according to this information [108]. When population pressure is taken into account in continuum models one obtains the Porous–Fisher equation, which replaces the constant diffusivity in the Fisher–KPP equation by a density-dependent function that increases as a power-law of the density. One of the most interesting features about this model is the appearance of compactly supported solutions, which give rise to the sharp invasion fronts observed in tissue formation experiments [28, 56, 80, 81]. Of course, there are additional effects which can play an important role in collective cell motility and have been modelled using extensions of the mentioned equations, such as cell–cell adhesion [6, 74, 87, 162], viscoelastic forces [2, 22, 112], interactions with the extracellular matrix [27, 68, 101], heterogeneity in cell size [126, 127], and cell cycle dynamics [188]. However, integrating even one of these mechanisms inevitably increases model complexity.

Mathematical models can thus be more or less complex depending on the available data and the required level of biological detail, and they are a powerful tool to explore the impact of different biological mechanisms on collective cell movement. So-called *identifiability analysis* methods [64, 66] provide a systematic approach to balance model complexity against the type and amount of experimental data, and are a first step towards the estimation of model parameters. We say that a model is *structurally identifiable* if different parameter values yield different model predictions. Hence, this is an intrinsic property of the model which depends on whether, given infinite ideal data, one can identify single values for the model parameters. Such formal structural identifiability analysis is possible for systems of ordinary differential equations [120, 176], and for certain families of PDEs (e.g. age-structured [177], and linear/semi-linear reaction-diffusion-advection equations [29]), but is more challenging for nonlinear reaction-diffusion equations. Added to this, biological data is never infinite nor ideal which limits how much insight we can gain from structural identifiability.

As a result, here we explore the question of *practical identifiability* [64] of two simple reaction-diffusion continuum models — namely the Fisher–KPP and Porous–Fisher equations — using data from recent tissue formation experiments [112, 114]. Practical identifiability deals with finite and possibly noisy data, and depends on the inference method, but at its core is motivated by the same question: *can we confidently identify es-*

timates for the different model parameters? Here we follow the ideas in [115, 188] and use a Bayesian approach in order to obtain posterior distributions for the different model parameters. Poor identifiability in a Bayesian context is thus associated with very broad posterior distributions indicating high uncertainty for the associated parameters [187]. Our work reveals that both models can be suitably calibrated to reproduce the dynamics of freely expanding epithelia, with the different model parameters being practically identifiable in all considered settings.

We structure the chapter as follows; first, we describe the two continuum models and the inference approach taken. Then, we estimate the different model parameters using comprehensive experimental data of the growth of large, circular epithelia (*Heinrich et al.* [112]).

2.2 Simple models of tissue growth

We start by looking at simple models describing the growth of a single epithelial monolayer tissue. We denote cell density in the tissue by a continuous variable $\rho(\mathbf{x}, t)$ which depends on space $\mathbf{x} \in \mathbb{R}^2$ and time $t \geq 0$. Cell density is assumed to change due to cell movement and local proliferation. Mass conservation implies then that the density ρ satisfies the continuity equation

$$\partial_t \rho + \nabla \cdot \mathbf{j} = r\rho f(\rho), \tag{2.1}$$

where the flux $\mathbf{j}$ determines how cells move, r is the proliferation rate, and $f(\rho)$ is a crowding function which describes how density–dependent effects reduce net growth. For simplicity we consider logistic growth given by $f(\rho) = 1 - \rho/K$, with K a saturation density or carrying capacity. Note that epithelial tissues are well characterised to undergo contact inhibition of proliferation, where cell cycling decreases as cell density increases [112, 173] and hence the logistic growth assumption is reasonable — see also [189] for other possibilities.

A very simple model can be motivated by assuming that cells move randomly following Brownian motion, which corresponds to the well-known Fick's law of diffusion, $\mathbf{j} = -D\nabla\rho$. In this case, we obtain the Fisher–KPP equation

$$\partial_t \rho = D\Delta\rho + r\rho\left(1 - \frac{\rho}{K}\right). \tag{2.2}$$

This model and related ones are particularly relevant to describe tissue growth due to the presence of travelling wave solutions — with suitable initial conditions — which are characterised by an invasion front of fixed shape that propagates at a constant speed [163]. While travelling waves are well understood in one spatial dimension, they are less studied in two. Following [141], we sketch a simple argument indicating that the wave speed in two dimensions approaches the classical one-dimensional value. In polar coordinates and under radial symmetry (with radial coordinate $\mathfrak{r}$), Eq. (2.2) becomes

$$\partial_t \rho = D\left(\partial_{\mathfrak{r}}^2 \rho + \frac{1}{\mathfrak{r}}\partial_{\mathfrak{r}}\rho\right) + r\,\rho\left(1 - \frac{\rho}{K}\right).$$

Let $R(t)$ denote the front position. Note that near the front, $\mathfrak{r} \approx R(t)$, so the curvature term is small: $\mathfrak{r}^{-1}\partial_{\mathfrak{r}}\rho \sim R(t)^{-1}\partial_{\mathfrak{r}}\rho$. Passing to a travelling wave coordinate $\xi = \mathfrak{r} - c(R)\,t$, where $c(R)$ is the local speed of a cylindrical front of radius R, yields

$$D\,\frac{\mathrm{d}^2\rho}{\mathrm{d}\xi^2} + \left(c(R) + \frac{D}{R(t)}\right)\frac{\mathrm{d}\rho}{\mathrm{d}\xi} + r\,\rho\left(1 - \frac{\rho}{K}\right) = 0.$$

This coincides with the one-dimensional travelling wave equation if

$$c(R) = c - \frac{D}{R(t)},$$

with c the one-dimensional wave speed. Hence a radially symmetric front accelerates toward c as $R(t) \to \infty$, remaining bounded above by c. As in the Fisher–KPP case, a linear analysis at the leading edge determines the minimal admissible speed, giving the classical bound

$$c \geq c_{\mathrm{FKPP}} := 2\sqrt{Dr}\,.$$

A more realistic model, however, should account for the fact that cell movement is not completely random and can be influenced by the local cell density. A standard approach in order to incorporate crowding effects into Eq. (2.1) results from the assumption that the velocity is proportional to the gradient of the density, so that cells move down population density gradients. In other words, we write the flux as $\mathbf{j} = \rho\mathbf{v}$, where $\mathbf{v}$ represents the cell velocity and now assume that $\mathbf{v} = -D\nabla\rho$. This gives the following Porous–Fisher equation

$$\partial_t \rho = D\,\nabla\cdot(\rho\nabla\rho) + r\rho\left(1 - \frac{\rho}{K}\right). \tag{2.3}$$

When there is no proliferation ($r = 0$), Eq. (2.3) corresponds to a specific case of the well-known porous–medium equation [212]. This equation is also related to Darcy's law which links the velocity with the population pressure: $\mathbf{v} = -\nabla P(\rho)$. For the general porous-medium equation, pressure and density are related via the power-law function $P(\rho) \sim \rho^{m-1}$, depending on the exponent m. During this work and unless stated otherwise, we will assume $m = 2$. Note that in the limit $m \to 1$, one obtains the linear diffusion case with $P(\rho) \sim \log \rho$.

Here, by population pressure we mean a constitutive, density-dependent driver of cell motility only — an analogue of physical pressure that encodes crowding constraints. It biases motion from densely populated regions toward less crowded ones, thereby promoting spreading from high to low density zones. We use it solely to modulate cell motility, and not proliferation or other cellular processes.

From a microscopic point of view, where one focuses on individual cell trajectories, Eq. (2.2) corresponds to the continuum limit of a system of non-interacting agents which move randomly and can proliferate with a density–dependent probability. The porous–medium equation with $m = 2$ can also be derived from microscopic movement rules when one takes into account volume exclusion [41, 98, 169], starting from on-lattice [15, 86] and also from off-lattice agent-based models [56, 77, 78, 166]. Further, the case with $m = 3$ can be identified as the mean-field limit of a system of interacting agents with a particular diffusive scaling [220] and has also been suggested as *the simplest model* to relate the dispersal velocity to both the density and its gradient [205].

Eq. (2.3) also predicts travelling waves in two spatial dimensions. In radial geometry, curvature contributes only an $O(1/R(t))$ correction near a front of radius $R(t)$, so the radial wave speed accelerates toward the one-dimensional speed. Because diffusion is degenerate at $\rho = 0$, these fronts are sharp, and their speed is selected by nonlinear dynamics rather than linear leading-edge behaviour. The one-dimensional equation in travelling wave coordinates reads

$$D \frac{\mathrm{d}}{\mathrm{d}\xi}\left(\rho \frac{\mathrm{d}\rho}{\mathrm{d}\xi}\right) + c\frac{\mathrm{d}\rho}{\mathrm{d}\xi} + r\,\rho\left(1 - \frac{\rho}{K}\right) = 0.$$

In this case, the explicit solution can be found explicitly by using the ansatz $\mathrm{d}\rho/\mathrm{d}\xi =$

$-a(1-\rho/K)$ [163], which solves exactly the equation above if $a^2 = Kr/2D$ and

$$c = c_{\mathrm{PF}} := \sqrt{\frac{DKr}{2}}\,.$$

In the following, we connect Eqs. (2.2) and (2.3) with data from recent experiments studying the dynamics of expanding and colliding epithelial monolayer tissues. The two suggested models are solved numerically in two spatial dimensions with the finite-volume numerical scheme described in [12, 45] — see Appendix A.

2.3 Tissue expansion experiments and parameter estimation

In order to calibrate the two suggested models, we focus on the experiments by [112]. In these, *Heinrich et al.* characterised the expansion dynamics and growth of single circular epithelial tissues using an MDCK cell line[1]. Initially, cells are cultured in a silicone stencil for 18 hours and, after the stencil removal, tissues are allowed to freely expand for 46 hours, which enables for each cell to undergo 2-3 cell divisions given that the cell cycle duration is around 16 hours. Local densities are then quantified by counting the number of nucleus centroids — for more details we refer to [112]. For our analysis, we only consider the measured cell densities after the first six hours of the experiment so that effects caused by the stencil removal are negligible. In Figure 2.1A we show snapshots from one such experiment using a circular tissue with initial diameter of 3.4 mm — see Figure 2.1B for the quantified densities. The radial density profile resulting from averaging 11 experimental replicates is shown in Figure 2.1C. Datasets used to reproduce these figures were taken from [113]. See Figure 2.2 for individual density profiles at specific time points.

Our experimental data then consists of many individual measurements of the cell densities $\rho(\mathbf{x}, t)$, giving rise to the dataset $\mathcal{D} = \{\rho^{\mathcal{D}}(\mathbf{x}_i, t_j)\}_{i,j}$. Here, the different measurements are recorded every 20 minutes, while the positions $\{\mathbf{x}_i\}_i$ correspond to the

[1]MDCK (Madin–Darby Canine Kidney) epithelial cells are a standard model for collective cell migration because they form cohesive monolayers with strong junctions and exhibit robust, coherent tissue-level motion.°

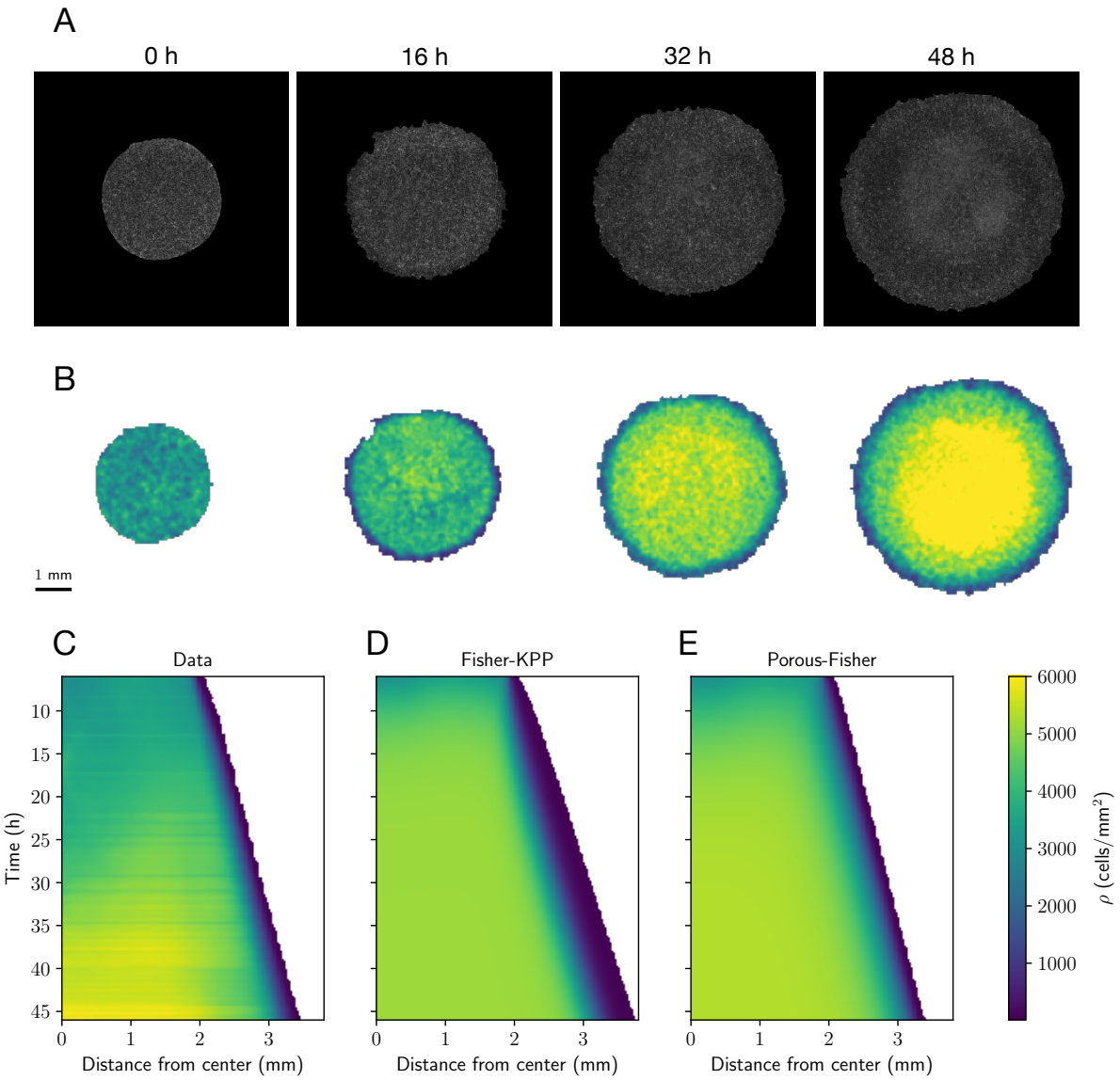


Figure 2.1: Expansions of single tissues and model predictions. (A) Microscopy images in phase-contrast at different times for the expansion of a circular tissue with initial diameter of 3.4 mm — taken from [113]. (B) Quantified experimental cell densities for the same expansion — data from [113]. (C) Experimental radial density profile obtained after averaging the expansions of 11 tissues with the same initial condition — from [112]. (D) Radial density profile from the Fisher–KPP model given by Eq. (2.2). (E) Radial density profile from the Porous–Fisher model given by Eq. (2.3). Model parameters correspond to the maximum posterior estimates. All densities thresholded at 10 cells/mm^2 for a better comparison between the Fisher–KPP and the Porous–Fisher model. See Figure 2.2 for individual density profiles at specific time points

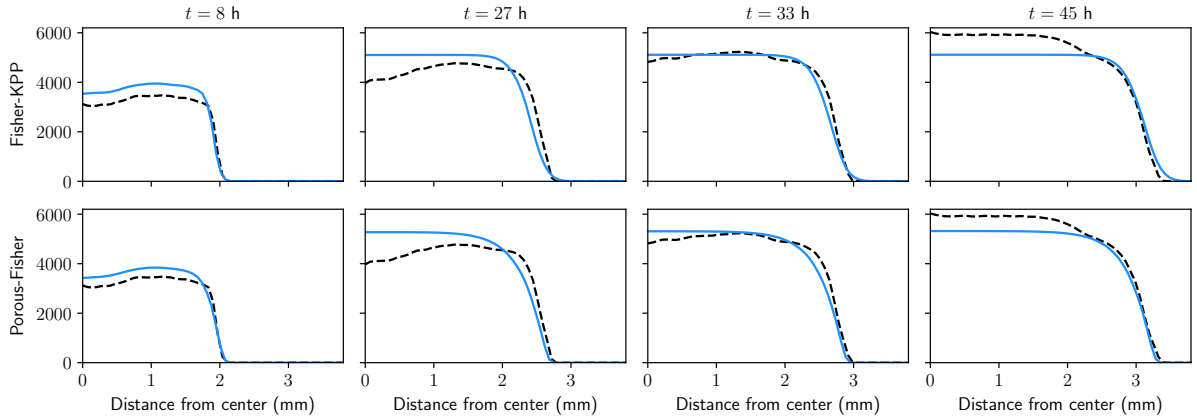


Figure 2.2: Comparing data and model predictions for tissue growth experiments [112] — individual snapshots corresponding to Figure 2.1. Top row shows predictions from the Fisher–KPP (linear diffusion) model; bottom row shows predictions from the Porous–Fisher (nonlinear diffusion) model. Blue lines represent numerical simulations using the maximum likelihood estimate for the model parameters. Black dashed lines represent the experimental density profiles.

centers of small voxels of 115×115 μm^2. In practice, and in order to keep the dimensionality of the data sufficiently low, we will only use the densities corresponding to the time points $t_j = 16, 26, 36, 46$ h. Moreover, and in order to minimise the effect of the stencil removal, the model takes as initial condition the quantified density profile six hours after stencil removal. In order to connect experimental data and models, we assume that the observations $\rho^{\mathcal{D}}$ are noisy versions of the model predicted density ρ. A common approach in mathematical biology [115, 188] is to impose that the observation errors are additive, independent and normally distributed with variance σ^2. In other words, we assume the following error model

$$\rho^{\mathcal{D}}(\mathbf{x}_i, t_j) = \rho(\mathbf{x}_i, t_j) + \varepsilon_{ij}, \quad \varepsilon_{ij} \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2). \tag{2.4}$$

While this is a commonly employed approach, it may be overly simplistic to assume the error model in Eq. (2.4). For example, model misspecification is likely to introduce spatial and temporal correlations [136]. Recent work advocates for a physically motivated multinomial measurement error model for count data [190, 145]. A more comprehensive quantification of the noise is left as a subject for future investigation.

2.3.1 Parameter estimation via maximum likelihood

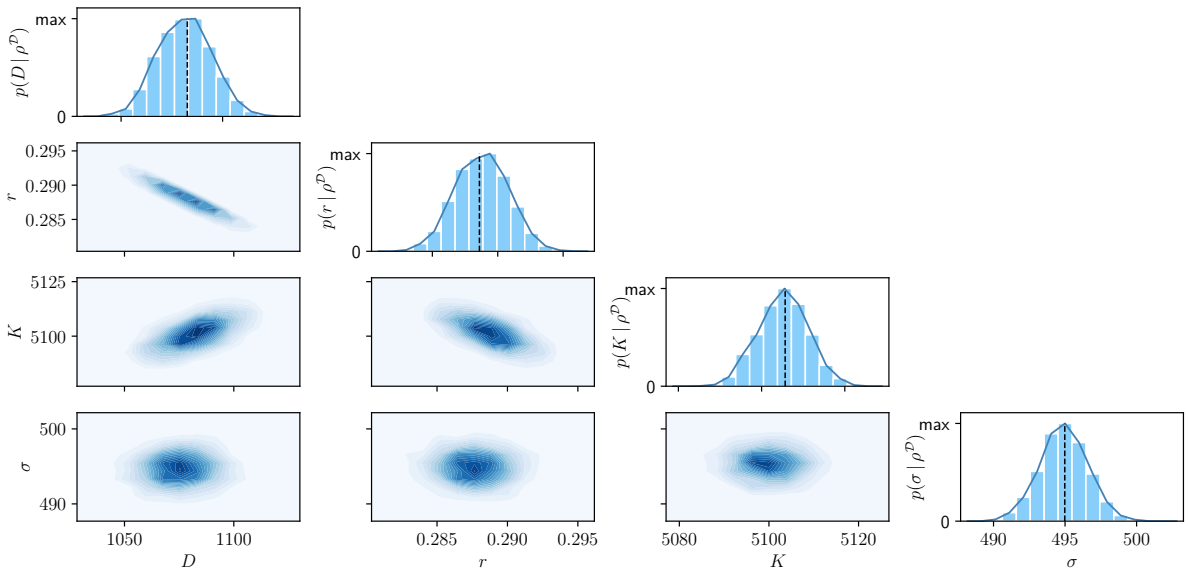


Figure 2.3: Results of the MCMC algorithm for the Fisher–KPP model given by Eq. (2.2). The diagonal plots represent the univariate marginal posterior distributions for each parameter. Below the diagonal we show the bivariate densities for every combination of parameters. Univariate posterior modes correspond to $(D, r, K, \sigma) = (1073 \pm 13\ \mu\text{m}^2/\text{h}, 0.289 \pm 0.002\ \text{h}^{-1}, 5113 \pm 6\ \text{cells/mm}^2, 492 \pm 2\ \text{cells/mm}^2)$, where the errors are given by one standard deviation, calculated from the posterior distributions. Black dashed lines indicate the maximum likelihood estimates for each parameter.

Both models — Eqs. (2.2) and (2.3) — have three parameters D, r, K to be estimated. Considering the variance of the observation error as an extra parameter, we can write them as a vector $\theta = (D, r, K, \sigma)$. With the error model given by Eq. (2.4) we can explicitly write the log-likelihood of observing the measured data as

$$\ell(\theta) = -\frac{1}{2}\sum_{i,j}\left(\log\left(2\pi\sigma^2\right) + \left(\frac{\rho(\mathbf{x_i}, t_j) - \rho^{\mathcal{D}}(\mathbf{x_i}, t_j)}{\sigma}\right)^2\right). \tag{2.5}$$

A direct approach to estimating the parameters in the two models consists of maximixing this log-likelihood as a function of the parameter vector θ, which gives a maximum likelihood estimator of the model parameters: $\theta_{\text{ML}} = \text{argmax}_\theta\, \ell(\theta)$. In the case of a fixed noise parameter σ, this is equivalent to minimizing the mean squared error between model and data. Note, however, that whenever a model is non-identifiable,

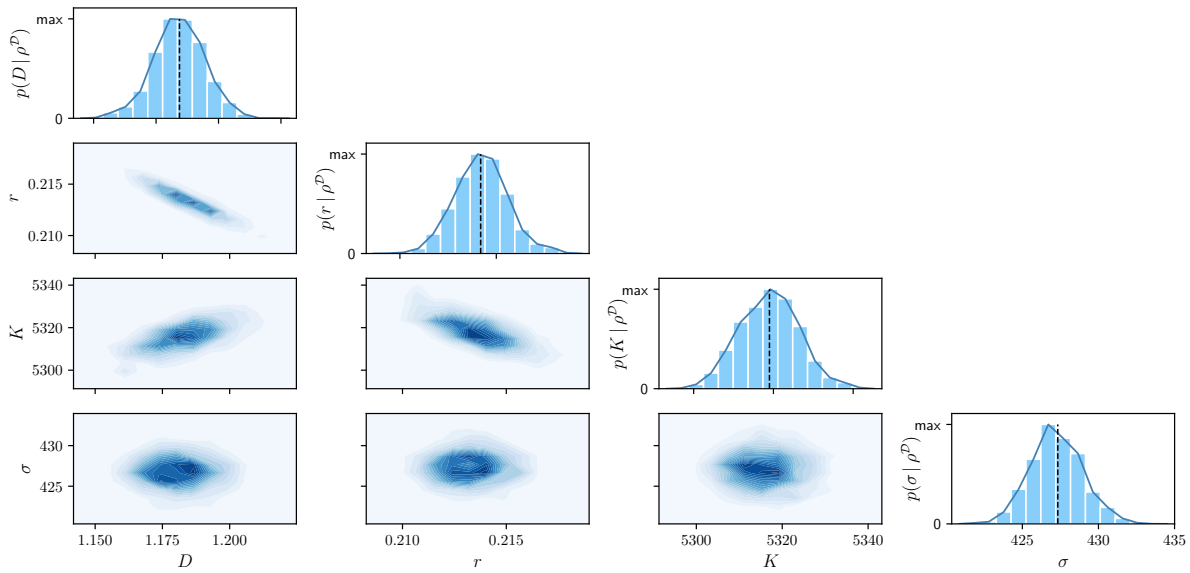


Figure 2.4: Results of the MCMC algorithm for the Porous–Fisher model given by Eq. (2.3). The diagonal plots represent the univariate marginal posterior distributions for each parameter. Below the diagonal we show the bivariate densities for every combination of parameters. Univariate posterior modes correspond to $(D, r, K, \sigma) = (1.18 \pm 0.01\ \mu\text{m}^2/(\text{cells}{\cdot}\text{h}), 0.214 \pm 0.001\ \text{h}^{-1}, 5319 \pm 7\ \text{cells}/\text{mm}^2, 427 \pm 2\ \text{cells}/\text{mm}^2)$, where the errors are given by one standard deviation, calculated from the posterior distributions. Black dashed lines indicate the maximum likelihood estimates for each parameter.

maximising the likelihood might lead to misleading results [187]. This is thus only a first step in our parameter inference analysis.

We perform the likelihood optimisation using the pyPESTO parameter inference toolbox [184]. This toolbox allows for local optimisation of the likelihood starting from an initial guess of θ_0. By randomly sampling a large number of initial vectors θ_0 we find the same local maximum in most of the optimisation runs. Additionally, this local maximum also maximises the likelihood among all the found local maxima. In order to generate initial guesses of θ_0 we sampled uniformly on log-scale using the parameter bounds $10^{-2.5} < r < 10^{1}\ \text{h}^{-1}$, $10^{3} < K < 10^{3.5}\ \text{cells}/\text{mm}^2$, $10^{1} < \sigma < 10^{3.5}\ \text{cells}/\text{mm}^2$, for both models; and $10^{2.5} < D < 10^{4.5}\ \mu\text{m}^2/\text{h}$ for the Fisher–KPP model, and $10^{-1.5} < D < 10^{-2.5}\ \mu\text{m}^2/(\text{cells}{\cdot}\text{h})$, for the Porous–Fisher model. The maximum likelihood estimators are indicated using dashed lines in Figures 2.3 and 2.4 for the Fisher–KPP and Porous–Fisher models, respectively.

As stated earlier, only the experimental cell densities corresponding to the time

points $t_j = 16, 26, 36, 46$ h were used for the likelihood calculation in Eq. (2.5). This was done to keep the computational costs of computing the maximum likelihood estimate at reasonable levels. Different choices of these time points yielded similar results for the maximum likelihood estimate θ_{ML}.

An alternative strategy would be to fit the model sequentially between successive frames, possibly at a finer temporal resolution, and treat the diffusion coefficient and proliferation rate as time-dependent parameters. This frame-to-frame calibration could capture transient changes in cell motility and division — particularly those induced by stencil removal — thereby alleviating the early-time overestimation of densities and the late-time underestimation observed with a single fit. A caveat is the estimation of the saturation density K, which is best informed by higher-density data [218]. This could be addressed by first estimating K from later time points (when densities are closer to saturation) and then fixing or constraining K while inferring D and r over earlier intervals.

2.3.2 Bayesian inference

Next, we explore the question of *practical identifiability* of the Fisher–KPP and Porous–Fisher models using a Bayesian approach. In order to capture uncertainty in the model parameters, we estimate the posterior distribution $P(\theta \mid \rho^{\mathcal{D}})$, which can be calculated from Bayes' theorem

$$P(\theta \mid \rho^{\mathcal{D}}) \propto P(\rho^{\mathcal{D}} \mid \theta)\pi(\theta),$$

where $P(\rho^{\mathcal{D}} \mid \theta)$ is the likelihood of observing the measured data given parameters θ, and $\pi(\theta)$ is the prior distribution of the parameter vector θ. We assume the error model given by Eq. (2.4), and hence the log-likelihood is given by Eq. (2.5). The priors for the two considered models are assumed to be uniform on log-scale using the bounds given in the previous section.

In order to infer the posterior distribution, we use a Metropolis-Hastings MCMC (Markov chain Monte Carlo) sampler with adaptive proposal covariance, which is also implemented in pyPESTO [184]. The Metropolis-Hastings MCMC algorithm is a simple and popular choice for exploring the parameter space [115, 188], in which a Markov Chain starts at position θ, and accepts a potential move to θ^* with probability

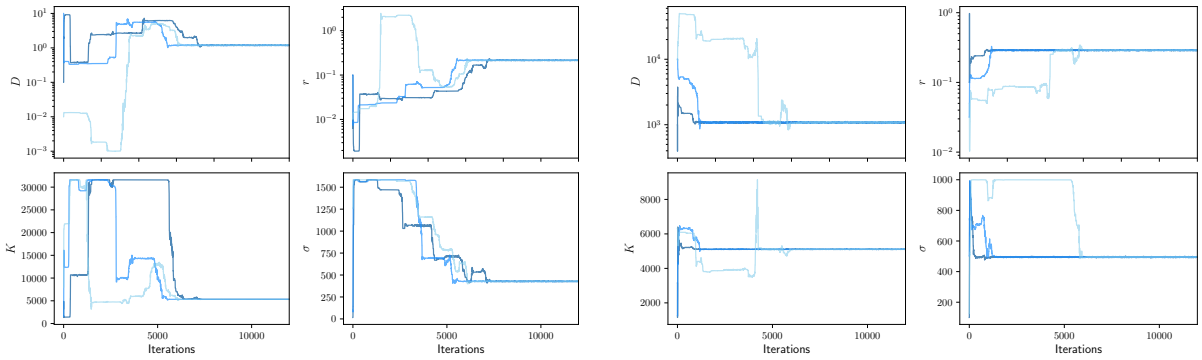


Figure 2.5: Typical Markov chain iterations of length 12,000 for the linear diffusion Fisher–KPP model (left) and the nonlinear diffusion Porous–Fisher model (right), showing the parameters D, r, K, and σ. For the linear model, the chains are initiated with $(D, r, K, \sigma) = (10^3, 0.1, 3200, 100)$, $(10^4, 0.03, 1300, 500)$, and $(10^4, 0.05, 6300, 130)$, respectively. The maximum univariate Gelman–Rubin diagnostic among the four parameters satisfies $\hat{R} < 1.01$ using the last 5000 chain iterations. For the nonlinear model, the chains are initiated with $(D, r, K, \sigma) = (10^{-1}, 0.1, 3200, 160)$, $(10, 0.1, 1600, 100)$, and $(10^{-2}, 0.02, 8000, 120)$, respectively, with $\hat{R} < 1.03$ for the final 5000 iterations.

$q = \min\{1, P(\theta^* \mid \rho^{\mathcal{D}}) / P(\theta \mid \rho^{\mathcal{D}})\}$. In this way, the Markov chain tends to move towards high values of the posterior distribution, while still allowing for transitions to regions of lower probability in order to move away from local maxima. In this context, poor identifiability of the parameters can be detected by Markov chains that fail to converge towards a unimodal peaked posterior distribution.

We run the MCMC algorithm starting from three different initial guesses of θ for both models. In all cases, the Markov chains converge rapidly to narrow and well-defined stationary distributions — see Figure 2.5 for plots of the chains and the univariate Gelman-Rubin convergence diagnostics. In particular, our typical Markov chain iterations are of length 12000. Taking the last 5000 iterations of the three chains in each model we obtain the posterior distributions $P(\theta \mid \rho^{\mathcal{D}})$. In Figures 2.3 (Fisher–KPP) and 2.4 (Porous–Fisher) we show a plot matrix representation of the univariate and bivariate marginal distributions, with unimodal and approximately symmetric univariate densities. We also observe an excellent agreement between the marginal univariate modes and the maximum likelihood estimates found in the last section. Note that for the two models, different combinations of the parameters D, r, K can result in the same invasion front speed, which explains the observed correlation between these parameters in the bivariate densities in Figures 2.3 and 2.4. However, we observe that there is

only one set of parameters maximising the likelihood, and that these parameters can be confidently identified given the small variance of the posterior distribution.

All identified parameters lie within the biologically feasible bounds. In the linear diffusion case (Fisher–KPP), the univariate modes are given by $(D, r, K, \sigma) = (1073\ \mu\text{m}^2/\text{h}, 0.29\ \text{h}^{-1}, 5113\ \text{cells}/\text{mm}^2, 492\ \text{cells}/\text{mm}^2)$. Using an average density of ~ 3000 cells/mm^2 the estimated proliferation rate is around $\sim 0.1\ \text{h}^{-1}$, which yields an estimated division time around 10 hours. This is consistent with the characteristic division time for MDCK cells of 16-18 hours given that this timescale can vary significantly with cell size [199]. The carrying capacity can also be related to the typical cell radius for MDCK cells. Although notable variability has been reported [223], the MDCK cell radius a is estimated to oscillate between $a \sim 6\ \mu$m and $a \sim 18\ \mu$m [99]. Assuming that maximum densities in the monolayer are associated with hexagonal close packing of cells, the maximum theoretical density is given by $K = 1/(2\sqrt{3}a^2)$ [213]. With our estimated carrying capacity this yields an estimate of $a \sim 8\ \mu$m, which again is consistent with previous measurements, at least for cells in the bulk of the tissue.

In the case of the Porous–Fisher model, we obtain the univariate modes $(D, r, K, \sigma) = (1.18\ \mu\text{m}^2/(\text{cells}\cdot\text{h}), 0.21\ \text{h}^{-1}, 5319\ \text{cells}/\text{mm}^2, 427\ \text{cells}/\text{mm}^2)$. Note that the proliferation related parameters r, K are very similar to the ones we estimated for the Fisher–KPP model. In this case, we estimate a cell divison time around ~ 11 hours, and a typical cell radius of $a \sim 7\ \mu$m, again within the known ranges. Note that for the Porous–Fisher model, the diffusion coefficient is density–dependent — $D(\rho) = D\rho$. Using an average density of ~ 3000 cells/mm^2, we also estimate an *average diffusion coefficient* which is three times larger than in the linear case. This larger average diffusion can be explained by accounting for the slower wave speed of the Porous–Fisher model when compared to that of the Fisher–KPP model — see Section 2.2. We also observe that the estimated noise related parameter σ is smaller in the Porous–Fisher case.

In summary, both models present well-defined and narrow posterior distributions for all the model parameters, with the parameter estimates being consistent with previous experimental measurements. Thus, we have shown via a Bayesian approach that all the model parameters appear to be identifiable. A more sophisticated approach aiming to use all the available data — instead of measurements every 10 hours — could

include for instance a mini-batch algorithm [149]. However, taking a subset of the data highlights that the models are practically identifiable, suggesting such approaches are not necessary in this case.

2.3.3 Almost identical predictions from different continuum models

Next, we explore to what extent the two considered models are able to reproduce the observed data. To do so, we solve numerically Eqs. (2.2) and (2.3) using the posterior modes that we estimated in the previous section. In order to minimise the possible impact of the stencil removal on cell motility [112], we use as initial condition the experimental density profile at time $t = 6$ h. The resulting radial density profiles are shown in Figure 2.1C-D — see also Figure 2.2.

First, we observe that both models yield very similar predictions with minor differences that are only noticeable near the expansion front. This is basically due to the fact that the solution of the Porous–Fisher model (2.3) presents a sharp front, in contrast with the exponential decay in space of the Fisher–KPP equation (2.2). Note that the Fisher–KPP model fails to accurately capture the behaviour of cell densities near the monolayer boundary, but the Porous–Fisher model, which accounts for population pressure, gives a more accurate description.

Secondly, we see that both models capture qualitatively the dynamics and growth of the expansions, but fail to capture the non-monotonic behaviour of the radial density profile for intermediate timescales. The experiments of *Heinrich et al.* [112] observed that this phenomenon is accentuated for smaller tissues. Moreover, for later times, the experiments report cell densities that are higher than the estimated carrying capacity. A quantification of the experimental density profile for longer timescales could help in estimating the carrying capacity. Moreover, we emphasize that both our models are minimal in the sense that they assume cell movement follows very simple rules. Accounting for cell–cell adhesion, which is known to play an important role in epithelial dynamics [112], or even for size variability as cells progress through the cell cycle [126, 154], could yield more accurate results.

In principle, we could quantify data stochasticity by generating posterior predictive trajectories — sampling parameters from their posterior distributions and adding

independent Gaussian observation noise. We chose not to do so here for three reasons. First, the posterior distributions are very narrow, so parameter uncertainty contributes negligibly to the predictive spread. Second, the Gaussian error model was adopted for likelihood tractability. For continuum models/data we expect spatial and temporal correlations in the residuals, so the associated uncertainty bands would likely misrepresent the true variability. Third, the error parameter σ is relatively small compared to the experimental densities and our datasets are averages over multiple experimental realisations, further reducing the impact of observation noise. For clarity of presentation, we therefore report deterministic simulations at the posterior modes and leave correlated, spatiotemporal noise modelling for future investigation.

All in all, these results show that both models, after being suitably calibrated, can explain equally well the data. Indeed, after evaluating the log-likelihood for both models at $\theta = \theta_{\mathrm{ML}}$, we did not report any significant difference. As we will see in Chapter 4, only under sufficiently complex experimental conditions — when a detailed physical description is required — can we distinguish between the models.

2.4 Discussion

In this chapter, we have focused on two main aspects of tissue growth modelling: the practical identifiability of the Fisher–KPP and Porous–Fisher models using a Bayesian approach. Using data from recent experiments studying the growth and expansion of single epithelial sheets [112], we were able to obtain well-defined posterior distributions for each of the model parameters with relatively narrow confidence intervals. Our work thus adds to a growing literature assessing the practical identifiability of similar models under a variety of different experimental conditions [28, 188].

In contrast with previous studies, and for the sake of conciseness, here we opted for using only a Bayesian MCMC approach. Another commonly used option is the profile likelihood method [188, 189], which requires the solution of an optimisation problem. This method, however, can yield similar results to the MCMC algorithm and significantly reduce computational time. Although the Bayesian method can be very helpful in performing uncertainty quantification, we believe that studies comparing a larger number of models may benefit from a likelihood-based approach.

Chapter 3

Quantifying cell cycle regulation by tissue crowding

In this chapter, we delve deeper into the tissue growth experiments by *Heinrich et al.* to study cell cycle dynamics during cell migration using cell cycle fluorescence markers. Building on the previous chapter, where we showed that the Fisher–KPP model captures the dynamics of expanding epithelial monolayers, we extend this model to account for cell cycle stages and their regulation. The work in this chapter is published as

"C. Falcó, D. J. Cohen, J. A. Carrillo & R. E. Baker. *Quantifying cell cycle regulation by tissue crowding.* Biophysical Journal, 124(6):923–932, 2025" [90],

using data originally published by *Heinrich et al.* [112], and *Simpson et al.* [188]. My contributions include model development and analysis, data curation, computational implementation, and writing of the publication. Code to solve the model and to perform the parameter estimation is available on Github. Data used to calibrate the model can also be found on Github and in [112].

3.1 Introduction

The coordination of cell proliferation across space and time is crucial for the emergence of collective cell migration, which plays a fundamental role in development, including tissue formation and morphogenesis, and also at later stages for tissue regeneration and homeostasis. Cells adapt their division rates in response to mechanical constraints within tissues [124, 199], allowing cell populations to self-organise and eventually form

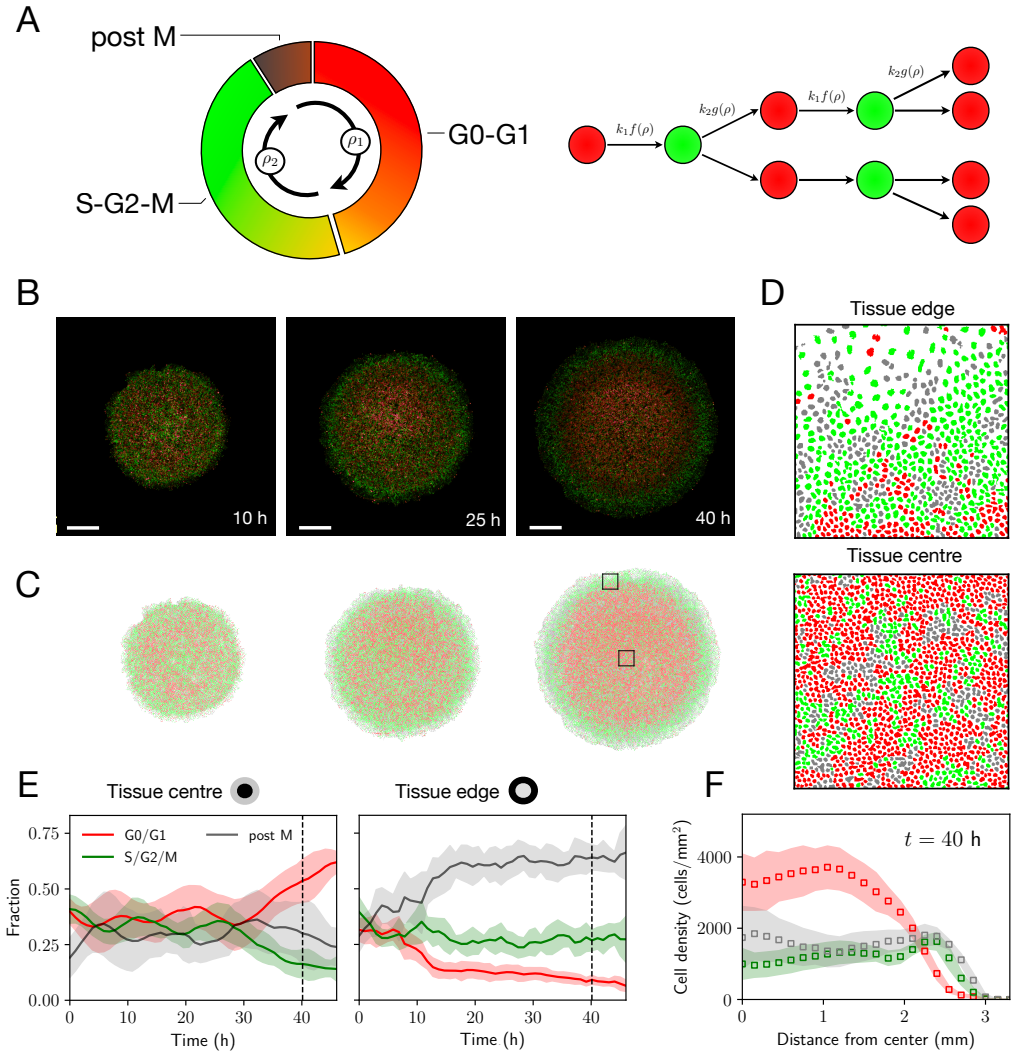


Figure 3.1: (A) Schematics of the FUCCI cell cycle marker system and model conceptualisation. (B) FUCCI fluorescence images from the experiments of *Heinrich et al.* [112] at different time points (adapted). Initial tissue diameter $\sim$ 3.4 mm. Scale bars correspond to 1 mm. (C) Segmented data showing G1 (red), S/G2/M (green), and post–mitotic (grey) cells. (D) Zoomed-in segmented data at the tissue edge and centre, corresponding to the black squares in (C). (E) Fraction of cell cycle state cells in the tissue centre and in the tissue edge — defined as regions extending $\sim$ 200 μm from the tissue center and tissue edge, respectively. (F) Density profiles in polar coordinates at $t = 40$ h, showing cells in G1, S/G2/M, and post–mitotic cells. (E) and (F) show the average of 11 independent tissue expansions with the same experimental initial condition, with shaded regions indicating one standard deviation with respect to the mean.

and maintain tissues and complex structures. Moreover, disruptions in the control of cell proliferation often result in tumour formation [153, 156, 167]. Although significant experimental efforts have been devoted to understand the mechanical regulation of

cell proliferation [107] and its interplay with collective cell migration, existing mathematical models have failed to describe these constraints and how they affect cell cycle progression [100, 188, 213].

In order to understand cell proliferation regulation, numerous experimental studies have explored how spatial and mechanical constraints within tissues affect different stages of the cell cycle. The cell cycle consists of four main stages, namely: the G1 phase, where cells grow and prepare for DNA replication; the S phase, during which DNA synthesis occurs; the G2 phase, characterised by further cell growth and preparation for mitosis; and finally, the M phase, where cell division takes place. Cells can also exit the cell cycle and enter G0, where they become quiescent. The experimental visualisation of cell cycle stages can be achieved via the widely used FUCCI cell cycle marker [182], which consists of red and green fluorescent proteins that are fused to proteins Cdt1 and Geminin, respectively. Cdt1 exhibits elevated levels during the G0/G1 phase and decreased levels throughout the remaining cell cycle stages, whereas Geminin shows high expression during the S, G2, and M phases; allowing thus to distinguish between these different stages — see Figure 3.1. Several extensions of the FUCCI system exist now [178]; for instance FUCCI4 allows for the simultaneous visualisation of the G1, S, G2, and M phases [13].

Experimental studies of cell migration are often performed in epithelia due to their strong cell-cell adhesion which gives rise to collective and cohesive motion. Moreover, they play a fundamental role in multicellular organisms as they serve as protective layers for various body surfaces and organs. Epithelial cell proliferation is regulated by mechanical forces, which can accelerate, delay, arrest, or re-activate the cell cycle. In particular, extensive research has focused on the G1-S boundary, revealing that intercellular tension can favour this transition [210], while tissue pressure can halt progression based on crowding [199].

More generally, the extracellular regulation of switches from G0 and G1, and within substages of G1, has been well-known for many years [170]. However, and contrary to initial assumptions, cells also have the ability to regulate progression through stages of the cell cycle following the G1-S transition in response to external cues. These external signals might involve not only mechanical forces [102], but also nutrients and growth

factors [58, 157]. In epithelia, this question was explored recently by *Donker et al.* [75], revealing a mechanical checkpoint in G2 which controls cell division. In particular, this checkpoint allows cells to regulate progression through G2, via sensing of local density, explaining why dense regions in epithelia contain groups of cells that are temporarily halted in G2.

Experimental studies employing FUCCI and variations of it have thus successfully linked mechanical constraints to cell cycle progression. These studies have employed qualitative analysis, direct measurements of cell cycle stage durations [75], or metrics associated with cell cycle progression, such as cell area [199], and Geminin/Cdt1 or EdU signals [118, 200]. However, these approaches omit a quantitative comparison between model and data, hence limiting the depth of mechanistic insights that can be derived.

Here, we present a quantitative investigation into the density–dependent regulation of cell cycle progression by sensing of local tissue density. First, we construct a mathematical model of cell cycle dynamics that accurately captures the impact of tissue crowding on cell cycle progression. By combining minimal mathematical modelling, Bayesian inference, and recent experimental data [112], we provide further evidence, consistent with previous experimental studies [75, 199], that density–dependent effects operate throughout the cell cycle and together serve as a regulating mechanism for the growth of epithelial tissues. Our work thus constitutes a systematic approach towards the quantification of density–dependent effects regulating cell cycle progression. Moreover, the obtained parameter estimates reveal an explicit relation between the duration of different cell cycle stages and tissue density, which is consistent with the experimental measurements of *Donker et al.* [75].

3.2 Mathematical models of cell cycle dynamics

We build on the model proposed by *Vittadello et al.* [213] to describe two cell populations, $\rho_1(\mathbf{x}, t)$ and $\rho_2(\mathbf{x}, t)$, in different stages of the cell cycle, for $\mathbf{x} \in \mathbb{R}^d$, and $t \geq 0$. We represent by ρ_1 the density of cells that are in G0/G1, while ρ_2 gives the density of cells in the S/G2/M phases of the cell cycle — see Figure 3.1. In the model, cell motility is described via linear diffusion, with a diffusion constant $D > 0$ for both cell

populations [214]. In order to effectively capture density–dependent effects controlling cell cycle progression, we assume that the transitions between different cell cycle stages are regulated by two *crowding functions*, $f(\rho)$ and $g(\rho)$, which depend on the total cell density $\rho = \rho_1 + \rho_2$. In particular, the transition rate from G1 to S is given by $k_1 f(\rho)$, while the division rate (from S/G2/M to G1) is given by $k_2 g(\rho)$, where $k_1, k_2 > 0$ are intrinsic rates of cell cycle progression. With this, the model reads

$$\begin{cases} \partial_t \rho_1 &= D\Delta\rho_1 - k_1\rho_1 f(\rho) + 2k_2\rho_2 g(\rho), \\ \partial_t \rho_2 &= D\Delta\rho_2 + k_1\rho_1 f(\rho) - k_2\rho_2 g(\rho), \end{cases} \tag{3.1}$$

where the factor of two in the equation for ρ_1 represents cell division into two daughter cells, and $\Delta = \sum_{i=1}^{d} \partial_{x_i}^2$ is the Laplacian operator in dimension d. These equations are solved first in polar coordinates (assuming radial symmetry in two spatial dimensions, $d = 2$) to describe epithelial tissue expansion experiments, and then in one spatial dimension ($d = 1$) to study scratch assay experiments [188], and travelling wave behaviour — see Appendix A for details on the numerical implementation. In all cases, we solve the model using no-flux boundary conditions.

In order to accurately capture density–dependent effects regulating cell cycle progression, we assume that f and g are non-increasing functions of the total density ρ. Again, this is motivated by the experimental observations of *Streichan et al.* [199] and *Donker et al.* [75]. Furthermore, we assume $f(0) = g(0) = 1$, so that k_1 and k_2 represent low–density transition rates. Note that setting $f = g \equiv 1$ gives rise to an exponential growth model (i.e. no dependence on density). On the other hand, choosing $f \equiv 1$ and $g(\rho) = (1 - \rho/K)_+$ we recover the *Vittadello et al.* model [213]. Here, we assume that $f(\rho)$ and $g(\rho)$ decrease linearly with the total cell density so that

$$f(\rho) = \left(1 - \frac{\rho}{K_1}\right)_+, \quad g(\rho) = \left(1 - \frac{\rho}{K_2}\right)_+, \tag{3.2}$$

where $K_1, K_2 > 0$ are constants controlling the impact of local cell density on cell cycle progression, and $(z)_+ = \max(z, 0)$. The specific form of these *crowding functions* is chosen here for simplicity, although other functions sharing the same properties show similar qualitative behaviour.

3.2.1 Bayesian inference

Following our previous work ([89], Chapter 2), we take a Bayesian approach [115, 184, 188] to calibrate the model given in Eqs. (3.1). In particular, given experimental measurements of the cell densities , and a vector of model parameters θ, we estimate the posterior probability distribution $p(\theta|\rho^{\mathcal{D}})$, which gives the probability density for the model parameters taking specific values.

All experimental datasets [112, 188] consist of direct measurements of the density of cells in the G1/post-M, and S/G2/M phases of the cell cycle. We denote these measurements by $\{\rho_1^{\mathcal{D}}(\mathbf{x}_i, t_j), \rho_2^{\mathcal{D}}(\mathbf{x}_i, t_j)\}_{i,j}$. The next step in order to estimate the different model parameters is to assume an error model, which relates experimental measurements with the model predictions given by the solutions of the model: $\rho_1(\mathbf{x}, t), \rho_2(\mathbf{x}, t)$. For simplicity, here we assume again that the residuals are independent and normally distributed

$$\rho_1^{\mathcal{D}}(\mathbf{x}_i, t_j) - \rho_1(\mathbf{x}_i, t_j) \overset{\text{iid}}{\sim} N(0, \sigma_1^2)\,,$$
$$\rho_2^{\mathcal{D}}(\mathbf{x}_i, t_j) - \rho_2(\mathbf{x}_i, t_j) \overset{\text{iid}}{\sim} N(0, \sigma_2^2)\,,$$

where σ_1 and σ_2 are parameters to be estimated from the data.

The white noise assumption has the advantage that simple likelihood-based methods can be used for inference. In particular, the log-likelihood of observing the data, given specific model parameters θ, can be written as

$$\ell_{\mathcal{D}}(\theta) = -\frac{1}{2}\sum_{k=1}^{2}\sum_{i,j}\left(\log\left(2\pi\sigma_k^2\right) + \left(\frac{\rho_k^{\mathcal{D}}(\mathbf{x}_i, t_j) - \rho_k(\mathbf{x}_i, t_j)}{\sigma_k}\right)^2\right).$$

To explore parameter identifiability, we follow a Bayesian approach. We use a MCMC sampler with adaptive proposal covariance to infer the posterior distributions. This is implemented in the parameter estimation toolbox pyPESTO [184]. For the tissue expansion experiments [112], we assume a log-uniform prior on D, k_1, k_2, with bounds: $10^1\ \mu\text{m}^2/\text{h} < D < 10^4\ \mu\text{m}^2/\text{h}$; $10^{-4}\ \text{h}^{-1} < k_1, k_2 < 10^1\ \text{h}^{-1}$. This assumption allows us to consider a broad range of orders of magnitude for these parameters, although simpler uniform priors could also be used. For the parameters $K_1, K_2, \sigma_1, \sigma_2$, uniform priors with the following bounds were used: $0 < K_1, K_2 < 20000$ cells/mm^2,

$0 < \sigma_1, \sigma_2 < 2000$ cells/mm^2. For the scratch assay data, we follow [188] and assume $\sigma_1 = \sigma_2 = \sigma$, and a uniform prior in all model parameters with the following conservative bounds: $0 < D < 2000$ μm^2/h, $0 < k_1, k_2 < 0.2$ h^{-1}, $0 < K_1, K_2 < 30000$ cells/mm^2, $0 < \sigma < 4000$ cells/mm^2. Figure 3.2 shows typical MCMC iterations for the *Heinrich et al.* experiments.

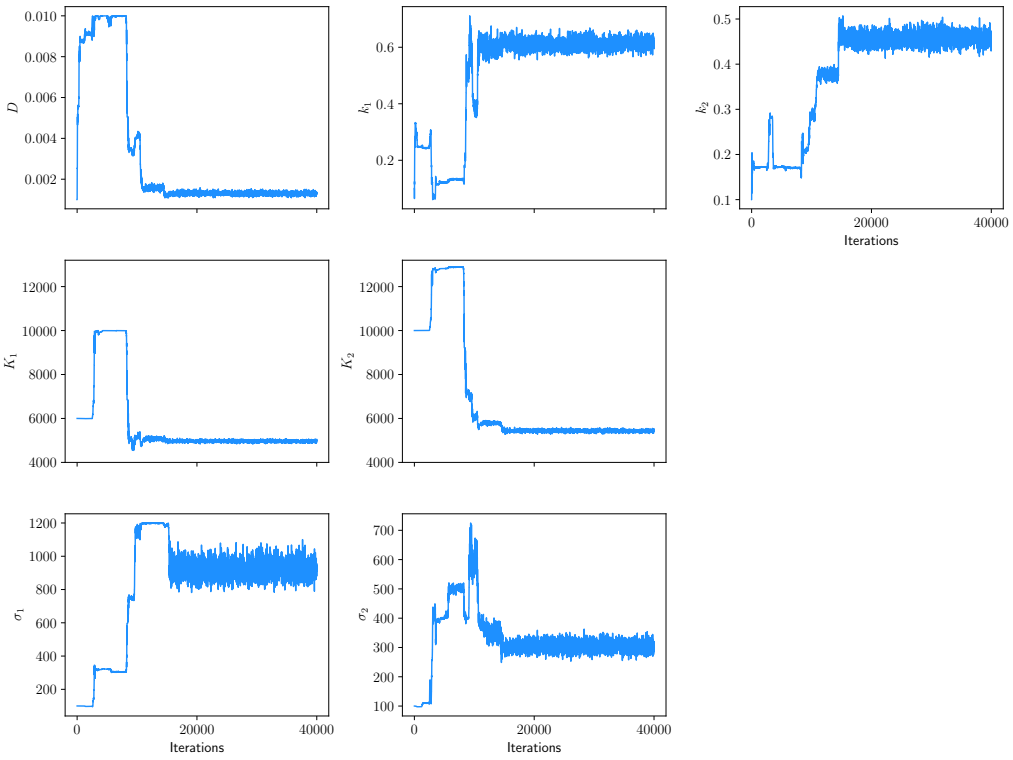


Figure 3.2: MCMC iterations for the tissue expansions experimental data. Parameters D, k_1, k_2, K_1, K_2 correspond to the model presented Eqs. (3.1), and σ_1, σ_2 are error model parameters.

In this chapter, we again report deterministic simulations evaluated at the posterior modes of all parameters, without propagating observation noise or parameter uncertainty when comparing with experimental data. This choice follows our approach in Chapter 2 — see the discussion there for details and justification.

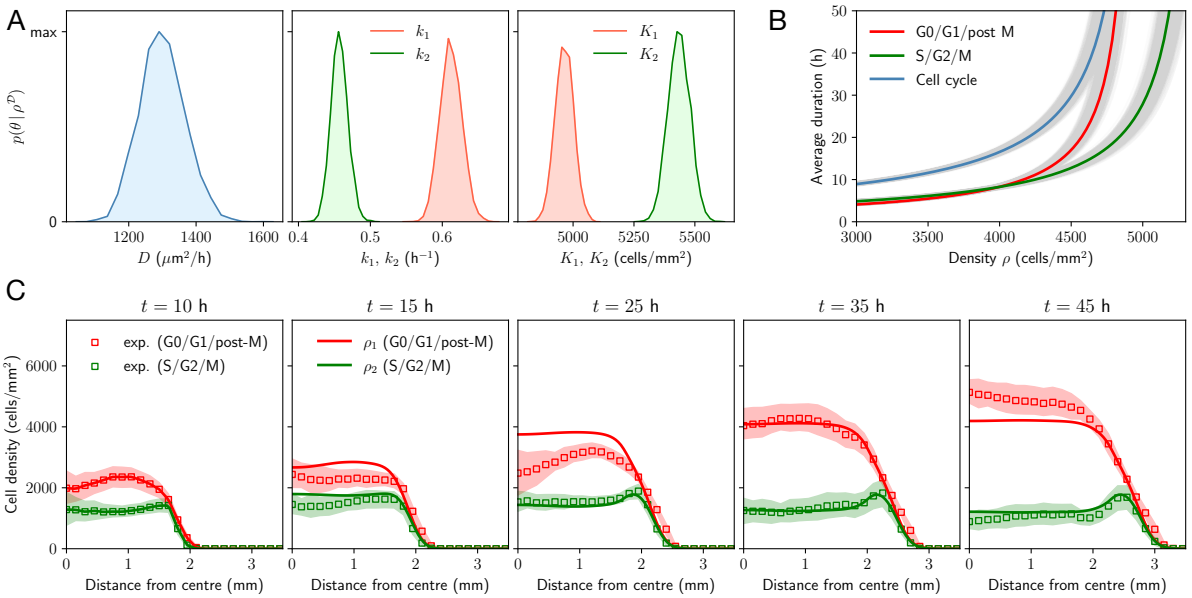


Figure 3.3: Density–dependent effects regulate cell cycle dynamics in epithelial tissue expansion experiments [112]. Parameter estimation and model-data comparison for the model given by Eqs. (3.1) and 3.2. (A) Univariate marginal posterior distributions for the model parameters. Posterior modes are given by $(D, k_1, k_2, K_1, K_2) = (1300 \pm 66\ \mu\text{m}^2/\text{h}, 0.612 \pm 0.015\ \text{h}^{-1}, 0.457 \pm 0.011\ \text{h}^{-1}, 4965 \pm 38\ \text{cells}/\text{mm}^2, 5435 \pm 45\ \text{cells}/\text{mm}^2)$, where errors correspond to one standard deviation. (B) Estimated duration of the G0/G1/post M (red) and S/G2/M (green) phases, as well as the whole cell cycle (black), as a function of cell densities. Solid lines correspond to posterior modes and shaded regions are obtained sampling from the posterior distribution. (C) Comparing data and model predictions. Squares represent the estimated cell density obtained by averaging eleven experimental realisations, which we use to calibrate the model. Shaded regions denote one standard deviation with respect to the mean. Numerical simulations in polar coordinates were obtained by using the posterior modes as parameter values.

3.3 Results

3.3.1 Tissue expansion experiments

We compare our model predictions to the experiments performed by *Heinrich et al.*[112] studying the expansion and growth dynamics of a single circular epithelial tissue — see Figure 3.1B. In these experiments, MDCK cells expressing the FUCCI markers are cultured in a silicone stencil for 18 hours and, after stencil removal, the cell population is allowed to freely expand for 46 hours. Given that the average cell cycle duration for MDCK cells is around 16 hours, this enables each cell to potentially undergo 2-3

cell divisions during the experiment. Local densities are then quantified by segmenting the fluorescence images in ImageJ and counting the number of nucleus centroids — Figure 3.1C. Note that post–mitotic cells do not fluoresce and appear dark, which makes the FUCCI system unreliable for cell counting. To quantify the density of post–mitotic cells, *Heinrich et al.* used a convolutional neural network to identify nuclei from phase contrast images [132] — see [112] for more details. Moreover, and in line with previous work [89], the model takes as initial condition the quantified density profile ten hours after stencil removal, so that the impact of the stencil on the dynamics is reduced. Note that after this time, cell densities near the tissue centre are relatively high ($\sim$ 3500 cells/mm^2, which corresponds to around 50-70% of the maximum saturation density for MDCK [89, 112]) and a fraction of cells in this region are likely to be found in a quiescent state due to contact inhibition of locomotion and proliferation [173].

The experiments by *Heinrich et al.* [112] reveal a higher density of cells in G0/G1 at the centre of the tissue, where the total cell density is also higher — see Figure 3.1D-E. The tissue edge, in contrast, is characterised by a larger number of cells which are preparing to divide (green) or are directly post–mitotic (grey). This agrees with previous observations of epithelial cells, which are known to control progression from G1 to S in response to spatial constraints [199]. Note, however, that the density of cells in S/G2/M in the tissue centre is low but non-zero, even at later times in the experiment — see Figure 3.1D-F — as observed also by *Donker et al.* [75].

For the sake of simplicity, here we consider post–mitotic cells (grey in Figure 3.1) and cells in G0/G1, as one single cell population. Quantifying post–mitotic cell density is crucial in order to estimate both K_1 and K_2 in Eqs. (3.2), given that these parameters are measures of contact inhibition of proliferation, typically associated with regions of higher cell density [218].

To calibrate the model, we fitted to the estimated cell density obtained by averaging eleven experimental realisations. We show the univariate marginal posterior distributions corresponding to the model parameters in Figure 3.3A, confirming that all model parameters are practically identifiable. In particular, all marginal posteriors show well-defined and unimodal distributions, with a relatively narrow variance.

The posterior distributions in Figure 3.3A are not only useful to inform further model predictions, but also give insights into the fundamental mechanisms underlying cell proliferation. In particular, given the intrinsic transition rate from G1 to S, k_1, and the constant K_1 in Eqs. (3.2), we can estimate the average duration of the combined G1/post-M phase, for a given fixed density ρ, as $1/k_1 f(\rho) = 1/(k_1(1-\rho/K_1)_+)$. Analogously, the estimated average duration of the S/G2/M phases is given by $1/k_2 g(\rho) = 1/(k_2(1-\rho/K_2)_+)$. Note, however, that these are only estimates of the timescales associated with different cell cycle stages. Put together, these estimates predict for a range of densities between 4000-4500 cells/mm^2, a population doubling time of 14-20 hours. In Figure 3.3B we plot these timescales as a function of the density ρ, observing how the duration of the different cell cycle stages increases with density. These results confirm again, in line with previous experimental measurements [75, 199], that cell cycle dynamics are tightly regulated by density–dependent effects. In particular, these estimates are consistent with the experimental measurements of *Donker et al.* [75] – taking into account that the initial cell densities in our datasets are around $\rho \sim 3500$ cells/mm^2. At very low densities, however, our estimates predict a relatively short cell cycle duration. This suggests that the shape of the *crowding functions* f and g might be closer to a constant function in this regime.

In Figure 3.3C we show numerical solutions of the model (Eqs. (3.1) and (3.2)), taking the posterior modes as parameter values. These confirm that the model can describe cell cycle dynamics inside expanding epithelial tissues. Notably, the model captures the tissue expansion speed, as well as the S/G2/M density peak near the edge of the tissue, which results from density–dependent effects regulating the cell cycle. We also note that this type of density profile is possible in the model when crowding-dependent effects are stronger in the early stages of the cell cycle (G1/post M) and weaker in the latter ones (S/G2/M). In terms of Eqs. (3.2) this requires having $K_1 < K_2$, which is correctly identified from the data.

We observe that the model overestimates the experimental density for early times of the experiment and, as a result of the model fit, underestimates it at later times. This is likely due to the transient behaviour that cells exhibit immediately after stencil removal [121, 122], which could have an impact on cell behaviour even after the first ten hours

of expansion, as suggested also in previous studies [89]. However, we emphasise that tissue edge motion can be well described by the model.

A similar behaviour is reported when the model is compared to a second set of experiments performed by *Heinrich et al.* [112]. In this case, we use the obtained parameter estimates to describe the expansion of initially smaller epithelial monolayers (initial diameter $\sim$ 1.7 mm). We highlight that the mathematical model can capture the expansion dynamics near the tissue edge as well as the expansion speed (see Figure 3.4), even though model parameters were inferred from the large tissue expansions.

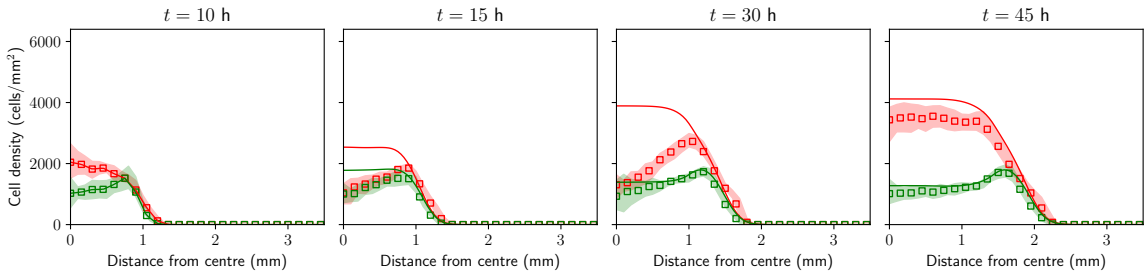


Figure 3.4: Comparing data (squares) and model predictions (solid lines) for small tissue expansions. Shaded regions denote one standard deviation with respect to the mean, obtained by averaging five experimental realisations. Numerical simulations in polar coordinates were obtained by using the parameter values obtained from the large tissue expansions (posterior modes), and no-flux boundary conditions.

3.3.2 Tissue colonisation experiments

Our model, together with the experiments of *Heinrich et al.* [112], reveals the intrinsic connection between tissue crowding and cell cycle progression, showcasing how this interplay can give rise to spatiotemporal patterns of cell proliferation in growing tissues. Next, we show how the model can be used to study and describe similar patterns observed in several other experimental studies using FUCCI and variants of it.

Streichan et al. [199] show, using a tissue barrier assay, how the cell cycle can be reactivated by allowing cells to migrate and colonise free space — see top row in Figure 3.5A. These experiments are initialised by growing MDCK-2 FUCCI cells in the G0/G1 phase within a removable barrier. After barrier removal, the tissue quickly colonises the available space, and cells behind the barrier, which were initially in G0/G1,

reactivate their cycle by entering S phase. On the other hand, cells located further behind the barrier remain at high density and do not progress through the cell cycle.

By solving numerically Eqs. (3.1) on a one-dimensional domain — see bottom row in Figure 3.5A — we immediately observe how a model accounting for density–dependent regulation predicts similar behaviour to that observed experimentally[1]. In particular, and as inferred from the experimental data of *Heinrich et al.* [112], the calibrated model predicts that crowding-dependent effects have a greater impact at the G1-S transition, compared to the S/G2/M phases. In terms of the model and the choice of *crowding functions* (Eqs. (3.2)), this once again requires $K_1 < K_2$.

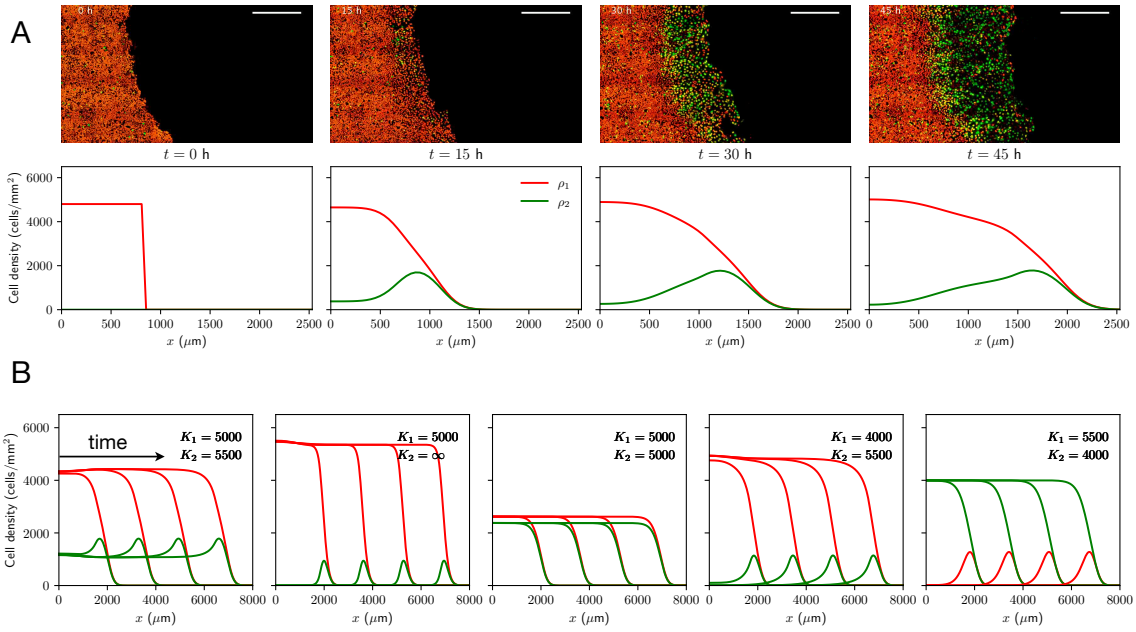


Figure 3.5: Cell cycle regulation by tissue crowding impacts cell migration. (A) Comparison with the tissue colonisation experiments of *Streichan et al.* [199] (top row, adapted with permission). Scale bars correspond to 500 μm. Bottom row shows numerical solutions of Eqs. (3.1) on a one-dimensional domain of length 3000 μm with no-flux boundary conditions, and initial conditions: $\rho_1(x,0) = 4800$ cells/mm^2 for $x < 850$ μm and $\rho_1(x,0) = 0$ cells/mm^2 otherwise; $\rho_2(x,0) = 0$. Parameter values correspond to the posterior modes in Figure 3.3. (B) Travelling wave solutions of Eqs. (3.1) for different values of K_1 and K_2 and at time points $t = 50$, 100, 150, 200 h. Units of K_1 and K_2 are cells/mm^2. Initial conditions: $\rho_1(x,0) = \rho_2(x,0) = 500$ cells/mm^2 for $x < 850$ μm, and $\rho_1(x,0) = \rho_2(x,0) = 0$ cells/mm^2 otherwise. In all cases, all parameters except for K_1 and K_2 are fixed (taken from posterior modes).

[1]Note that the experimental images from *Streichan et al.* [199] do not show post–mitotic cells which appear dark in the FUCCI system, and that the total cell density (including post–mitotic cells) was used to estimate the parameters in the model.

3.3.3 Cell cycle regulation and cell migration

Given that assuming $K_1 < K_2$ seems necessary in order to obtain biologically realistic model predictions, *what role does tissue crowding play in shaping cell migration patterns?* We explore this question by varying the values of K_1 and K_2 in Eqs. (3.2) — Figure 3.5B. First, we observe that the density of S/G2/M peaks near the tissue edge when $K_2 > K_1 > 0$ and remains low in the tissue bulk as long as $K_2 \gg K_1$. However, for $K_2 \sim K_1$, the height of this peak decreases and the fraction of S/G2/M cells in the tissue bulk increases. On the other hand, when we assume a higher influence of density during S/G2/M relative to G1/post-M ($K_2 < K_1$), we observe that the tissue centre shows a higher fraction of cells in S/G2/M, in contrast with previously reported observation of contact inhibition of proliferation [173].

3.3.3.1 Minimum travelling wave speed

The numerical solutions in Figure 3.5B suggest that low-density initial conditions lead to travelling wave solutions in one spatial dimension: $\rho_1(x-ct)$, $\rho_2(x-ct)$, with $c > 0$ being the wave speed, and x denoting the spatial coordinate. Standard arguments predict the existence of a minimum travelling wave speed in terms of only three model parameters. We thus look for travelling wave solutions in the model given by Eqs. (3.1) in one spatial dimension. We assume that the *crowding functions* $f(\rho)$ and $g(\rho)$ are non-increasing with ρ, and non-negative. In the comoving reference frame, we can write: $\rho_1(x,t) = U_1(z)$, $\rho_2(x,t) = U_2(z)$, where $z = x - ct$. By denoting $U = U_1 + U_2$, $V_1 = U_1'$, $V_2 = U_2'$, the model reduces to

$$\begin{cases} U_1' &= V_1\,, \\ DV_1' &= -cV_1 + k_1 U_1 f(U) - 2k_2 U_2 g(U)\,, \\ U_2' &= V_2\,, \\ DV_2' &= -cV_2 - k_1 U_1 f(U) + k_2 U_2 g(U)\,, \end{cases} \tag{3.3}$$

where the primes indicate differentiation with respect to z.

The set of steady states of system (3.3) consists of the origin $(U_1, V_1, U_2, V_2) = (0,0,0,0)$ and any state of the form $(\gamma, 0, U^* - \gamma, 0)$, with $f(U^*) = g(U^*) = 0$ and $0 \leq \gamma \leq U^*$. Note that whenever $f(\rho)$, $g(\rho) > 0$ for all $\rho \geq 0$, the latter does not exist.

As usual with linear diffusion models, the stability of the origin gives a lower bound on the wave speed c. In particular the Jacobian of system (3.3) at the origin reads

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ k_1/D & -c/D & -2k_2/D & 0 \\ 0 & 0 & 0 & 1 \\ -k_1/D & 0 & k_2/D & -c/D \end{pmatrix}.$$

The eigenvalues λ_i of the linearized system about this point satisfy the polynomial equation

$$\lambda^4 + \frac{2c}{D}\lambda^3 + \left(\left(\frac{c}{D}\right)^2 - \frac{k_1+k_2}{D}\right)\lambda^2 - c\frac{k_1+k_2}{D^2}\lambda - k_1k_2 = 0\,.$$

By defining

$$\gamma^{\pm} = \left(\frac{c}{D}\right)^2 + \frac{2}{D}\left[k_1 + k_2 \pm \sqrt{k_1^2 + k_2^2 + 6k_1k_2}\right],$$

the roots of this quartic polynomial can be expressed as

$$\lambda_1^{\pm} = \frac{1}{2}\left(-\frac{c}{D} \pm \sqrt{\gamma^+}\right), \quad \lambda_2^{\pm} = \frac{1}{2}\left(-\frac{c}{D} \pm \sqrt{\gamma^-}\right).$$

We seek biologically realistic solutions with $U_1, U_2 \geq 0$, and hence the eigenvalues must be real. In particular, this demands $\gamma^{\pm} \geq 0$, which establishes the minimum travelling wave speed found in [213]

$$c_{\min} = \sqrt{2D\left(-k_1 - k_2 + \sqrt{k_1^2 + k_2^2 + 6k_1k_2}\right)}. \tag{3.4}$$

As in the Fisher–KPP equation, an analysis of the dispersion relation shows that when the initial condition decays faster than exponentially, e.g. compact support, the growth in the leading edge is governed by the linearised model. Consequently, the invasion is pulled by this leading edge and the minimal wave speed is linearly selected [213].

Interestingly, Eq. (3.4) suggests that the invasion speed is independent of cell cycle regulation, and only depends on cell motility (D), and the intrinsic, low-density growth rates (k_1 and k_2). However, we highlight that, as shown in the figure, crowding constraints play an important role in shaping collective migration patterns.

The expression for the minimum travelling wave speed facilitates a comparison between the two-stage model proposed here (Eqs. (3.1)), and conventional single-population models of cell migration of the form

$$\partial_t \rho = D\Delta\rho + r\rho F(\rho)\,,$$

where F is a non-increasing function of the total cell density ρ, satisfying $F(0) = 1$. The intrinsic growth rate of the population, r, is related to the intrinsic rates of cell cycle progression, k_1 and k_2, via $r^{-1} = k_1^{-1} + k_2^{-1}$. When $4r/(k_1 + k_2) \ll 1$, Eq. (3.4) can be approximated by

$$c_{\min} \sim 2\sqrt{Dr}, \quad r = \frac{k_1 k_2}{k_1 + k_2},$$

which agrees with the prediction of the well-known Fisher–KPP equation ($F(\rho) = 1 - \rho/K$ for a maximum cellular density $K > 0$) in one spatial dimension. Using the estimated parameter values we obtain $4r/(k_1 + k_2) \sim 0.98$, and in this case Eq. (3.4) predicts a minimum travelling wave speed of $c_{\min} \sim 33$ μm/h, while the Fisher–KPP approximation yields $c_{\min} \sim 26$ μm/h; both of them within the range of values measured by *Heinrich et al.* [112].

A better comparison with the two-stage model can be obtained by setting $rF(\rho) = \lambda(\rho)$, where $\lambda(\rho)$ is the dominant eigenvalue of the growth matrix

$$\begin{pmatrix} -k_1 f(\rho) & 2k_2 g(\rho) \\ k_1 f(\rho) & -k_2 g(\rho) \end{pmatrix},$$

as given by Eqs. (3.1). In this case,

$$c_{\min} = 2\sqrt{D\lambda(0)},$$

where $\lambda(0) = (-k_1 - k_2 + \sqrt{k_1^2 + k_2^2 + 6k_1 k_2})/2$, agreeing with the prediction from Eq. (3.4).

3.3.3.2 Study of travelling wave solutions

More generally, we noted that $c_{\min}$ does not depend on the choice of *crowding functions* f and g; however, crowding constraints have an impact on the observed migration patterns (Figure 3.5B). To understand how growth and cell cycle regulation lead to the patterns observed experimentally, we investigate travelling wave solutions in a simplified setting.

A commonly used approach to obtain approximate solutions for travelling waves is the so-called Canosa's method [42]. This procedure is a standard singular perturbation technique, and consists of a transformation $y = -z/c$, where $D/c^2 := \varepsilon$ is treated as a small parameter. The first-order perturbation in ε approximates, within a small error,

travelling solutions of the well-known Fisher–KPP equation, even though in this case ε is not necessarily small [163]. In our case, by using the esimated parameters and a wave speed of 30 μm/h, we obtain $\varepsilon \sim O(1)$. We highlight, however, that the lowest order approximation in ε provides an excellent approximation of the travelling wave — see Figure 3.6.

By using the transformation $y = -z/c$, we can write system (3.3) as

$$\frac{\mathrm{d}U_1}{\mathrm{d}y} - \varepsilon\frac{\mathrm{d}^2U_1}{\mathrm{d}y^2} + k_1U_1f(U) - 2k_2U_2g(U) = 0\,, \tag{3.5}$$

$$\frac{\mathrm{d}U_2}{\mathrm{d}y} - \varepsilon\frac{\mathrm{d}^2U_2}{\mathrm{d}y^2} - k_1U_1f(U) + k_2U_2g(U) = 0\,. \tag{3.6}$$

Observe that, given the sign of the transformation $y = -z/c$, we need to impose the following boundary conditions

$$U_1(-\infty) = U_2(-\infty) = 0, \quad U_1(+\infty) = \alpha, \quad U_2(+\infty) = U^* - \alpha\,,$$

with $f(U^*) = g(U^*) = 0$ and $0 \leq \alpha \leq U^*$. For the choice of f and g in the main text ($f(U) = (1 - U/K_1)_+$ and $g(U) = (1 - U/K_2)_+$ with $K_1 < K_2$), we expect $U^* = K_2$ and $\alpha \in [0, K_2]$.

Although the analysis as $\varepsilon \to 0$ looks like a singular perturbation problem, setting $\varepsilon = 0$ gives a valid first-order approximation. This is due to the fact that the nonlinear terms in Eqs. (3.5) and (3.6) vanish at both boundaries [163]. Hence, we can look for a regular perturbation expansion in both U_1 and U_2. By denoting the order $O(1)$ solutions as u_1 and u_2 we obtain

$$\frac{\mathrm{d}u_1}{\mathrm{d}y} = -k_1u_1f(u) + 2k_2u_2g(u)\,, \tag{3.7}$$

$$\frac{\mathrm{d}u_2}{\mathrm{d}y} = k_1u_1f(u) - k_2u_2g(u)\,, \tag{3.8}$$

where $u = u_1 + u_2$. In Figure 3.6), we compare the approximate solutions obtained by solving this system with the full travelling wave solutions. We highlight that the lowest order approximation provides an excellent approximation of the travelling wave shape.

Next, and in order to make analytical progress, we set f and g to be Heaviside functions: $f(u) = H(K_1 - u)$ and $g(u) = H(K_2 - u)$. This model is not an approximation of the model presented here, but a simplification which preserves the same qualitative

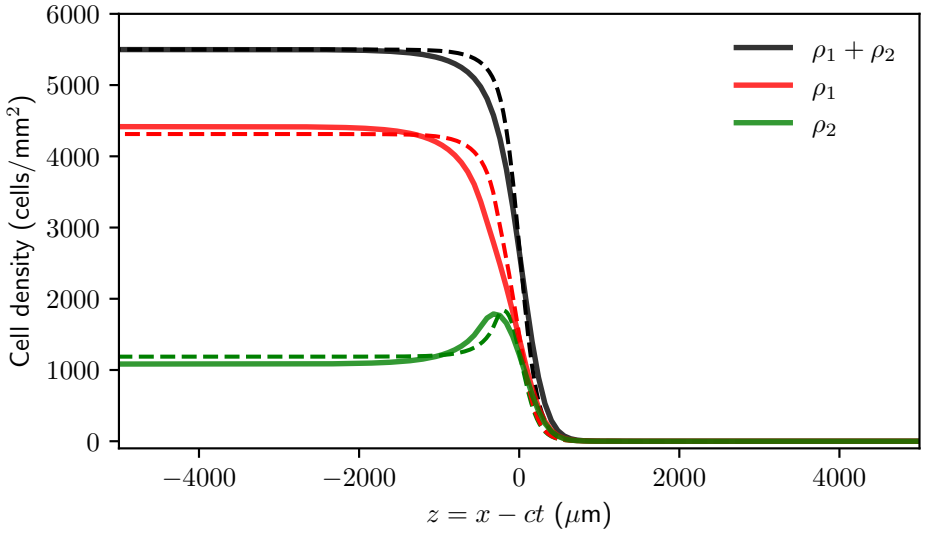


Figure 3.6: Comparison of travelling wave solutions obtained from the partial differential model (solid lines), given by Eqs. (3.1), and the order $O(1)$ approximation (dashed lines), obtained from solving the ordinary differential equations (3.7) and (3.8). Model parameters are taken from posterior distribution modes.

behaviour. Hence, we expect that the observed phenomena show similar dependence on the model parameters; this will be numerically confirmed later. In particular, note that this simplified model also describes two density checkpoints, at the G1-S boundary, and during the G2/M phases. The parameters K_1 and K_2, respectively, quantify the cell density associated with these checkpoints.

We rewrite Eqs. (3.7) and. (3.8) in terms of the variables (u, u_2),

$$\frac{\mathrm{d}u}{\mathrm{d}y} = k_2 u_2 g(u)\,,$$
$$\frac{\mathrm{d}u_2}{\mathrm{d}y} = k_1(u - u_2)f(u) - k_2 u_2 g(u)\,.$$

Depending on the relative values of the total cell density, u, and the density checkpoints parameters K_1, K_2, we distinguish three possible cases. As inferred from the experimental data, we assume $K_1 < K_2$.

Tissue edge ($u < K_1 < K_2$). In this region $f(u) = g(u) = 1$ and we can write

$$\frac{\mathrm{d}u}{\mathrm{d}y} = k_2 u_2\,,$$
$$\frac{\mathrm{d}u_2}{\mathrm{d}y} = k_1 u - (k_1 + k_2)u_2\,.$$

The solution at the tissue edge reads

$$u(y) = e^{-(k_1+k_2)y/2}\left[Ae^{\gamma y/2} + Be^{-\gamma y/2}\right],$$
$$u_2(y) = e^{-(k_1+k_2)y/2}\left[\frac{A\left(\gamma-(k_1+k_2)\right)}{2k_2}e^{\gamma y/2} - \frac{B\left(\gamma+(k_1+k_2)\right)}{2k_2}e^{-\gamma y/2}\right],$$

where A, B are constants to be determined, and $\gamma = \sqrt{(k_1+k_2)^2+4k_1k_2}$. Imposing boundary conditions at $y \longrightarrow -\infty$, and noting that $\gamma-(k_1+k_2) > 0$, we obtain $B = 0$. Hence, for $u < K_1$ both solutions are increasing exponentials. Without loss of generality we set $U(0) = K_1$, giving $A = K_1$, and thus

$$u(y) = K_1 e^{(\gamma-k_1-k_2)y/2},$$
$$u_2(y) = \frac{K_1\left(\gamma-(k_1+k_2)\right)}{2k_2}e^{(\gamma-k_1-k_2)y/2}. \tag{3.9}$$

Intermediate region ($K_1 < u < K_2$). In this region $f(u) = 0$ and $g(u) = 1$, leading to

$$\frac{\mathrm{d}u}{\mathrm{d}y} = k_2u_2,$$
$$\frac{\mathrm{d}u_2}{\mathrm{d}y} = -k_2u_2.$$

Hence, $u_2(y) = Ce^{-k_2y}$ for a constant C, which can be found by continuity at $z = 0$. We obtain

$$u_2(y) = \frac{K_1\left(\gamma-(k_1+k_2)\right)}{2k_2}e^{-k_2y}. \tag{3.10}$$

Given that this is a decreasing exponential, we have found that the peak in S/G2/M cell density, ρ_2^{edge}, occurs at $y = 0$. This is, $\rho_2^{\mathrm{edge}} = u_2(0)$.

Since in this region $u + u_2$ is a constant, which can be found by continuity at $y = 0$, we also obtain

$$u(y) = K_1 + u_2(0) - u_2(y).$$

Tissue bulk ($K_1 < K_2 < u$). Now we have $f(u) = g(u) = 0$ and hence we can write

$$\frac{\mathrm{d}u}{\mathrm{d}y} = \frac{\mathrm{d}u_2}{\mathrm{d}y} = 0.$$

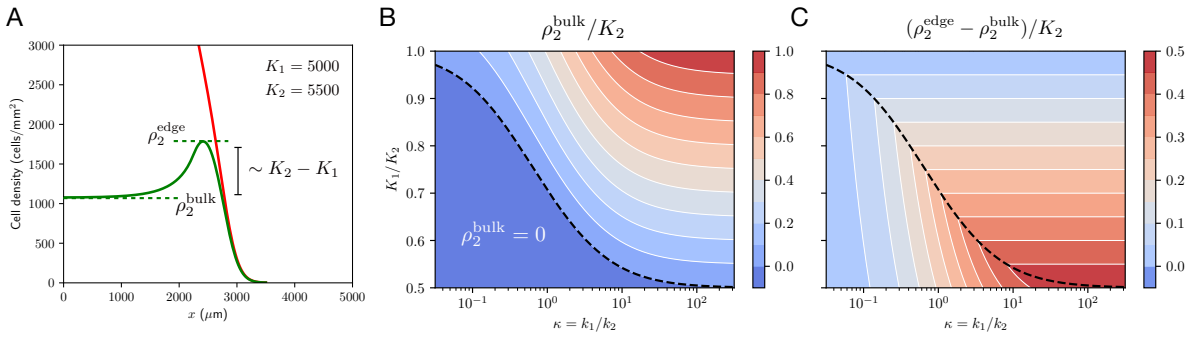


Figure 3.7: Cell cycle transition rates (k_1, k_2), and crowding constraints (K_1, K_2) determine cell proliferation patterns in growing tissues. (A) Schematic of travelling wave solutions near the tissue edge. (B)-(C) Approximated S/G2/M cell densities at the tissue edge and tissue bulk as a function of the ratios $\kappa = k_1/k_2$ and K_1/K_2. The black dashed line corresponds to the curve $K_1/K_2 = \alpha(\kappa) = 2/(\sqrt{\kappa^2 + 6\kappa + 1} - \kappa - 1)$.

In the tissue bulk, both densities are constant and $u = K_2$. By using continuity, and the solutions from the intermediate region, we find

$$\rho_2^{\text{bulk}} := u_2(z) = (K_1 - K_2 + u_2(0))_+ , \tag{3.11}$$

where we impose positivity of $u_2(z)$.

In particular, we find that, whenever $\rho_2^{\text{bulk}} > 0$, the S/G2/M cell density difference between the tissue edge and the bulk satisfies

$$\rho_2^{\text{edge}} - \rho_2^{\text{bulk}} = K_2 - K_1 .$$

In this case, and by combining Eqs. (3.9), (3.10), and (3.11), we obtain the full solution

$$u_2(y) = \begin{cases} \dfrac{K_1(\gamma - (k_1 + k_2))}{2k_2} \exp\left(\dfrac{\gamma - (k_1 + k_2)}{2} y\right), & y \le 0, \\ \dfrac{K_1(\gamma - (k_1 + k_2))}{2k_2} e^{-k_2 y}, & 0 < y \le y^*, \\ \dfrac{K_1(\gamma - (k_1 + k_2))}{2k_2} - (K_2 - K_1), & y > y^*, \end{cases} \tag{3.12}$$

for $y = -(x - ct)/c$, and y^* defined from Eq. (3.10): $k_2 y^* = \log(\rho_2^{\text{edge}}/\rho_2^{\text{bulk}})$.

The analysis above for the simpler model suggests that the density of S/G2/M cells in the tissue bulk, ρ_2^{bulk}, only depends on the ratio of cell cycle progression rates, $\kappa = k_1/k_2$, and on the ratio of densities associated with crowding constraints, K_1/K_2,

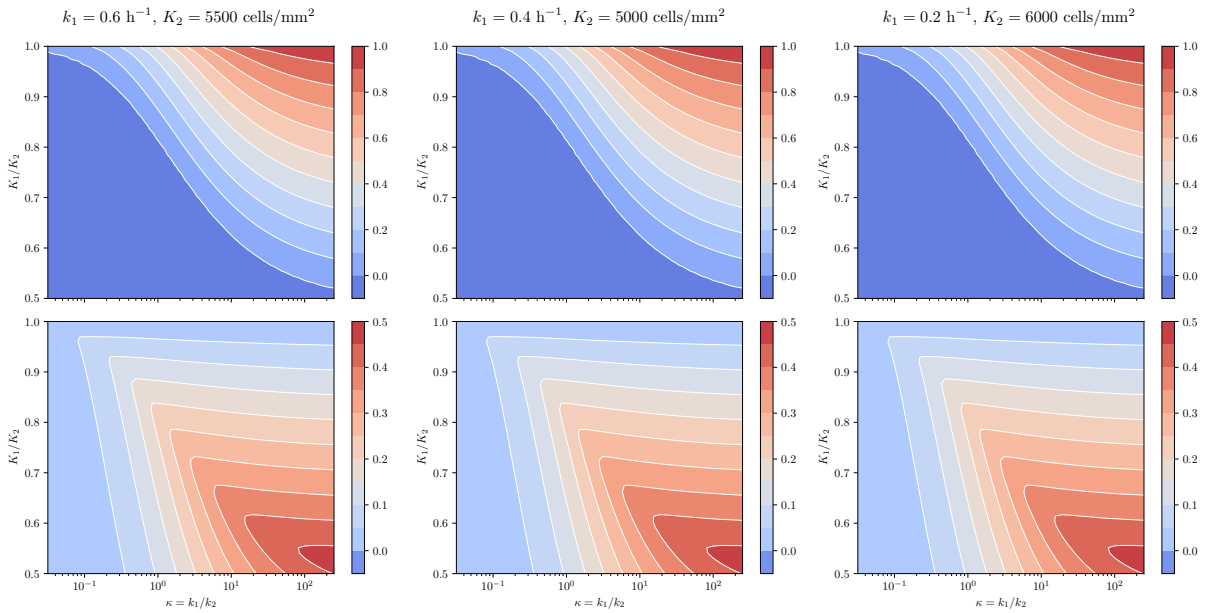


Figure 3.8: S/G2/M cell densities at the tissue edge and tissue bulk as a function of the ratios $\kappa = k_1/k_2$ and K_1/K_2 for the full cell cycle model with $f(\rho) = (1 - \rho/K_1)_+$, $g(\rho) = (1 - \rho/K_2)_+$. These results confirm that ρ_2^{edge} and ρ_2^{bulk} are determined by the two ratios of parameters: k_1/k_2 and K_1/K_2.

(Figure 3.7). In particular, we obtain

$$\frac{\rho_2^{\text{bulk}}}{K_2} \sim \begin{cases} \dfrac{K_1}{K_2}\dfrac{\sqrt{\kappa^2+6\kappa+1}-\kappa+1}{2} - 1\,, & K_1/K_2 > \alpha(\kappa)\,, \\ 0\,, & K_1/K_2 \leq \alpha(\kappa)\,, \end{cases}$$

where $\alpha(\kappa) = 2/(\sqrt{\kappa^2+6\kappa+1} - \kappa - 1)$. A similar dependence with respect to the model parameters is observed numerically for the model given by Eqs. (3.1) and (3.2) (see Figure 3.8). For our estimated parameters, the expression above predicts $\rho_2^{\text{bulk}}/K_2 \sim 0.3$, which is consistent with experimental observations. We also highlight that, as long as $K_1 < K_2$, and k_1 and k_2 are of a similar order of magnitude, this expression predicts that the number of cells in S/G2/M in the tissue bulk will be small in comparison to the number of cells in G1/post-M (Figure 3.7B). In particular, note that $\rho_2^{\text{bulk}} \to 0$ as $K_2 \to \infty$.

Interestingly, the travelling wave analysis also reveals that, when $\rho_2^{\text{bulk}} > 0$, the difference in S/G2/M cell density between the tissue bulk, ρ_2^{bulk}, and the tissue edge, ρ_2^{edge}, depends only on the difference of densities associated to crowding constraints at

the G1-S and G2-M boundaries (Figure 3.7C),

$$\rho_2^{\text{edge}} - \rho_2^{\text{bulk}} \sim \begin{cases} K_2 - K_1\,, & K_1/K_2 > \alpha(\kappa)\,; \\ \rho_2^{\text{edge}}\,, & K_1/K_2 \leq \alpha(\kappa)\,; \end{cases}$$

where

$$\rho_2^{\text{edge}} \sim \frac{K_1}{K_2}\frac{\sqrt{\kappa^2+6\kappa+1}-\kappa-1}{2}\,.$$

For our estimated parameters, we obtain $\rho_2^{\text{edge}} - \rho_2^{\text{bulk}} \sim 500$ cells/mm^2, again consistent with the experimental observations. These analytical expressions confirm the impact of density–dependent effects on cell migration and suggest that differences in the regulation of cell cycle stages contribute to the emergence of cell proliferation patterns.

Finally, we confirm numerically (see Figure 3.8) that a similar parametric dependence holds in the full model described by Eqs. (3.1) and (3.2), reinforcing the idea that these two ratios are key determinants of the spatial distribution of cell cycle stages during tissue growth.

3.3.4 Density–dependent effects and experimental design

We have demonstrated that our model can be calibrated to experimental data collected by *Heinrich et al.* [112] to provide confident estimates of all parameters and, from there, used to extract and quantify crowding constraints regulating the cell cycle. An obvious question to ask is whether the model parameters could also be confidently estimated from other datasets, in particular where the cell density remains much lower and the impact of tissue crowding is reduced. To explore this question, we attempt to estimate the model parameters (including K_1 and K_2) using data from a low-density scratch assay with 1205Lu melanoma cells — see Figure 3.9. We highlight that melanoma cells are highly metastatic and often display uncontrolled and invasive migration, in contrast to the highly collective and regulated movement exhibited by epithelial cells. In this experiment, tissues are seeded at an initial density of ~ 400 cells/mm^2 (5% of the theoretical maximum packing density [213]), and data is collected every 16 hours, over two full days, allowing cells to potentially undergo 1-2 cell cycles. The posterior distributions obtained for the different model parameters reveal estimates for D, k_1 and k_2 that

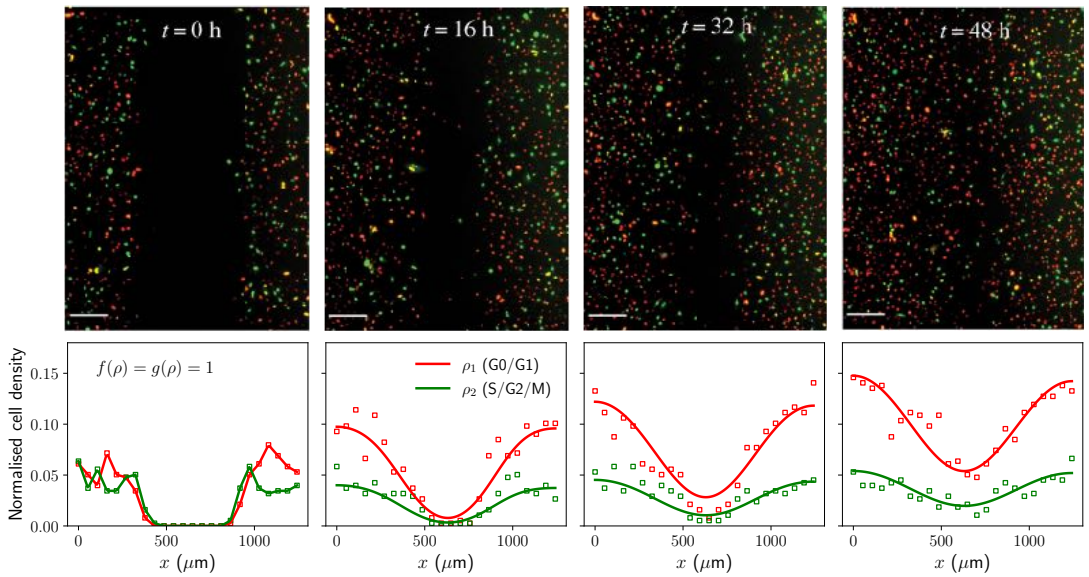


Figure 3.9: Absence of density–dependent effects in a low-density scratch assay experiment (1205Lu melanoma cells). In this case, an exponential growth model can reproduce the experimental data. Density is normalised by using the theoretical maximum density corresponding to hexagonal close packing of cells [213]. Top row is adapted from [188], with scale bars corresponding to 200 μm. Numerical solutions of Eqs. (3.1) with $f(\rho) = g(\rho) = 1$ on a one-dimensional domain. Parameters are estimated using the experimental data from [188] — posterior modes from Figure 3.11.

are consistent with previous studies [188]. However, the low experimental densities do not allow for the quantification of density–dependent effects — the parameters K_1 and K_2 cannot be estimated with any degree of confidence (see Figure 3.10). This non-identifiability of K_1 and K_2 suggests the use of a simpler model, which assumes that cell cycle progression is independent of density–dependent effects ($f(\rho) = g(\rho) = 1$) and hence is only valid in the low-density regime. Indeed, when calibrated to data from the low-density scratch assay it provides confident parameters estimates (Figure 3.11), and an excellent agreement with the experimental data — see Figure 3.9. This result clearly illustrates both the key role that mathematical modelling can play in the experimental design process, and the importance of considering parameter identifiability in the process of model construction.

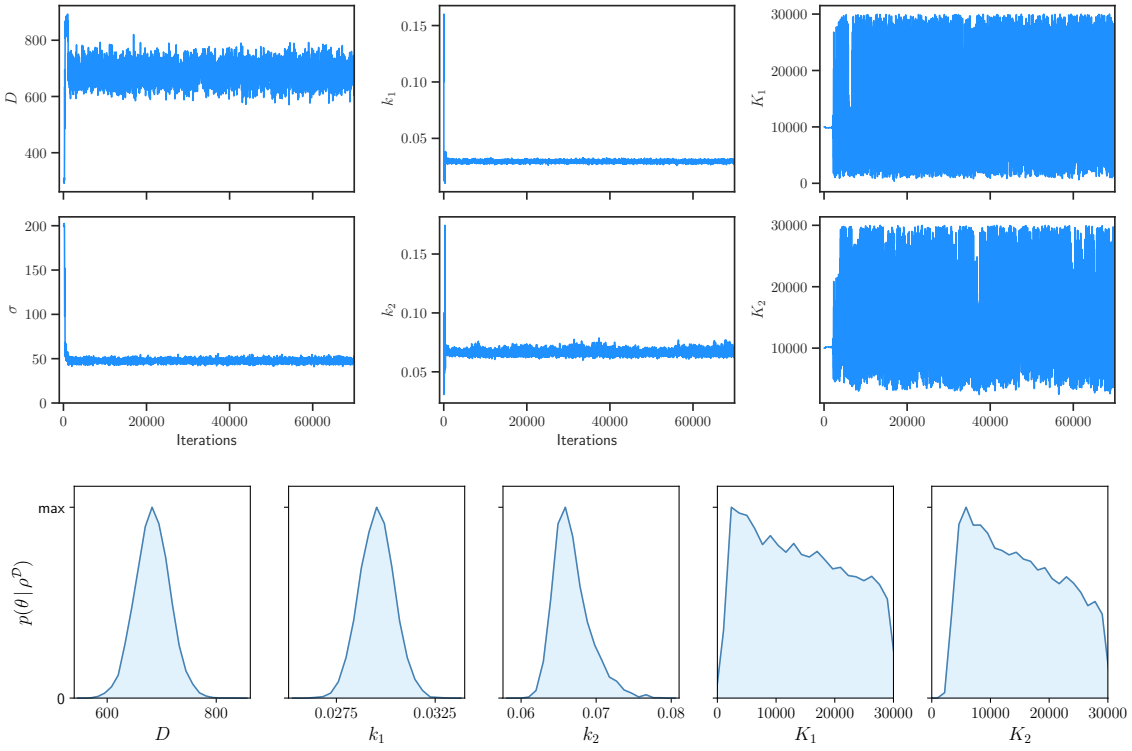


Figure 3.10: MCMC iterations and univariate marginal posterior distributions obtained by using low-density scratch assay data [188]. With this dataset, K_1 and K_2 are practically non-identifiable.

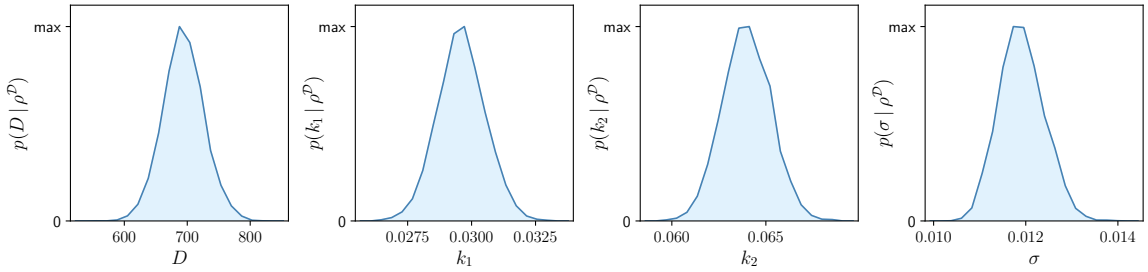


Figure 3.11: Univariate marginal posterior distributions for the exponential growth model ($f(\rho) = g(\rho) = 1$) obtained by using low-density scratch assay data [188]. In this case, the model has three parameters (D, k_1, k_2) and one error model parameter (σ, normalised by the theoretical maximum density assuming hexagonal packing of cells [213]).

3.4 Discussion and outlook

We have presented a new mathematical model of cell migration with cell cycle dynamics which captures and quantifies cell cycle regulation by sensing levels of tissue

crowding. In line with previous experimental studies, by combining minimal modelling and Bayesian inference, we confirm that cell cycle progression is monitored via crowding constraints [75, 199], and present a systematic approach towards the quantification of interactions regulating cell proliferation. Our model is capable of quantifying cell cycle data from experiments using the FUCCI system, and enables the extraction of mechanistic insights into how individual cells regulate proliferation based on population-level measures.

The model presented here offers several applications to further our understanding of cell-cell interactions in cell proliferation. In particular, our model presents a systematic way to quantify the impact of drugs and gene knockouts/knockdowns interfering with cell proliferation. By using parameter estimation techniques, applied to different experimental datasets, we can gain insights into the regulatory roles of specific genes in the cell cycle. Another possible application concerns the study of cell migration in biomaterials incorporating cadherin proteins, which have recently been shown to slow down cell cycle dynamics [200]. Furthermore, generalisations of the FUCCI system could allow for a finer representation of the different cell cycle stages — for instance, FUCCI4 [13] allows for the simultaneous visualisation of the four stages of the cell cycle. In line with these methodologies, extensions of our model (Eqs. (3.1)) to multi-stage cell populations are straightforward, and could enable a more exhaustive explanation of the role of spatial constraints across all four cell cycle stages [75].

In the case of *Heinrich et al.*'s experiments [112], the excellent imaging quality allowed us to perform an accurate quantification of the cellular density profiles. This, in turn, facilitated model development and the subsequent inference of model parameters from the data, with the estimated parameters showing a low uncertainty. While the parameter identifiability of such mathematical models can be evaluated a priori under the assumption of infinite ideal data [29, 177], biologically realistic datasets are finite, and often contain a significant level of noise, which can, in certain instances, constrain the ability to confidently estimate model parameters. More generally, and as we have illustrated, practical constraints in the experimental data often relate to the level of model complexity which can be inferred from experiments and the confidence in model parameter estimates.

Continuum models are a widely adopted approach for describing cell migration. However, these models come with limitations: they tend to neglect local structure, especially in situations involving multiple cell populations. Such local structure can be observed in Figure 3.1; (C) and (D) show some degree of local correlation in the cell phases, however this phenomenon is lost when averaging radially to obtain the density profiles in (E). Agent-based models [44, 118, 129] can help mitigate some of these issues, by providing more understanding of the generation and maintenance of spatial structure, but at the cost of increased computational times for simulation and inference, additional model parameters, and limited analytical tractability. We emphasise, however, that cell cycle dynamics appear to be globally desynchronised, as observed in previous studies [165], and so our differential equation-based model remains appropriate for this study, where the data is generated by averaging over a number of experimental replicates.

The model presented here is minimal in the sense that it assumes that cell movement is random, and it ignores basic cell-cell interactions such as cell-cell adhesion, which are typical of epithelial cell migration. While local cell density is likely to have an impact on cell motility [112], previous work shows that for individual expanding epithelial tissues, the linear diffusion model provides a good approximation [89]. Note, however, that it is important to account for population pressure and its impact on cell movement when considering tissue–tissue interactions [114]. Additional research is needed to determine whether more complicated models [56, 87], incorporating cell-cell adhesion and other basic interactions offer deeper mechanistic understanding. Moreover, the model given in Eqs. (3.1) assumes that at low densities, the duration of each of the cell cycle stages follows an exponential distribution. While this assumption contradicts experimental observations [192, 219] and can be mitigated by representing the cell cycle as a multi-stage process [100, 221], such models break the cell cycle into a very large number of stages, limiting the potential for calibration to experimental data. Additional investigation is required to understand the extent to which more complicated models can provide further insights into how cells coordinate proliferation and migration to give rise to complex collective behaviours. For example, in the context of the cell cycle, an option is to explicitly incorporate cell cycle stage via the use of an

age-structured model [131] that includes density–dependent regulation. Our results indicate that adopting a quantitative approach [144], that carefully examines quantitative data through the lens of mathematical modelling and Bayesian inference, can help provide answers to this question.

Chapter 4

Quantifying tissue–tissue collisions

In this chapter, we turn to a second set of experiments by *Heinrich et al.* [114], in which tissue–tissue interactions play a key role. In these experiments, epithelial monolayers are allowed to grow until they collide, generating mechanical pressure on one another. As we will see, the Fisher–KPP model is no longer sufficient to capture the dynamics observed in this setting. To accurately describe the system, it becomes necessary to incorporate population pressure into the model. The work in this chapter is published as

"C. Falcó, D. J. Cohen, J. A. Carrillo & R. E. Baker. *Quantifying tissue growth, shape and collision via continuum models and Bayesian inference.* Journal of the Royal Society Interface, 20(204):20230184, 2023" [89].

My contributions include model development, data curation, computational implementation, and writing of the publication. Experimental data corresponding to tissue–tissue collisions (Figure 4.3) was provided by *Heinrich et al.* Animated movies corresponding to the numerical simulations in this chapter can be found on Figshare. Code used to perform the parameter estimation, solve numerically the models, and create the animations is available on Github.

4.1 Introduction

Cells do not live in isolation; they coexist and organise to form tissues and, ultimately, functional organs. During growth and repair, epithelial cells sense their surroundings, exchange mechanical and biochemical cues, and navigate collectively rather than

as solitary agents. In Chapter 2 we calibrated two simple and well-known reaction-diffusion models — the Fisher–KPP and Porous–Fisher equations — against single-tissue spreading assays. We showed that, once calibrated, both models reproduce the observed expansion of an isolated epithelial monolayer with comparable accuracy. The key difference lies in their treatment of density-dependent motility: while Fisher–KPP assumes a constant diffusivity, the Porous–Fisher model incorporates crowding effects through a nonlinear diffusion term that vanishes at low cell densities. For isolated tissues, this distinction proved non-essential, either model sufficed.

The situation changes dramatically when two or more monolayers are allowed to grow and collide [114]. Here, the interface between neighbouring tissues becomes a dynamic, deformable region where mechanical forces, contact inhibition of locomotion [173], and density-dependent migration all interplay. In this chapter we demonstrate that, under these conditions, the Fisher–KPP model fails to capture the spatiotemporal evolution of the collision front, whereas the Porous–Fisher model — with its minimal yet crucial account of cell–cell interactions — accurately predicts both transient and steady-state patterns. Despite having only three free parameters, the Porous–Fisher equation proves strikingly versatile: it not only fits binary collision experiments, but also described the final tessellation of complex, multi-tissue composites that are of direct relevance to tissue engineering applications [89].

Our approach is structured as follows. First, building on the identifiability analysis of Chapter 2, we verify that the parameter sets inferred from single-tissue data can qualitatively describe binary collision experiments. Second, we test the two models against an extensive library of collision/healing assays reported by *Heinrich et al.* [114], which span a wide range of initial geometries and relative tissue sizes. Finally, leveraging the calibrated Porous–Fisher model, we quantify how interface velocity, curvature, and pressure gradients govern the self-organisation of multiple epithelial sheets and predict their long-time configurations.

By establishing the Porous–Fisher framework as the minimal continuum model that reconciles both isolated growth and tissue–tissue collisions, this chapter provides a robust theoretical foundation for designing and interpreting experiments in wound healing and tissue regeneration. Moreover, the insights gained here can potentially

guide the development of strategies aimed at steering collective migration in tissue engineering.

4.2 Modelling tissue–tissue collisions

Although we will always work with homotypic tissues (i.e. of the same cell type), it is particularly useful to identify a system consisting of multiple homotypic tissues with a model that accounts for several interacting cell populations. In our case, the tissues are composed of the same cell populations initially seeded at distinct spatial locations. Note, however, that the models presented below can account also for heterotypic tissue experiments. We denote the different species or tissues by ρ_i for $i = 1, \dots, n$ with n being the total number of species. In the linear diffusion Fisher–KPP model we assume that each species follows random motion and hence the diffusive part in the PDE remains unaffected. Taking into account that proliferation is limited by the total population density, we may write for $n = 2$

$$\begin{cases} \partial_t \rho_1 = D\Delta\rho_1 + r\rho_1 \left(1 - \dfrac{\rho_1 + \rho_2}{K}\right), \\ \partial_t \rho_2 = D\Delta\rho_2 + r\rho_2 \left(1 - \dfrac{\rho_1 + \rho_2}{K}\right). \end{cases} \tag{4.1}$$

For the nonlinear diffusion Porous–Fisher model, we can write the total population pressure as $P(\rho_1, \rho_2) = D\,(\rho_1 + \rho_2)$. With this, the two-species model becomes

$$\begin{cases} \partial_t \rho_1 = D\,\nabla \cdot (\rho_1 \nabla\,(\rho_1 + \rho_2)) + r\rho_1 \left(1 - \dfrac{\rho_1 + \rho_2}{K}\right), \\ \partial_t \rho_2 = D\,\nabla \cdot (\rho_2 \nabla\,(\rho_1 + \rho_2)) + r\rho_2 \left(1 - \dfrac{\rho_1 + \rho_2}{K}\right). \end{cases} \tag{4.2}$$

Extensions of these models to an arbitrary number of species, $n > 2$, are straightforward. The existence theory for cross-diffusion systems of the type of (4.2) is studied in [37, 72]. Note also that as a result of the population pressure term, system (4.2) gives sharp boundaries separating both species for initially segregated data [21, 52] which, again, motivates its use to reproduce the experiments in [114].

4.2.1 Reproducing experimental tissue collisions

In the next sections, we explore numerically the two proposed models under different initial conditions. We start with a qualitative study of some of the experiments performed by *Heinrich et al.* [114] and follow with a more quantitative analysis of collisions between rectangular tissues.

4.2.1.1 Binary tissue–tissue collisions

We first test the two proposed models in binary tissue–tissue collisions. In order to do so, here we choose different initial shapes for the two colliding tissues, namely we study circle–circle, rectangle–rectangle and circle–rectangle collisions. We also analyse the case of two colliding circles with different initial radii. See Figure 4.1A for the experimental initial and final configurations. We emphasise that, in contrast with all the shown numerical simulations of our models, the colours in the experimental snapshots are only used to label each different tissue and do not quantify cell densities.

We numerically solve Eqs. (4.2) for the four mentioned initial conditions and with the parameters that we estimated from the previous experiments [112]. The numerical scheme is identical to the one-species case [12, 45] — see also Appendix A. As expected, in all four studied configurations the Porous–Fisher model shows sharp boundaries separating the two tissues after collision, and the observed patterns are nearly identical to the experimental final configurations after a simulation time equivalent to around 60 hours (Figure 4.1B). Note that in contrast with the experimental snapshots, Figure 4.1B shows quantitative cell densities.

When, instead of the Porous–Fisher model accounting for population pressure (4.2), we use the Fisher–KPP model (4.1), we still observe patterns that resemble the experimental configurations. However, recall that in this case cells do not sense local pressure and are free to move in all directions, which results in a region where cells from both tissues can mix. Note that in this case, no sharp boundary between tissues is observed either — Figure 4.1C. Even though the Fisher–KPP model fails to reproduce density profiles near the collision boundary, it still can capture qualitatively the density profiles in the bulk of the tissue, where the population density gradient becomes more uniform. Hence, after suitable calibration, both the Fisher–KPP and the Porous–Fisher

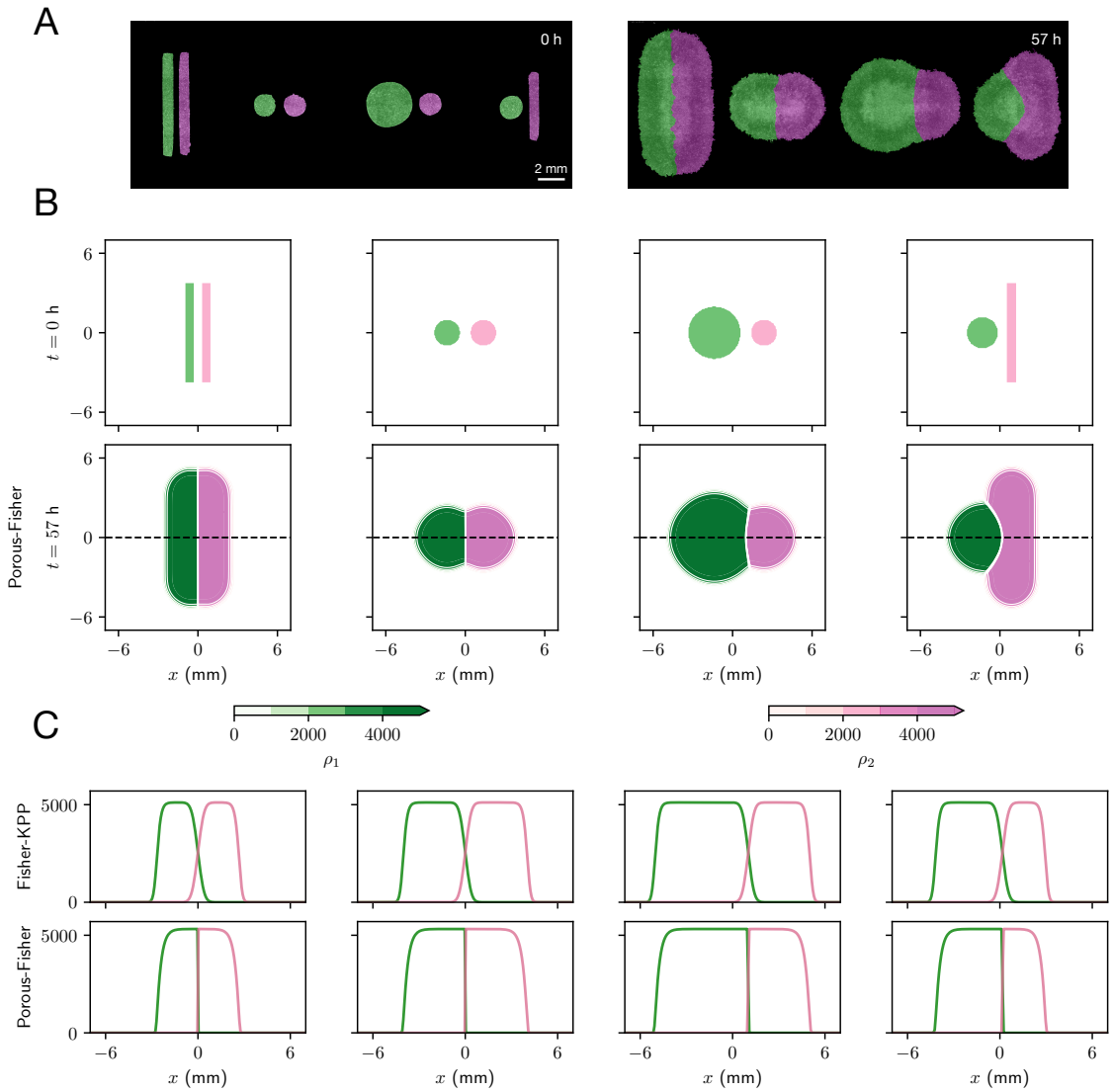


Figure 4.1: Reproducing tissue–tissue collisions with different geometries — animated movies available at Figshare. Accounting for population pressure correctly predicts the sharp boundaries observed in experiments. (A) Experimental results for initial conditions with different tissue geometries. Figures adapted from [114] (Creative Commons License). (B) Initial conditions and numerical simulations for the Porous–Fisher model (Eqs. (4.2)) at $t = 57$ h. (C) Comparison of the Fisher–KPP model (Eq. (4.1)) and the Porous–Fisher model (Eq. (4.2)). Solutions corresponding to the black dashed lines in (B). Parameter estimates given in Chapter 2: $(D, r, K) = (1073\ \mu\text{m}^2/\text{h}, 0.29\ \text{h}^{-1}, 5113\ \text{cells}/\text{mm}^2)$ for the Fisher–KPP model and $(D, r, K) = (1.18\ \mu\text{m}^2/(\text{cells}{\cdot}\text{h}), 0.21\ \text{h}^{-1}, 5319\ \text{cells}/\text{mm})^2$ for the Porous–Fisher model.

models show similar behaviour in this region far from the collision boundary and the propagating front. However, note that the Fisher–KPP model does not capture any interaction between colliding tissues and hence fails to describe the observed behaviour

near the collision boundary.

Observe also that collisions shown in Figure 4.1B that occur between tissues with the same shape (rectangle–rectangle and circle–circle collisions) were initialised with tissues of the same density. As reported experimentally in [114], these initial conditions result in the formation of a fixed sharp boundary that does not move in time. However, when collisions between tissues with different densities occur, then the denser tissue pushes the less dense tissue resulting in a boundary displacement which can be measured experimentally. For collisions between tissues with different shapes the collision boundary can also show a similar behaviour, as shown in Figure 4.1. In the next sections we study this phenomenon quantitatively using the Porous–Fisher model (4.2). Of course, given that linear diffusion fails to predict a sharp boundary between colliding tissues, this boundary displacement cannot be estimated from the Fisher–KPP model (4.1). Before moving to the study of collision boundary dynamics, we analyse a further set of more complex tissue collision experiments, which make evident the limitations of the simple Fisher–KPP model.

4.2.1.2 Multi-tissue collisions

We have showed that, after suitable calibration, both the Fisher–KPP and the Porous–Fisher models show similar behaviour in regions of tissue that are far from boundaries. However, under more complex experimental conditions where tissue boundary dynamics become important, the predictive power of the Fisher–KPP model becomes more limited.

These differences between the Fisher–KPP (4.1) and the Porous–Fisher model (4.2) become more evident when multiple tissues collide simultaneously. Here we focus on the experiments performed by *Heinrich et al.* [114] shown in Figure 4.2, where eight homotypic circular tissues are initially set apart on a hexagonal lattice. The initial configuration is also represented in Figure 4.2 alongside the solutions predicted by the two proposed models after 57 hours. From these results, it becomes evident that the Fisher–KPP model is not suitable to describe complex interactions between tissues. In contrast, accounting for population pressure does yield the predicted behaviour with a final pattern nearly identical to that observed experimentally.

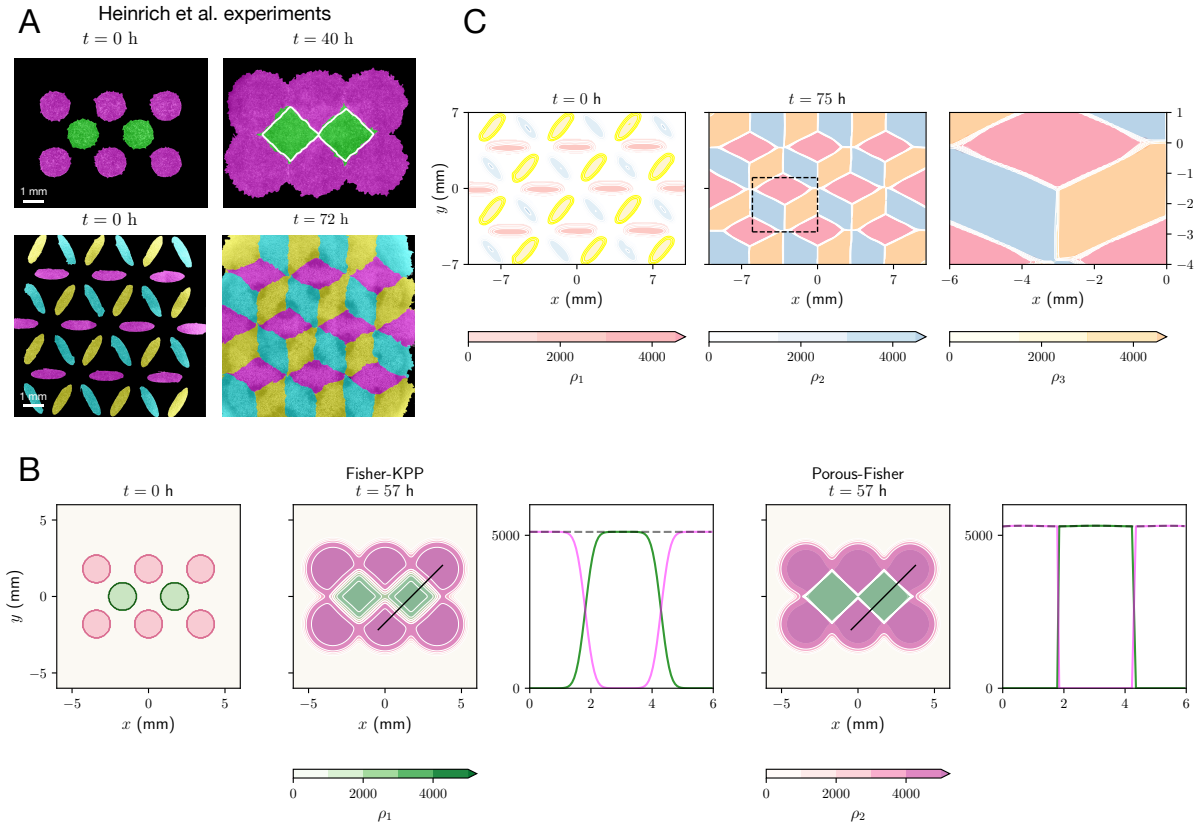


Figure 4.2: Reproducing complex tissue collisions observed in *Heinrich et al.* experiments — animated movies available at Figshare repository. The Fisher–KPP model cannot reproduce complex multi-tissue collisions. (A) Experimental multi-tissue collisions, adapted from [114] (Creative Commons License) (B) Multi-tissue collision between eight homotypic circles for both the Fisher–KPP (4.1) and the Porous–Fisher (4.2) models. Density profiles are taken along the black solid lines. Black dashed lines represent the total population density, $\rho_1 + \rho_2$. Note that numerical simulations use parameter estimates obtained from different experiments — posterior modes in Figures 2.3 and 2.4. (C) Tri-tissue tesselation inspired by Escher's artwork and reproduced experimentally also by *Heinrich et al.* [114]. Here we show numerical simulations of the Porous–Fisher model. Rightmost panel zooms in the region indicated in the middle panel, and shows sharp boundaries. Model parameters correspond to posterior modes (Figure 2.4).

The Porous–Fisher model (4.2) can thus predict the behaviour observed in complex experimental settings with multiple tissues colliding. A numerical simulation of an extension of (4.2) to three species is depicted also in Figure 4.2. This last experiment mimics the self-assembly of a tri-tissue composite designed in [114].

4.2.2 Quantifying collisions between rectangular tissues

As mentioned earlier, collisions between two rectangular tissues result in the formation of a sharp boundary. Whenever the two rectangles are identical — i.e. have the same shape and density — the tissue boundary does not move and coincides with the centroid of the combined tissue. However, *Heinrich et al.* observed that using larger or denser tissues results in a boundary displacement in the direction of the smaller or less dense tissue — see Figure 4.3A for their experimental data. As we will show later, the Porous–Fisher model also predicts a boundary displacement when there is a width/density mismatch between the initial tissues.

Here, we focus on the Porous–Fisher model, and explore to what extent it can reproduce the observed experimental data. In order to perform a quantitative comparison of model and experiments, we calibrate again Eqs. (4.2) by using the data corresponding to a collision between identical rectangular tissues (control case in Figure 4.3A). After carrying out parameter estimation, we explore how the model performs in collisions of rectangular tissues with relative mismatches in either the width or number of cells (density and width mismatch in Figure 4.3A). For simplicity, and after having determined that our model is practically identifiable, we estimate the parameters using a maximum likelihood approach, as explained in previous sections, by comparing experimental and simulated cell densities. The initial densities are taken from experimental data, which in the control case (rectangles with equal density and equal width) are identical to those in Figure 4.1A.

For this set of experiments, the maximum likelihood estimate yields $(D, r, K) = (3.26\ \mu\text{m}^2/(\text{cells}\cdot\text{h}), 0.11\ \text{h}^{-1}, 4077\ \text{cells}/\text{mm}^2)$, which gives an approximate cell radius of $\sim 8\ \mu$m. Observe that the diffusion parameter D and the proliferation rate r show notable differences with respect to the previous set of experiments. In particular these parameters suggest faster migration and slower proliferation, while the front speed remains more or less constant with respect to the case of a single tissue expansion. Note however, that as we proved in the previous sections, the Porous–Fisher model is practically identifiable and hence, although different parameter combinations result in the same invasion speed, we can confidently identify a set of parameters which maximises the likelihood of observing our data.

In fact, the differences in the parameters estimated from the two experiments [112, 114] could be explained by accounting for the transient regime that occurs immediately after the stencil removal. This short timescale is estimated to last around 6-10 hours, which we remove in order to calibrate the model. However, if we only take into account the first 20 hours of the experiment, the maximum likelihood procedure yields very different estimates for the model parameters, which suggests that the experimental collision time could be smaller than this transient timescale. A similar behaviour has been reported for scratch assays with a cancer cell line [121].

After the model is calibrated using the control case data, we can simulate Eqs. (4.2) under different settings by varying the initial conditions. We study collisions between two rectangular tissues with an initial density (2600 vs 1800 cells/mm^2) or width (1000 vs 500 μm) mismatch. In Figure 4.3B we plot the density profiles obtained from the numerical simulations, which show an excellent agreement with the experimental data once both tissues have collided. At early times however, and in line with our discussion above, the model cannot reproduce the observed experimental dynamics. In particular tissue–tissue collisions occur around eight hours before they are observed in the experiments. The agreement between model and data becomes more evident upon visualising individual snapshots from these density profiles (see Figure 4.3C). Note that here, in the numerical simulation of both the density and width mismatch cases, we use the parameters estimated from a collision between identical rectangles.

4.2.3 Population pressure gradients drive boundary displacement

As discussed earlier, the Porous–Fisher model produces a sharp boundary separating the two colliding tissues. When the two tissues are not identical, there is a population pressure gradient at this boundary, which yields a net displacement with velocity $\mathbf{v} = -\nabla P(\rho)$. The nonlinear diffusion model assumes $P(\rho) \sim \rho$ and thus the boundary will move in the direction of the less crowded tissue. This translates, of course, into a wider tissue pushing a more narrow one, or a denser tissue pushing a less dense one.

Our numerical simulations also reveal this behaviour (Figure 4.3D), giving a larger boundary displacement the larger the width/density mismatch. When using parameters inferred from the control case, the total boundary displacement that the model pre-

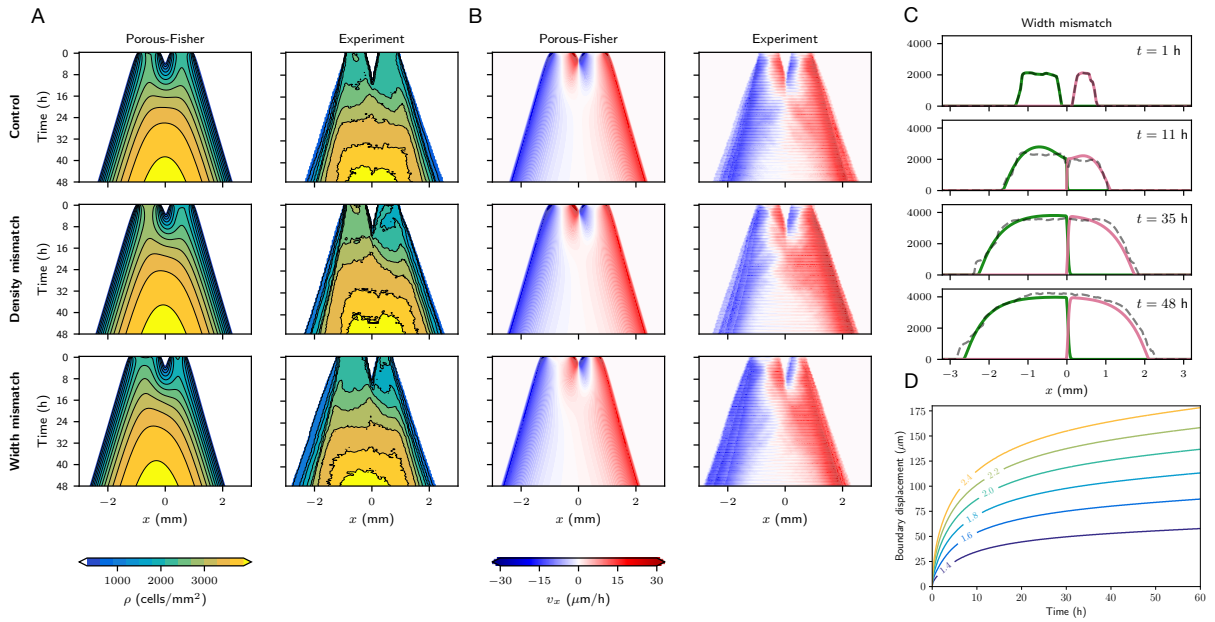


Figure 4.3: Quantifying rectangle–rectangle collisions. (A) Experimental density and velocity profiles resulting from rectangle–rectangle collisions — data provided by *Heinrich et al.* [114]. (B) Numerical simulations of the Porous–Fisher model (4.2) with parameters $(D, r, K) = (3.26\ \mu\text{m}^2/(\text{cells}\cdot\text{h}),\ 0.11\ \text{h}^{-1},\ 4077\ \text{cells}/\text{mm}^2)$. Initial conditions taken from the experimental data. (C) Comparing experimental density (dashed) with model prediction (solid) for the width mismatch case. Note that although parameters are estimated from the control case, there is an excellent agreement between model and data when the initial conditions are modified. (D) Boundary displacement predicted by the model, for different density mismatches, labelled in different colours.

dicts falls short with respect to the experimental measures of *Heinrich et al.* by around 60-100 μm [114], which accounts for less than 5% of the final tissue width after collision. We believe that uncertainty associated with the experimental measures might have a minor impact on these results, as the boundary location can be determined experimentally up to subcellular accuracy and is then averaged over the collision axis — given that different parts of the tissue might not collide at the same time. However, the transient behaviour that cells exhibit after stencil removal can have a more significant effect on the later dynamics [121, 122], especially if this timescale is of the order of the collision time.

Another aspect which could have a more important influence on tissue boundary dynamics from the modelling perspective is the choice of the pressure function $P(\rho)$. In the Porous–Fisher model, cells move following population pressure gradients, mov-

ing away from crowded regions with a pressure function that is assumed to depend linearly on the density. However, using a more general pressure function would also give similar qualitative results but with possibly different dynamics. Note that a logarithmic dependence $P(\rho) \sim \log \rho$ [114] is not suitable for this problem as it corresponds to the case of random cell movement in which there is no sharp boundary separating the tissues.

More generally, one could consider pressure functions that grow as a power-law function of the density, $P(\rho) \sim \rho^{m-1}$ for $m > 1$. For large values of the exponent m, cells only move when the density gradient is large, while in the limit $m \to 1$ we recover the linear diffusion case. Considering this pressure-density relationship yields a porous-medium equation with proliferation for the evolution of the density, which also produces sharp boundaries between colliding tissues for $m > 1$. Hence one could ask how does boundary displacement depend on the relationship between pressure and density — i.e. on the exponent m. In the next section, we study this dependence analytically for small or no proliferation. Beyond analytical insights, we believe that it would be interesting to explore the practical identifiability of the exponent m, and whether considering a more general pressure-density relationship could give more accurate tissue boundary dynamics [144].

4.2.4 Boundary displacement for cross-diffusion systems

We start by assuming that there is no proliferation, and hence the equations governing the dynamics of the two tissues read

$$\begin{cases} \partial_t \rho_1 = D\, \nabla \cdot (\rho_1 \nabla (\rho_1 + \rho_2)) \,, \\ \partial_t \rho_2 = D\, \nabla \cdot (\rho_2 \nabla (\rho_1 + \rho_2)) \,. \end{cases}$$

The total population density, $\rho = \rho_1 + \rho_2$, satisfies then a porous medium equation with exponent two

$$\partial_t \rho = D\, \nabla \cdot (\rho \nabla \rho) = \frac{D}{2}\, \Delta \rho^2.$$

Note that this argument also works for more general versions of the population pressure, which in this case was assumed to grow linearly with the total population density: $P(\rho) \sim \rho$. More generally, if we assume that the pressure inside each tissue increases

as a power-law function of the total population density, $P(\rho) \sim \rho^{m-1}$, we obtain the cross-diffusion system

$$\begin{cases} \partial_t \rho_1 = D\,\nabla \cdot \Big(\rho_1 \left(\rho_1 + \rho_2\right)^{m-2} \nabla \left(\rho_1 + \rho_2\right) \Big) , \\ \partial_t \rho_2 = D\,\nabla \cdot \Big(\rho_2 \left(\rho_1 + \rho_2\right)^{m-2} \nabla \left(\rho_1 + \rho_2\right) \Big) . \end{cases}$$

The resulting equation for the total population density $\rho = \rho_1 + \rho_2$ is now a porous-medium equation with exponent m

$$\partial_t \rho = D\,\nabla \cdot \Big(\rho^{m-1} \nabla \rho \Big) = \frac{D}{m}\,\Delta \rho^m .$$

Solutions to the porous-medium equation in the whole space tend to the self-similar solution given by the Barenblatt profile [57], with known explicit expressions [212]. For simplicity we deal with the one-dimensional case, although the argument works similarly in higher spatial dimensions. In that case, the solution[1] $\rho(x,t)$ tends as $t \to \infty$ to

$$\rho(x,t) = t^{-\alpha} \psi \left(|x|\, t^{-\alpha} \right) ,$$

where $\alpha = 1/(m+1)$ and

$$\psi(y) = \Big(C - \kappa z^2 \Big)^{1/(m-1)} , \quad |y| < \sqrt{\frac{C}{\kappa}},$$

with $\kappa = \alpha(m-1)/2D$ and the constant C determined by conservation of mass. Note that ρ propagates with a free boundary whose radius is given by $r(t) = \sqrt{C/\kappa}\, t^{\alpha}$. Imposing conservation of mass $M = \int_{-r(t)}^{r(t)} \rho(x,t)\, \mathrm{d}x$ yields an expression for the constant C

$$C^{1/(m-1)} \sqrt{\frac{C}{\kappa}} = \frac{\Gamma\left(\frac{1}{m-1} + \frac{3}{2}\right)}{\Gamma\left(\frac{m}{m-1}\right)} \frac{M}{\sqrt{\pi}}.$$

4.2.4.1 Population pressure and boundary displacement

We have seen that when the pressure is given by a power-law function of the density — $P(\rho) \sim \rho^{m-1}$ — the total population density ρ is described by a porous-medium equation with exponent m, whose asymptotic solution is known. Further, we also know that, due to the population pressure term, the two species do not mix but stay segregated [21, 52]. This means that there is an interface separating them whose position

[1] Assuming a proper rescaling of time and space to fix the constant in the porous medium equation

is given by $b(t)$. In order to find the position of this interface we just need to impose mass conservation for one of the tissues: $M_2 = \int_{b(t)}^{r(t)} \sigma(x,t)\,\mathrm{d}x$ — where the masses of the two colliding tissues are given by M_1 and M_2. By doing so, we obtain

$$M_2 = \int_{b(t)}^{r(t)} \sigma(x,t)\,\mathrm{d}x = t^{-\alpha} C^{1/(m-1)} \int_{b(t)}^{r(t)} \left(1 - \frac{\kappa x^2}{C t^{2\alpha}}\right)^{1/(m-1)} \mathrm{d}x$$
$$= \frac{\Gamma\left(\frac{1}{m-1} + \frac{3}{2}\right)}{\Gamma\left(\frac{m}{m-1}\right)} \frac{M}{\sqrt{\pi}} \int_{b(t)/r(t)}^{1} \left(1 - y^2\right)^{1/(m-1)} \mathrm{d}y\,. \tag{4.3}$$

Eq. (4.3) predicts the behaviour of $b(t)$. By noting that the integral cannot depend on time, we obtain that $b(t)$ is a fraction of the expansion radius, $b(t)/r(t) = \ell < 1$, and hence $b(t)$ grows at the same rate as $r(t)$

$$b(t) \sim t^{\alpha}, \quad \alpha = \frac{1}{m+1}\,.$$

Note that this equation relates the form of the equation of state — $P(\rho) \sim \rho^{m-1}$ — with the boundary displacement, which can be measured experimentally.

The ratio ℓ can be found by solving the relation

$$\frac{2M_2}{M} = \frac{\int_{\ell}^{1} \left(1 - y^2\right)^{1/(m-1)} \mathrm{d}y}{\int_{0}^{1} \left(1 - y^2\right)^{1/(m-1)} \mathrm{d}y},$$

or in terms of the Beta function

$$\frac{M_1 - M_2}{M} = \frac{\mathrm{B}\left(\ell^2; 1/2, \frac{m}{m-1}\right)}{\mathrm{B}\left(1/2, \frac{m}{m-1}\right)},$$

where $\mathrm{B}(\ell^2\,; a, b)$ is the incomplete Beta function evaluated at ℓ^2. It is clear then that if $M_2 = M_1 = M/2$, then $\ell = 0$ and that $M_2 < M_1$ (density or width mismatch) allows for boundary displacement with $0 < \ell < 1$.

4.2.4.2 Boundary displacement in the presence of proliferation

In a setting where the cell division time is smaller than the timescale of the experiment one can consider a system equivalent to the ones in the previous section but with a linear proliferation term

$$\begin{cases} \partial_t \rho_1 = D\,\nabla \cdot \left(\rho_1 \left(\rho_1 + \rho_2\right)^{m-2} \nabla \left(\rho_1 + \rho_2\right)\right) + r\rho_1\,, \\ \partial_t \rho_2 = D\,\nabla \cdot \left(\rho_2 \left(\rho_1 + \rho_2\right)^{m-2} \nabla \left(\rho_1 + \rho_2\right)\right) + r\rho_2. \end{cases}$$

The equation for the total population density $\rho = \rho_1 + \rho_2$ reads

$$\partial_t \rho = \frac{D}{m} \Delta \rho^m + r\rho.$$

Under the change of variables $\rho = \tilde{\rho} e^{rt}$ and $\tau = (e^{r(m-1)t} - 1)/r(m-1)$ [108], the equation for the total population density reduces to

$$\partial_\tau \tilde{\rho} = \frac{D}{m} \Delta \tilde{\rho}^m.$$

Now the boundary satisfies

$$b(t) \sim \tau^\alpha \sim \left(e^{r(m-1)t} - 1 \right)^\alpha,$$

which shows the same power-law behaviour for $t \ll 1/r(m-1)$, but grows exponentially for $t > 1/r(m-1)$.

Note however that the exponential growth approximation eventually becomes biologically unrealistic as at some point proliferation will become density-limited. At this point, the densities at both tissues are uniform, driving the boundary to slow down until it finally stops moving.

4.3 Discussion and outlook

Our analysis establishes the Porous–Fisher framework as the minimal continuum description that reconciles single-tissue growth with the far richer phenomenology of tissue–tissue collisions. By fitting a single parameter set to a set of collision assays we demonstrated that the model predicts both transient interface kinetics and the eventual tessellation of multi-tissue composites. The fact that these predictions hold across widely differing initial geometries underscores the central role of density-dependent motility when tissues meet and interact via mechanical forces.

From a modelling perspective we have proposed a systematic method that is based on spatially resolved cell densities. This extension moves beyond the heuristic front-tracking method of *Heinrich et al.* [114], which was limited to equal-density tissues and simple shapes, and could not interrogate the internal structure of each monolayer. In contrast, our approach yields time-series of full density fields, thereby supplying the extra constraints needed to infer all three Porous–Fisher parameters and to validate the

model under a much wider array of experimental conditions. Such quantitative power is a prerequisite for the design and assembly of tissue composites, where one wishes to predict how prescribed initial arrangements will self-organise.

While the Porous–Fisher equation captures the essential crowding feedback, it abstracts away many mechanisms that are known to modulate collective migration. A natural next step is therefore to enrich the pressure term or to add advective stresses that encode *cell–cell adhesion* [56, 87], or *heterogeneity in cell size* [126]. Each extension is biologically motivated and, in principle, straightforward to implement; yet whether the resulting parameters are *structurally* or even *practically* identifiable is far from obvious. Indeed, recent work has shown that models only marginally more complex than ours can already suffer severe non-identifiability [188]. Future progress will therefore hinge on combining richer mechanistic hypotheses with rigorous model-selection and identifiability analyses.

Chapter 5

A local continuum model of cell–cell adhesion

Having shown how density-dependent proliferation and large-scale population pressure shape collective cell behaviours, we now ask how cell–cell adhesion functions as a fundamental mechanism of cellular self-organisation. Differential adhesion between distinct cell types serves as our guiding paradigm: it has a rich catalogue of sorting patterns, and provides a simple test for any mechanistic theory. By taking a short-range interactions limit in earlier non-local models, we obtain a local model of cell–cell adhesion, preserving the sorting phenomenology while reducing model complexity dimensionality. This simplification makes the model easier to implement numerically and more amenable to calibration to quantitative data. This chapter is based on

"C. Falcó, R. E. Baker & J. A. Carrillo. *A local continuum model of cell–cell adhesion.* SIAM Journal on Applied Mathematics, 84(3):S17–S42, 2024" [87],

recently highlighted as a SIGEST article in

"C. Falcó, R. E. Baker & J. A. Carrillo. *A nonlocal-to-local approach to aggregation–diffusion equations.* SIAM Review, 67(2):353–372, 2025" [88].

My contributions involve formulating and analysing the model, implementing the numerical scheme, and writing the manuscript.

5.1 Introduction

From the formulation of the differential adhesion hypothesis (DAH) by Malcolm Steinberg more than 50 years ago [194, 195, 196, 197], there have been many experimental and modelling efforts to understand adhesion-based cellular self-organisation. Differential adhesion between cell populations is now understood as a fundamental mechanism for the formation of tissue shapes during morphogenesis, maintenance and repair, as it allows cells to sort and arrange themselves into complex patterns [109, 198, 207, 208]. On the other hand, mathematical models of differential adhesion not only provide helpful insights into experimental work but have also proven to be interesting from an analytical point of view, motivating a number of further theoretical studies.

The experimental evidence of adhesion-based sorting provided by Townes and Holtfreter [119, 206] inspired Steinberg to formulate the DAH [194, 195, 196, 197]. These experiments showed that dissociated cell populations from amphibian embryos could self-organise and arrange themselves into a pattern with two distinguished cell types. Based on the analogy that cells behave as immiscible fluids with different surface tensions, just like oil and water, Steinberg developed a thermodynamic model that could explain this behaviour in terms of the relative strengths of cell–cell adhesion bonds. Under this setting then, when different cell populations are mixed, they self-organise in order to minimise the total adhesion energy of the tissue. This framework has been able to explain for instance the bullseye pattern seen in multiple experiments [76, 96, 130], in which two cell populations sort into two concentric spheroids, with the most adherent cells comprising the inner one — see Figure 5.1. The DAH and further experimental evidence supporting it are reviewed in [95] —- see also [208] for a more modern perspective.

In this chapter, we present a local continuum model of aggregation–diffusion phenomena, where interactions between species are determined by surface tension parameters. Specifically, in the context of differential cell–cell adhesion, the local model reproduces the different patterns described by the differential adhesion hypothesis, in terms of differential surface tension, as originally proposed by Steinberg. This local model is derived from a general nonlocal model in the limit of short-range interactions and is expressed through a system of thin-film type equations. In its simplified form,

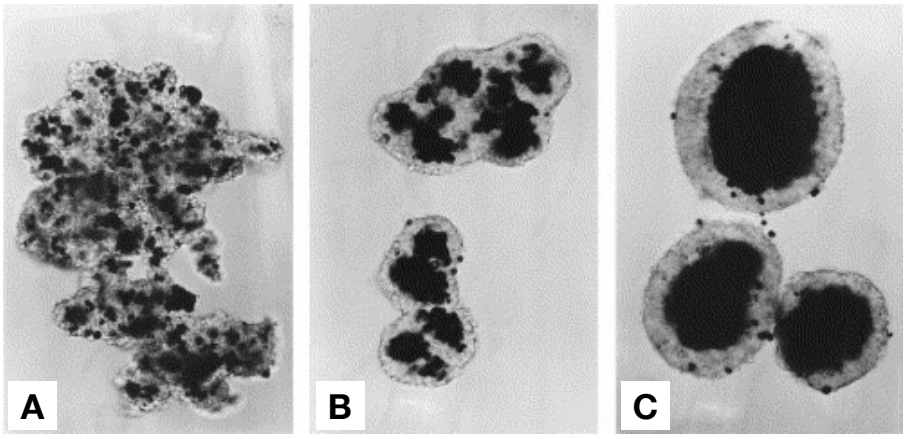


Figure 5.1: Illustrative cell sorting pattern observed in experiments mixing chick embryo neural retina cells (light) and pigmented retinal epithelial cells (dark). (A) After 5 h, cells form mixed aggregates. (B) After 19 h the pigmented retinal cells are almost engulfed by the neural retinal cells, which are located in the interior of the aggregates. (c) After two days the mixture of cell population has formed a bullseye/engulfment pattern. Adapted from [7] (with permission).

the model has four parameters, each of which has a clear physical interpretation, both in the context of nonlocal models and as surface tensions in thin-film equations [164]. Additionally, the local model offers the advantage of being more analytically tractable than nonlocal models, providing explicit stationary solutions even for two interacting species.

5.1.1 Nonlocal models of cell–cell adhesion

Mathematical models describing cell–cell adhesion have taken different approaches by considering either interfacial energy contributions, a tissue bulk modulus, or short–range attraction in the form of nonlocal interactions [2]. Many individual-based models have been used for adhesion–based patterning, e.g. cellular Potts models [116, 130], vertex models [3, 111], and particle–based models [46, 215], to name but a few. While these have been successful in reproducing the observed experimental patterns, discrete models present difficulties, namely the computational cost involved in simulating them and the lack of analytic insights for large numbers of cells. Continuum models, in principle, can offer a solution to these issues, but it was not until two decades ago that this was achieved in the context of cell–cell adhesion.

The first continuum model capable of reproducing cell sorting phenomena was proposed by *Armstrong et al.* [6]. This model combines random motion, resulting in linear diffusion, with cell–cell adhesion, represented by a nonlocal attractive term. Cell–cell adhesion is modelled by assuming cells within a certain distance are attracted to each other. However, one limitation of this approach is that it fails to predict full segregation or sharp boundaries, as it assumes random motion. To overcome this, the linear diffusion term can be replaced with a density-dependent diffusion term that accounts for population pressure [162]. These nonlinear diffusion models are commonly used to describe crowding effects in mathematical biology [41, 77, 108] and have been derived from individual-based models [78] as well as on-lattice models [15, 86]. The modified model, along with variations incorporating density-limited mobilities [56], has proven more effective in capturing the behaviour of adhesion-based pattern formation.

For the sake of conciseness here we do not explicitly derive the mentioned nonlocal models, but we refer to [62] for further and more detailed explanations. A fairly general nonlocal model related to the ones above can be derived as the thermodynamic limit of a system of interacting particles [56]. In this model, cells interact with other cells via a strong repulsion at short distances due to the volume-filling effect of the cell nuclei, and by attraction at medium distances. The latter is linked to the size of the cell and its protrusions or filopodia, and represents adhesive forces. We refer to Chapter 1 for a discussion on how to derive this nonlocal equation as the mean-field limit of an interacting particle system [166].

In the case of differential adhesion between two populations with densities ρ and η [56], we can apply similar ideas to derive the two-species model

$$\begin{cases} \partial_t \rho = \nabla \cdot (\rho \nabla (W_{11} * \rho + W_{12} * \eta + \epsilon(\rho + \eta))) , \\ \partial_t \eta = \nabla \cdot (\eta \nabla (W_{21} * \rho + W_{22} * \eta + \epsilon(\rho + \eta))) , \end{cases} \tag{5.1}$$

where W_{11} and W_{22} represent the self-adhesion potentials, while W_{12} and W_{21} describe the cross-adhesion interactions. The parameter $\epsilon > 0$ quantifies the strength of the localised repulsion. The existence of solutions for this system is established in [97], with further details in [73] for the case without cross-diffusion. A common assumption is that the cross-interaction is symmetric, i.e., $W_{12} = W_{21}$, and that the potentials share the same functional form: $W_{ij} = K_{ij} W$, where W is a typical attractive potential and the

$K_{ij} \geq 0$ are constants representing the cell–cell adhesion strengths. This assumption is consistent with previous nonlocal models of cell–cell adhesion [6, 56, 162]. Although these models have been successful in understanding the underlying biology, they are mathematically complex, reducing analytical insight and limiting the extent to which they can be calibrated and used to describe experimental data.

5.1.2 Outline

In this chapter, we follow the approach in [17] to derive a local model of cell–cell adhesion from Eqs. (5.1). However, the main focus of this paper is not to compare local and nonlocal models, but to study the local model and investigate its consistency with adhesion-based pattern formation and the differential adhesion hypothesis. By formally taking the limit of short-range interactions in the general nonlocal model described by Eqs. (5.1), we obtain a system of thin-film-like equations that model the evolution of the two cell populations

$$\begin{cases} \partial_t \rho &= -\nabla \cdot (\rho \nabla (\kappa \Delta \rho + \alpha \Delta \eta + \mu \rho + \omega \eta)) , \\ \partial_t \eta &= -\nabla \cdot (\eta \nabla (\alpha \Delta \rho + \Delta \eta + \omega \rho + \eta)) . \end{cases} \tag{5.2}$$

The parameters in the system, $\kappa, \alpha, \mu \geq 0$, $\omega \in \mathbb{R}$, can be related to the potentials of the nonlocal model, W_{ij}, and to the strength of the volume-filling mechanism, but can also be understood as relative surface tensions, as in the thin-film equation. In this setting, one could ask whether differential tension — analogous to differential adhesion — in the model is sufficient to give rise to the patterns seen in the Steinberg experiments. Interestingly, we show that it is possible to identify parameter regimes for each one of the different observed configurations (Figure 5.4) with the cross-interaction parameters α and ω playing a major role in the behavior of the local model (see Figure 5.5).

This chapter is structured in two parts. First we derive and study the local model for one cell population, including linear stability, numerical simulations of the model in one and two dimensions, and the calculation of steady states and associated energy minimisers. Then, we extend these ideas and derive the model for two interacting cell populations, Eqs. (5.2). We show via numerical simulations that we can recover the patterns predicted by the DAH. Again, in the local model for two-species, explicit

stationary solutions are available and offer a direct interpretation for cell sorting phenomena. Finally, we summarise our findings and discuss some other advantages of the local model compared to previously used nonlocal models.

5.2 One-species model

We begin with the nonlocal model derived in Chapter 1

$$\partial_t \rho = \nabla \cdot (\rho \nabla (\epsilon \rho + W * \rho)) \ . \tag{5.3}$$

Recall that the $\epsilon\rho$ term represents a localised repulsive force at the origin and the potential W is assumed to be purely attractive and radially symmetric.

Many current models of adhesion only take into account interactions between cells that are separated by less than a maximum *sensing radius*. Here, we build on the idea that for large populations, the *sensing radius* is much smaller than the typical length scale of the population and hence attractive forces between cells are given by a short-range interaction potential. Hence, we set $W(\mathbf{x}) = -a^{-d}\varphi(\mathbf{x}/a)$ with a a scaling parameter which dictates the range of attraction, and φ a sufficiently smooth function defined in $\mathbb{R}^d$. As $a \to 0$, the potential W tends to a Dirac delta function supported at the origin. We further assume that the function φ satisfies several conditions:

1. $\varphi(\mathbf{x}) = \varphi(|\mathbf{x}|)$ and $\varphi(r)$ is a non-increasing function for $r > 0$, meaning that W is both symmetric and attractive.

2. φ approaches a constant as $r \to \infty$. Without loss of generality we assume that this constant is zero.

3. The moments $M_n = \int_{\mathbb{R}^d} |\mathbf{x}|^n \varphi(\mathbf{x}) \, \mathrm{d}\mathbf{x}$ decay suitably fast.

Omitting the time dependence and using the change of variable

$$(W * \rho)(\mathbf{x}) = -a^{-d} \int_{\mathbb{R}^d} \varphi\left(\frac{\mathbf{y}}{a}\right) \rho(\mathbf{x} - \mathbf{y}) \, \mathrm{d}\mathbf{y} = -\int_{\mathbb{R}^d} \varphi(\mathbf{y}) \rho(\mathbf{x} - a\mathbf{y}) \, \mathrm{d}\mathbf{y},$$

we can consider the limit of short-range attraction and expand $\rho(\mathbf{x} - a\mathbf{y})$ as a Taylor series for small values of the scaling parameter a:

$$\begin{aligned}(W * \rho)(\mathbf{x}) = &- \rho(\mathbf{x}) \int_{\mathbb{R}^d} \varphi(\mathbf{y})\, d\mathbf{y} + a \int_{\mathbb{R}^d} (\nabla\rho(\mathbf{x}) \cdot \mathbf{y})\, \varphi(\mathbf{y})\, d\mathbf{y} \\ &- \frac{a^2}{2} \int_{\mathbb{R}^d} \left(\mathbf{y}^T \cdot H_\rho(\mathbf{x})\mathbf{y}\right) \varphi(\mathbf{y})\, d\mathbf{y} + o(a^2);\end{aligned}$$

where $H_\rho(\mathbf{x})$ is the Hessian matrix of ρ.

We will only keep the first two non-zero terms in the expansion as these should give a good approximation for short-range interactions [82]. For the first term in the Taylor expansion we simply have $\rho \int_{\mathbb{R}^d} \varphi = M_0\rho$, and we also note that the terms with odd order derivatives of ρ vanish due to the symmetry assumption on the potential. Then the error term in the expression above is $O(a^4)$. The next non-vanishing term in the series contains the second-order derivatives of ρ and reads

$$\begin{aligned}\int_{\mathbb{R}^d} \left(\mathbf{y}^t \cdot H_\rho(\mathbf{x})\mathbf{y}\right) \varphi(\mathbf{y})\, d\mathbf{y} &= \sum_{i=1}^{d}\sum_{j=1}^{d} \frac{\partial^2\rho}{\partial x_i \partial x_j} \int_{\mathbb{R}^d} y_i y_j \varphi(\mathbf{y})\, d\mathbf{y} \\ &= \sum_{i=1}^{d} \frac{\partial^2\rho}{\partial x_i^2} \int_{\mathbb{R}^d} y_i^2 \varphi(\mathbf{y})\, d\mathbf{y} \\ &= \frac{1}{d}\left(\int_{\mathbb{R}^d} |\mathbf{y}|^2 \varphi(\mathbf{y})\, d\mathbf{y}\right) \sum_{i=1}^{d} \frac{\partial^2\rho}{\partial x_i^2} \\ &= \frac{M_2}{d}\Delta\rho,\end{aligned}$$

where we used again that φ is symmetric. Putting this all together gives

$$W * \rho = -M_0\rho - \frac{M_2 a^2}{2d}\Delta\rho + O(a^4 M_4).$$

Using only the first two terms in the approximation in Eq. (5.3) yields

$$\frac{\partial\rho}{\partial t} = -\nabla \cdot \left(\rho\nabla\left(\tilde{M}\Delta\rho + (M_0 - \epsilon)\rho\right)\right), \tag{5.4}$$

with $\tilde{M} = M_2 a^2/2d$. Note that the approximation makes sense as long as the moments M_n of higher order ($n \geq 4$) are small compared to M_2, and the scaling parameter a is small. We emphasize here though, that the goal of this chapter is not to compare (5.4) with the nonlocal model (5.3), but to study possible behaviors of the local model in Eq. (5.4). We refer to [49, 82] for rigorous justification of this limit.

Two relevant observations can be made here. First, note that the sign of $M_0 - \epsilon$ gives the relative strength of repulsive and attractive forces. For $M_0 - \epsilon < 0$, localised repulsion is the dominant interaction, while for $M_0 - \epsilon > 0$, the short-range attractive forces overcome repulsion. Here, we focus on the latter case, since it allows biological aggregation. In fact, with our choice of diffusion and aggregation potential W, Eq. (5.3) only has stationary states in the $M_0 - \epsilon > 0$ case [38]. As we will see, our analysis suggests that this is also the case for the local model given by Eq. (5.4).

Second, and as it was already remarked in [6], the fourth order term has a dampening effect on the PDE. In the absence of this term, one obtains an ill-posed problem due to the negative diffusion coefficient. Therefore, it does not seem possible to have a second-order model of cell–cell adhesion, thus making evident the need for a fourth-order approximation. A similar phenomenon happens in [5] when one takes the continuum limit of a microscopic model incorporating cell–cell adhesion.

Before moving onto further considerations, and in order to facilitate the analysis, we nondimensionalise Eq. (5.4). Under a suitable rescaling — for instance, set $\rho \mapsto \tilde{M}\rho$ and $\mu^2 = (M_0 - \epsilon)/\tilde{M}$ — the model can be written as

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot \left(\rho \nabla \left(\Delta \rho + \mu^2 \rho \right) \right), \tag{5.5}$$

with $\mu^2 > 0$, according to our previous considerations. This model resembles a Cahn–Hilliard [85] or thin-film type equation where the parameter μ^{-2} plays the role of surface tension [164], which reintroduces the idea of modelling tissues as having fluid-like properties, as originally proposed by Steinberg in his model. These considerations will become more relevant later on when we discuss systems of two-species.

The thin-film equation governs the evolution of the thickness of a thin fluid layer on a surface. Equations of the type of Eq. (5.5) appear as the lubrication approximation of a gravity-driven Hele-Shaw cell [103, 104]. Depending on the sign of μ^2, the equation is referred as *long-wave unstable* or *long-wave stable*. The sign of μ^2 characterizes the linear stability of the constant steady state — more details will be discussed later. This model also falls under a larger family of thin-film equations, whose properties have been well-studied — see [137, 138, 139] for an exhaustive study of the steady states, [191] for stability of self-similar solutions, and [19, 20] for long-time behaviour of solutions and regularity.

Associated with the local model, we also have the free energy

$$\mathcal{F}[\rho] = \frac{1}{2}\int_\Omega \left(|\nabla\rho|^2 - \mu^2\rho^2\right) \mathrm{d}\mathbf{x}. \tag{5.6}$$

With this in mind, Eq. (5.5) can be written as a gradient flow with respect to the 2-Wasserstein metric (see for instance [55, 155, 183])

$$\frac{\partial\rho}{\partial t} = \nabla\cdot\left(\rho\,\nabla\frac{\delta\mathcal{F}}{\delta\rho}\right).$$

Observe that by integrating by parts formally,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}[\rho] &= \int_\Omega \nabla\rho\cdot\nabla\left(\frac{\partial\rho}{\partial t}\right)\mathrm{d}\mathbf{x} - \mu^2\int_\Omega \rho\,\frac{\partial\rho}{\partial t}\,\mathrm{d}\mathbf{x}\\ &= \int_\Omega \rho\,\nabla\left(\Delta\rho + \mu^2\rho\right)\cdot\nabla\left(\frac{\delta\mathcal{F}}{\delta\rho}\right)\mathrm{d}\mathbf{x}\\ &= -\int_\Omega \rho\left|\nabla\frac{\delta\mathcal{F}}{\delta\rho}\right|^2\mathrm{d}\mathbf{x} \le 0\,,\end{aligned}$$

and hence the energy is non-decreasing in time. Here we used the first and third order boundary conditions

$$\partial_\nu\rho = \partial_\nu\Delta\rho = 0 \quad \text{on } \partial\Omega, \tag{5.7}$$

where ν is the exterior normal of Ω.

5.2.1 Linear stability analysis

Equation (5.5) admits as steady states any spatially homogeneous solution. One of the first biologically relevant questions that arises from this model, is whether aggregations are possible as in the case of Eq. (5.3). To investigate this question we perform linear stability analysis on the spatially homogeneous solution $\rho(\mathbf{x},t) = \rho_0$. In order to do so, we consider a perturbation $\rho(\mathbf{x},t) = \rho_0 + \tilde{\rho}(\mathbf{x},t)$ and linearize the resulting equation. By setting $\tilde{\rho}(\mathbf{x},t) \propto e^{i\mathbf{k}\cdot\mathbf{x}+\sigma(\mathbf{k})t}$ one finds the dispersion relation

$$\sigma(\mathbf{k}) = \rho_0|\mathbf{k}|^2\left(\mu^2 - |\mathbf{k}|^2\right).$$

In fact, the resulting linearized equation is identical to that of the standard Cahn–Hilliard equation [84]. A necessary condition for the formation of non-trivial stationary states is then $\mathrm{Re}(\sigma(\mathbf{k})) > 0$ for certain values of the wave vector $\mathbf{k}$, which results in the upper bound: $|\mathbf{k}| < \mu$.

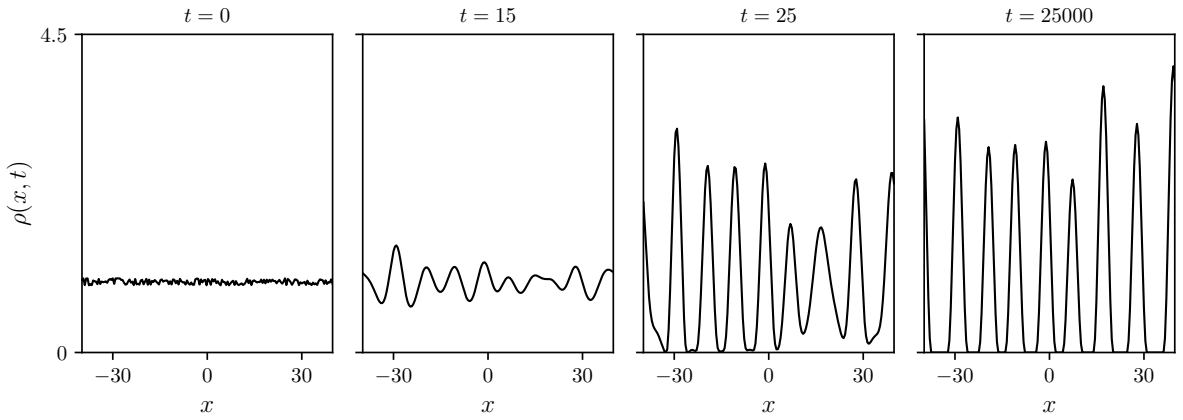


Figure 5.2: Aggregation is possible in the local model as long as $\mu^2 > 0$. Numerical simulations with periodic boundary conditions and parameters: $\mu^2 = 1$, $L = 40$, $\delta x = 0.2$, $\delta t = 0.01$. Initial data corresponds to the spatially homogeneous steady state $\rho(x, 0) = 1$ for $x \in [-L, L]$ plus a small perturbation.

From the instability condition we see that in the case where $\mu^2 < 0$, the homogeneous steady state is linearly stable and thus aggregation is not possible. As mentioned earlier, this case happens when $-\int_{\mathbb{R}^d} W < \epsilon$, for which Eq. (5.3) has no stationary states either [38]. We note that these considerations are not new, as Eq. (5.5) belongs to a broader class of thin-film equations, whose linear stability is well-established [138]. From now on, we always consider the local model in the long-wave unstable regime $\mu^2 > 0$.

5.2.2 Numerical experiments

In this section, we explore numerically some basic properties of the local model. For that purpose we use a numerical scheme based on that in [11], which we briefly outline in Appendix A. Moreover, we run all of our simulations on a domain $[-L, L]$, where L is specified individually for every experiment and we assume periodic boundary conditions on ρ and its derivatives.

We start by showcasing some of the results in the previous section and whether aggregations of cells are possible in this model. As expected from our derivations, the spatially homogeneous steady state is unstable in the case $\mu^2 > 0$. To test this prediction, we use a suitably large domain, $L = 40$, and perform simulations using as initial densities a slightly perturbed homogeneous steady state (Figure 5.2). We see that small

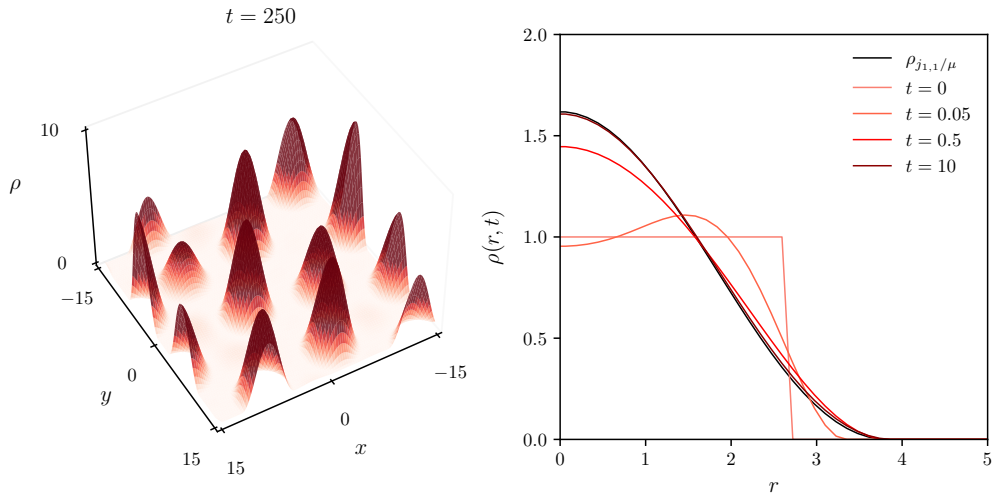


Figure 5.3: (left) Aggregation in the two-dimensional local model. Initial data is $\rho(x, 0) = 1$ plus a small perturbation. Density configuration at $t = 250$. (right) Convergence to steady state in the two-dimensional model. Radial density profiles at different time points and stationary solution given by $\rho_{j_{1,1}/\mu}(r)$. Simulation parameters are $\delta t = 0.01$, $\delta x = 0.1$, $\delta y = 0.1$, $\mu = 1$ and domain specifications: $L = 15$ (left) and $L = 5$ (right).

perturbations rapidly lead to spatial patterning that mimics previous models of cell–cell adhesion [6]. While cell densities change very rapidly at early times, as they evolve towards different peaks, smaller density bumps disappear at a very low rate, so that the system only reaches the stationary configuration after much longer times. In particular, the local model shows similar behaviour to nonlocal models with compactly supported interaction potentials, which usually give rise to stationary states with multiple separated aggregates [45]. The distance separating different cell aggregates is of course larger than the *sensing radius* in the potential. A similar pattern appears in the two-dimensional case (Figure 5.3).

5.2.3 Steady states and energy minimisers

Next, we combine both analytical and numerical insights in order to study the stationary solutions of the local model. Note that steady states of (5.5) satisfy the equation

$$\Delta\rho + \mu^2\rho = C, \tag{5.8}$$

with C a constant that could be different on each connected component of $\mathrm{supp}(\rho)$. Observe that as a result of the gradient flow structure, calculating steady states is equiva-

lent to the problem of finding critical points of the energy, Eq. (5.6), which are given by the condition: $-\delta\mathcal{F}/\delta\rho = C$.

Here, and motivated by our first numerical simulations, we assume that steady states are supported on finite unions of compact sets. In particular, and as we will see later in further numerical experiments, we assume that for smaller domains — and not too disperse initial data — there is only one connected component of $\mathrm{supp}(\rho)$. This is what in the thin-film equation literature is called the *droplet steady state* [191]. We consider the one- and two-dimensional cases separately.

One-dimensional case. In the one-dimensional case, general solutions to Eq. (5.8) read

$$\rho(x) = A\cos(\mu x) + B\sin(\mu x) + \frac{C}{\mu^2}.$$

Imposing that stationary states are both symmetric and invariant under translations, we can without loss of generality set $B = 0$. We now write $\mathrm{supp}(\rho) = [-b, b]$ and use mass conservation to find

$$\rho_b(x) = A\left(\cos(\mu x) - \cos(\mu b)\right), \quad A = \frac{m\mu}{2\left(\sin(\mu b) - \mu b\cos(\mu b)\right)}, \tag{5.9}$$

where $m = \int_\Omega \rho(x,0)\mathrm{d}x$. With this, one finds a family of compactly supported steady states parametrised by b. Analogously, we could parametrise these steady states by their *touchdown angle*, given by $\rho'(b)$. Note that in order to preserve positivity of solutions we need $\mu b \in (0, \pi]$.

However, numerical solutions show that for a given mass m, and in the cases where $\mathrm{supp}(\rho)$ has only one connected component, solutions of Eq. (5.4) tend to a unique steady state. We conjecture here that this steady state corresponds to the energy minimiser. In order to find it, we calculate the energy given by Eq. (5.6) of the family of steady states in Eq. (5.9)

$$\mathcal{F}[\rho_b] = \frac{m^2\mu^3}{2}\frac{\cos(\mu b)}{\sin(\mu b) - \mu b\cos(\mu b)}.$$

Note that $\mathcal{F}[\rho_b]$ is a decreasing function of μb on $(0, \pi]$ and hence the energy minimiser corresponds to the case where $\mu b = \pi$:

$$\rho_{\pi/\mu}(x) = \frac{m\mu}{2\pi}\left(\cos(\mu x) + 1\right), \quad |x| \leq \frac{\pi}{\mu}.$$

Observe here that the minimum of the energy $\mathcal{F}[\rho_b]$ is obtained when the zero contact angle condition $\rho'(b) = 0$ is satisfied, in the same way as in [17]. We emphasise here that while we have been able to identify the steady state with the lowest energy, in general this is a very complex problem — see [139] for an exploration of the energy landscape for a larger family of thin-film type equations.

Two-dimensional case. The calculations here are very similar to the one-dimensional case. Again based on numerical simulations we assume that steady states have radial symmetry and write $r = |\mathbf{x}|$. With this, general solutions $\rho = \rho(r)$ to Eq. (5.8) read

$$\rho(r) = AJ_0(\mu r) + BY_0(\mu r) + \frac{C}{\mu^2},$$

with J_n and Y_n being Bessel functions of the first and second kind, respectively. Imposing regularity at the origin yields $B = 0$ and, once again, assuming that steady states are supported on a closed disk of radius b, we obtain

$$\rho_b(r) = A\left(J_0(\mu r) - J_0(\mu b)\right), \quad A = \frac{m\mu}{\pi b\left(J_1(\mu b) - \mu b J_0(\mu b)\right)}.$$

For positive solutions we need to impose that $\mu b \in (0, j_{1,1}]$ where $j_{1,1} \approx 3.832$ is the first zero of $J_0' = -J_1$. The energy minimisation argument also holds here. Note that

$$\mathcal{F}[\rho_b] = \frac{m^2\mu^4}{\pi}\frac{J_0(\mu b)}{\mu b\left(2J_1(\mu b) - \mu b J_0(\mu b)\right)},$$

is a decreasing function of μb on $(0, j_{1,1}]$, and hence the minimiser corresponds to the case $\mu b = j_{1,1}$ — or equivalently to the solution satisfying $\rho'(b) = 0$. This steady state reads

$$\rho_{j_{1,1}/\mu}(r) = \frac{m\mu^2}{\pi j_{1,1}^2}\left(1 - \frac{J_0(\mu r)}{J_0(j_{1,1})}\right), \quad |x| \leq \frac{j_{1,1}}{\mu},$$

with $J_0(j_{1,1}) \approx -0.403$. In Figure 5.3 we sketch the radial density profiles for simulations in a two-dimensional box of length $2L = 10$.

5.2.4 Contact angle of energy minimisers

Here we have found that the energy minimiser satisfies the zero contact angle condition $\rho'(b) = 0$. In fact, this can be justified using a perturbation argument in one dimension as in [17], and without the need for explicit expressions for the steady states. To see this, assume that ρ is a symmetric solution with compact support given by $\mathrm{supp}(\rho) =$

$[-b, b]$. Now perturb the support $\bar{b} = b + \delta b$, with $\delta b \ll b$ and assume that the solution is also perturbed according to $\bar{\rho} = \rho + \delta\rho$. Note that the condition $\int \bar{\rho} = \int \rho$ requires the perturbation to have zero total mass. If we calculate the free energy of the new solution we obtain up to first order in the perturbation

$$\mathcal{F}[\bar{\rho}] = \mathcal{F}[\rho] + \int_{-b}^{b} \delta\rho \, \frac{\delta \mathcal{F}}{\delta \rho} \, \mathrm{d}x + \delta b \cdot \rho'(b)^2.$$

If we assume that ρ is a minimiser, then the second term becomes zero, as ρ satisfies Eq. (5.8). If $\rho'(b) \neq 0$ then we can find solutions with lower energy, contradicting the fact that ρ is a minimiser. Hence $\rho'(b) = 0$.

These compactly supported, droplet-like steady states satisfy a zero contact angle, in contrast to classical passive wetting where partial wetting typically yields a finite contact angle [209]. By contrast, the corresponding nonlocal model can produce finite contact angles, because the local model truncation omits higher-order derivative contributions that are most important near the contact line. It would also be interesting to study how adhesion modifies wetting in the presence of active forces [106], which are critical for cell migration.

5.3 Extension to two interacting populations

Having analysed the model for a single population, we now extend it to two interacting cell populations. Our main goal here is to study if such model is able to reproduce the patterns seen in the Steinberg experiments, and whether this behaviour can be understood in terms of the model parameters.

In the case of two interacting populations, the self-adhesion of each species and the cross-adhesion between them determine the behaviour of the system. Depending on the relative strength of adhesive forces, experimentally it is seen that the two cell populations may evolve to one of four different configurations, that we represent in Figure 5.4.

5.3.1 A system of thin-film equations to model cell–cell adhesion

We proceed as in the one-species case and assume that the potentials W_{ij} in Eqs. (5.1) are attractive and scale according to a parameter a, which gives the range of interactions.

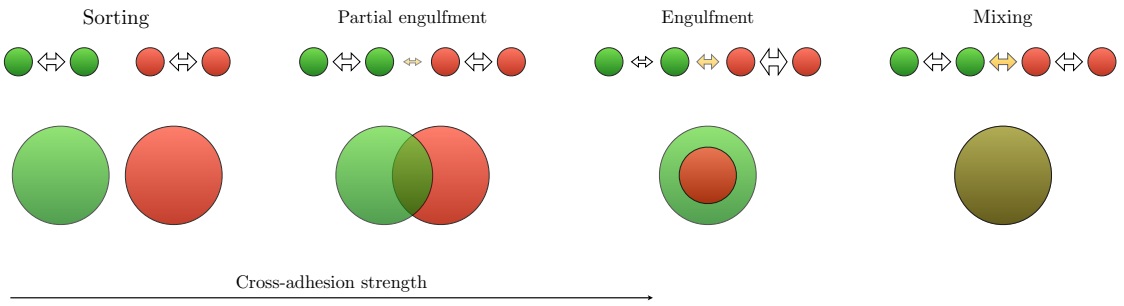


Figure 5.4: Possible configurations for Steinberg experiments in terms of the cross-adhesion and the self-adhesion of a system of two-species (adapted from [162]). In the weak cross-adhesion regime we might have two patterns depending on whether the cross-adhesion strength is strictly zero or positive. **Sorting** is observed when there is no cross-adhesion between the two-species, and **partial engulfment** when cross-adhesion is small compared to the self-adhesion of each population. When the cross-adhesion is stronger, the system might evolve to an **engulfment** pattern, where the more cohesive species is surrounded by the less cohesive one; or to complete **mixing** of the cell populations. The first corresponds to the case in which the cross-adhesion is stronger than the self-adhesion of one-species but weaker than the self-adhesion of the other one. The latter occurs when the cross-adhesion strength is comparable to both self-adhesion forces.

More precisely we impose $W_{ij}(\mathbf{x}) = -a^{-d}\varphi_{ij}(\mathbf{x}/a)$, with the functions φ_{ij} satisfying the conditions in the previous section. In the limit $a \to 0$, we can approximate $W_{ij} * f \approx -c_{ij}f - d_{ij}\Delta f$, where $f \in \{\rho, \eta\}$ and the constants c_{ij}, d_{ij} could be different for each potential. Note that c_{ij} is the volume of φ_{ij} and d_{ij} is related to its second moment

$$c_{ij} = \int_{\mathbb{R}^d} \varphi_{ij}(\mathbf{x})\,\mathrm{d}\mathbf{x}, \quad d_{ij} = \frac{a^2}{2d}\int_{\mathbb{R}^d} |\mathbf{x}|^2\varphi_{ij}(\mathbf{x})\,\mathrm{d}\mathbf{x}.$$

For simplicity we assume here that the cross-interaction potential is the same for the two-species $W_{12} = W_{21}$, which is a commonly used assumption in many models of cell–cell adhesion [6, 56]. Using these approximations in the two-species nonlocal model Eqs. (5.1), yields

$$\begin{cases} \partial_t\rho = -\nabla\cdot\left(\rho\nabla\left(\kappa_1\Delta\rho + \tilde{\alpha}\Delta\eta + \mu_1\rho + \tilde{\omega}\eta\right)\right), \\ \partial_t\eta = -\nabla\cdot\left(\eta\nabla\left(\tilde{\alpha}\Delta\rho + \kappa_2\Delta\eta + \tilde{\omega}\rho + \mu_2\eta\right)\right). \end{cases}$$

The model parameters can be understood in terms of the potentials W_{ij}. First note that the parameters in the fourth order terms, κ_1, κ_2 and $\tilde{\alpha}$, are directly related to the

second moments of the potentials. Hence, they only give information on the strength and range of attractive forces. Assuming that the potentials are all attractive, we have $\kappa_1, \kappa_2, \alpha \geq 0$. On the other hand, the parameters in the second order terms, μ_1, μ_2 and $\tilde{\omega}$, are both related to the volumes of the potentials, and to the strength of repulsive forces, which are given by ϵ. According to our considerations in Section 5.2.1, we assume that $-\int_{\mathbb{R}^d} W_{11} > \epsilon$ and also $-\int_{\mathbb{R}^d} W_{22} > \epsilon$, meaning that self-attraction overcomes repulsion in each of the populations. This gives $\mu_1, \mu_2 > 0$. However, cross-attraction between the two type of cell types could be weaker and thus $\tilde{\omega}$ could be either positive or negative.

In fact, and in order to facilitate the analysis, we can reduce the number of parameters with a suitable rescaling of the variables. For example, set $x \mapsto \xi x$, $t \mapsto Tt$, $\rho \mapsto \mu_2 \rho$, $\eta \mapsto \mu_2 \eta$, with $\xi^2 = \mu_2/\kappa_2$ and $T = \xi^2$. This yields the new rescaled parameters

$$\kappa = \frac{\kappa_1}{\kappa_2}, \quad \alpha = \frac{\tilde{\alpha}}{\kappa_2}, \quad \mu = \frac{\mu_1}{\mu_2}, \quad \omega = \frac{\tilde{\omega}}{\mu_2},$$

and the reduced system

$$\begin{cases} \partial_t \rho &= -\nabla \cdot (\rho \nabla (\kappa \Delta \rho + \alpha \Delta \eta + \mu \rho + \omega \eta)), \\ \partial_t \eta &= -\nabla \cdot (\eta \nabla (\alpha \Delta \rho + \Delta \eta + \omega \rho + \eta)). \end{cases} \tag{5.10}$$

Here, $\kappa \geq 0$ and $\mu > 0$ represent the relative self-adhesion strength of ρ with respect to η; while $\alpha \geq 0$ and $\omega \in \mathbb{R}$ give the relative strength of the cross-attraction forces.

Observe too that the local model, Eqs. (5.10), is essentially a system of two thin-film like equations, where the parameters κ and α can be understood as the relative tensions of one-species with respect to the other one, and of the interface separating the two populations. The parameters in the second order terms μ and ω are then related to the population pressure exerted by each cell type. Under this setting, it makes sense to ask whether differential tension, as originally proposed by Steinberg, can explain cell sorting phenomena. In other words, *can we identify relevant regimes for the four parameters $\kappa, \alpha, \mu, \omega$ such that the experimental patterns are recovered in the local model?* We highlight that the local model given in Eqs. (5.10) is not a phenomenological description emerging from the tissue-fluid analogy, but arises in the limit of short-range interactions of Eqs. (5.1), and hence it provides a direct connection between the original differential adhesion hypothesis and a physical model of cell–cell adhesion.

In order to avoid negative diffusion, we require the matrix

$$M = \begin{pmatrix} \kappa & \alpha \\ \alpha & 1 \end{pmatrix},$$

to be positive definite. Since $\kappa > 0$, this requires $\det M \geq 0$. This sets a limit on the strength of the cross-attraction: $0 \leq \alpha < \sqrt{\kappa}$.

5.3.2 Energy

Thanks to the symmetry in the cross-interaction terms given by α and ω, this system also exhibits a gradient-flow structure

$$\begin{cases} \partial_t \rho = \nabla \cdot \left(\rho \, \nabla \dfrac{\delta \mathcal{F}_2}{\delta \rho} \right), \\ \partial_t \eta = \nabla \cdot \left(\eta \, \nabla \dfrac{\delta \mathcal{F}_2}{\delta \eta} \right), \end{cases}$$

with respect to the 2-Wasserstein metric [51, 55, 155, 183] and the free energy

$$\mathcal{F}_2[\rho, \eta] = \int_\Omega \left(\frac{\kappa}{2} |\nabla \rho|^2 + \frac{1}{2} |\nabla \eta|^2 + \alpha \nabla \rho \cdot \nabla \eta - \frac{\mu}{2} \rho^2 - \frac{1}{2} \eta^2 - \omega \rho \eta \right) \mathrm{d}\mathbf{x}. \tag{5.11}$$

We note that the nonlocal model for two-species, given by Eqs. (5.1), also exhibits a gradient flow structure when the cross-interaction potentials are symmetrisable, providing variational schemes to prove the existence of solutions to the system [51, 73, 97].

As in the one-species case, we can formally integrate by parts to show that the energy is non-increasing in time

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \mathcal{F}_2[\rho, \eta] = {} & \kappa \int_\Omega \nabla \rho \cdot \nabla \left(\frac{\partial \rho}{\partial t} \right) \mathrm{d}\mathbf{x} + \int_\Omega \nabla \eta \cdot \nabla \left(\frac{\partial \eta}{\partial t} \right) \mathrm{d}\mathbf{x} \\ & + \alpha \int_\Omega \nabla \rho \cdot \nabla \left(\frac{\partial \eta}{\partial t} \right) \mathrm{d}\mathbf{x} + \alpha \int_\Omega \nabla \left(\frac{\partial \rho}{\partial t} \right) \cdot \nabla \eta \, \mathrm{d}\mathbf{x} \\ & - \mu \int_\Omega \rho \frac{\partial \rho}{\partial t} \mathrm{d}\mathbf{x} - \int_\Omega \eta \frac{\partial \eta}{\partial t} \mathrm{d}\mathbf{x} - \omega \int_\Omega \rho \frac{\partial \eta}{\partial t} \mathrm{d}\mathbf{x} - \omega \int_\Omega \frac{\partial \rho}{\partial t} \eta \, \mathrm{d}\mathbf{x} \\ = {} & \int_\Omega \rho \, \nabla \left(\kappa \Delta \rho + \alpha \Delta \eta + \mu \rho + \omega \eta \right) \cdot \nabla \left(\frac{\delta \mathcal{F}_2}{\delta \rho} \right) \mathrm{d}\mathbf{x} \\ & + \int_\Omega \eta \, \nabla \left(\alpha \Delta \rho + \Delta \eta + \omega \rho + \eta \right) \cdot \nabla \left(\frac{\delta \mathcal{F}_2}{\delta \eta} \right) \mathrm{d}\mathbf{x} \\ = {} & - \int_\Omega \rho \left| \nabla \frac{\delta \mathcal{F}_2}{\delta \rho} \right|^2 \mathrm{d}\mathbf{x} - \int_\Omega \eta \left| \nabla \frac{\delta \mathcal{F}_2}{\delta \eta} \right|^2 \mathrm{d}\mathbf{x} \leq 0. \end{aligned}$$

Again, we used the boundary conditions on ρ and η given by Eq. (5.7).

5.3.3 Numerical simulations for Steinberg experiments in one dimension

Here we study numerically whether the local model is able to reproduce the four different patterns observed in Steinberg experiments (Figure 5.4), namely: (i) mixing; (ii) engulfment; (iii) partial engulfment; (iv) sorting. To do so, we must determine which parameter ranges correspond to each of the observed patterns. This is simpler when we assume a particular shape for the potentials φ_{ij}. Let us then assume that these only differ by constants, i.e. $\varphi_{ij} = K_{ij}\varphi$, for constants $K_{ij} \geq 0$ satisfying $K_{12} = K_{21}$ and a given potential φ. This is actually the case in previously used nonlocal models [6, 53, 56], where the constants K_{ij} give the adhesive strengths of the two cell populations. Under these assumptions, the model parameters are directly related to the moments of φ and the constants ϵ and K_{ij}:

$$\kappa = \frac{K_{11}}{K_{22}}; \quad \alpha = \frac{K_{12}}{K_{22}}; \quad \mu = \left(\frac{M_0 - \epsilon/K_{11}}{M_0 - \epsilon/K_{22}}\right)\kappa; \quad \omega = \left(\frac{M_0 - \epsilon/K_{12}}{M_0 - \epsilon/K_{22}}\right)\alpha;$$

where M_0 is the volume of the potential φ. Note then that κ and α can be interpreted as relative adhesion strengths, as mentioned earlier. However, μ and ω are not only related to cell–cell adhesion but also to the strength of local repulsion due to volume exclusion.

We also assume without loss of generality that η is the less cohesive population and hence $K_{22} < K_{11}$. Then according to these expressions we have $\mu > \kappa > 1$. Parameter values outside of this range are also valid but interpreting the model in such cases becomes more challenging. For the cross-interaction parameters α and ω, one must distinguish two regimes depending on K_{12}, i.e. the strength of the cross-adhesion:

1. Weak cross-adhesion ($K_{12} < K_{22}$). In this case, we have $\omega < \alpha < 1$, as shown by the blue-shaded region in Figure 5.5. As the cross-adhesion strength decreases, ω becomes more negative, and α decreases as well.

2. Strong cross-adhesion ($K_{12} > K_{22}$). Conversely, when the cross-adhesion is stronger than the self-adhesion of the second cell type, we have $\omega > \alpha > 1$, as depicted in the red-shaded region in Figure 5.5. In this case, both α and ω increase with the cross-adhesion strength. Additionally, the quotient ω/α increases with K_{12}, meaning that in the limit of strong cross-adhesion, we should expect $\omega \gg \alpha$.

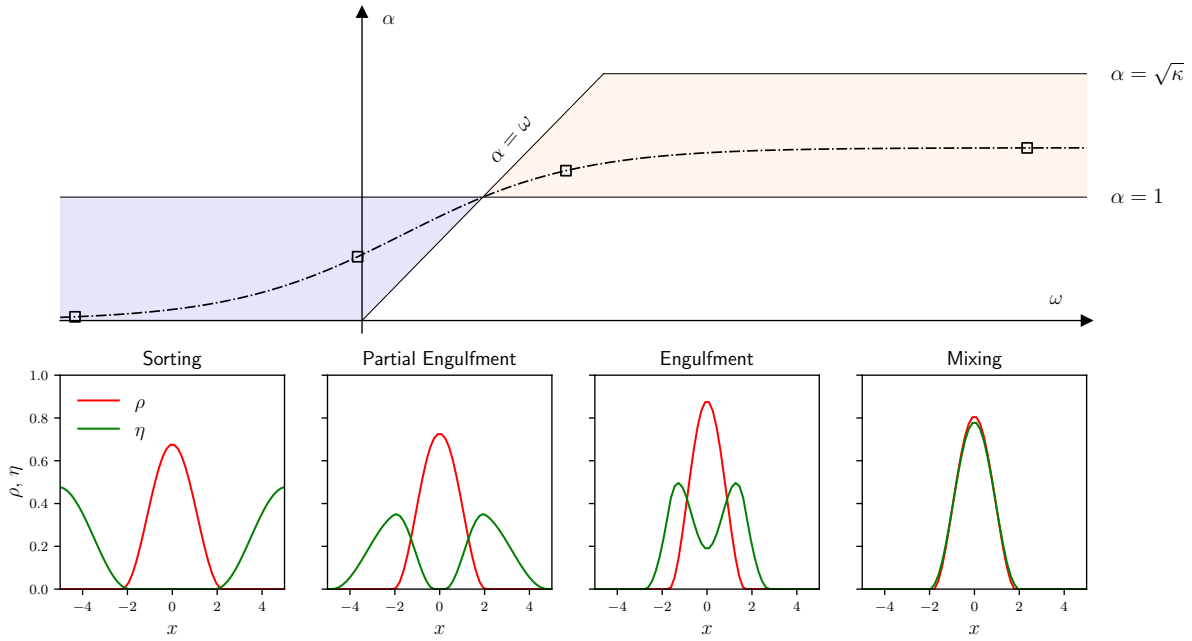


Figure 5.5: Understanding the impact of changing model parameters. Imposing that η is the less cohesive population implies $\mu > \kappa > 1$ as discussed in the text. We focus then on the cross-interactions. Parameter ranges for α and ω shown above: the blue-shaded region represents the weak cross-adhesion regime, while the red-shaded region corresponds to the case of strong cross-adhesion. Below we plot the numerically found steady states for the parameter values given by the square points: sorting, $\omega = -2.38$, $\alpha = 0.03$; partial engulfment, $\omega = -0.04$, $\alpha = 0.52$; engulfment, $\omega = 1.69$, $\alpha = 1.21$; mixing, $\omega = 5.51$, $\alpha = 1.40$. In every case $\kappa = 2$ and $\mu = 4$. We observe the different patterns seen in the Steinberg experiments and the transition from sorting to mixing as we increase the cross-adhesion, agreeing with the model interpretation. Numerical simulations performed on a domain of length $L = 5$ and $\delta x = 0.2$, $\delta t = 0.01$ with periodic boundary conditions and initial condition $\rho(x,0) = \eta(x,0) = \chi_{|x|<1.5}/2$. See [91] for an animated movie with the stationary states corresponding to each point in the dashed line.

Given the gradient flow structure of the local model in Eqs. (5.10), one could also understand these regimes by looking at the free energy $\mathcal{F}_2[\rho, \eta]$. We focus on the interaction terms in Eq. (5.11). In particular, it becomes evident that whenever $\omega < 0$, then in order to minimize the energy, both species will tend to separate so that the value of the product $\rho\eta$ is small. On the other hand, when $\omega > 0$, the two cell types will be attracted to each other, trying to maximize the value of $\rho\eta$.

We now simulate Eqs. (5.10) having in mind the above considerations. We start with numerical simulations in small domains and periodic boundary conditions, as shown in Figure 5.5. The figure suggests that our intuition of the model was correct, since the described regimes are able to replicate the four patterns observed in the Steinberg experiments. As predicted by the differential adhesion hypothesis, we observe that the two cell populations tend to separate when the cross-adhesion is weak. On the other hand, if the cross-adhesion is larger, the more cohesive population ρ gets engulfed inside η, and eventually the cross-adhesion is strong enough to drive mixing of the two.

The same patterns emerge in numerical simulations on larger domains. Here we choose parameter values corresponding to the two regimes that we described above, and show the solutions at different times in Figure 5.6. When the cross-adhesion is non-zero ($\alpha \neq 0$), steady states are composed of multiple compactly supported blobs or bumps. In the next section we will see that it is possible to find analytically the exact shape of each one of these bumps, given their individual masses. This is an advantage with respect to nonlocal models, where analytical solutions are only available for specific types of potentials [53]. Note however, that predicting the final mass of each of the bumps is difficult.

Observe too that solutions in the weak cross-adhesion regime — corresponding to the sorting and partial engulfment patterns — show very similar behaviour for early times (see Figures 5.6 and 5.7). When $\alpha = 0$ then both cell species tend to separate, converging to more or less sharply segregated solutions — which does not happen in previous nonlocal models that consider linear diffusion [6]. However, when cross-adhesion is small but strictly positive, the two populations move away from each other at early times and later organise themselves to form different aggregates, composed by

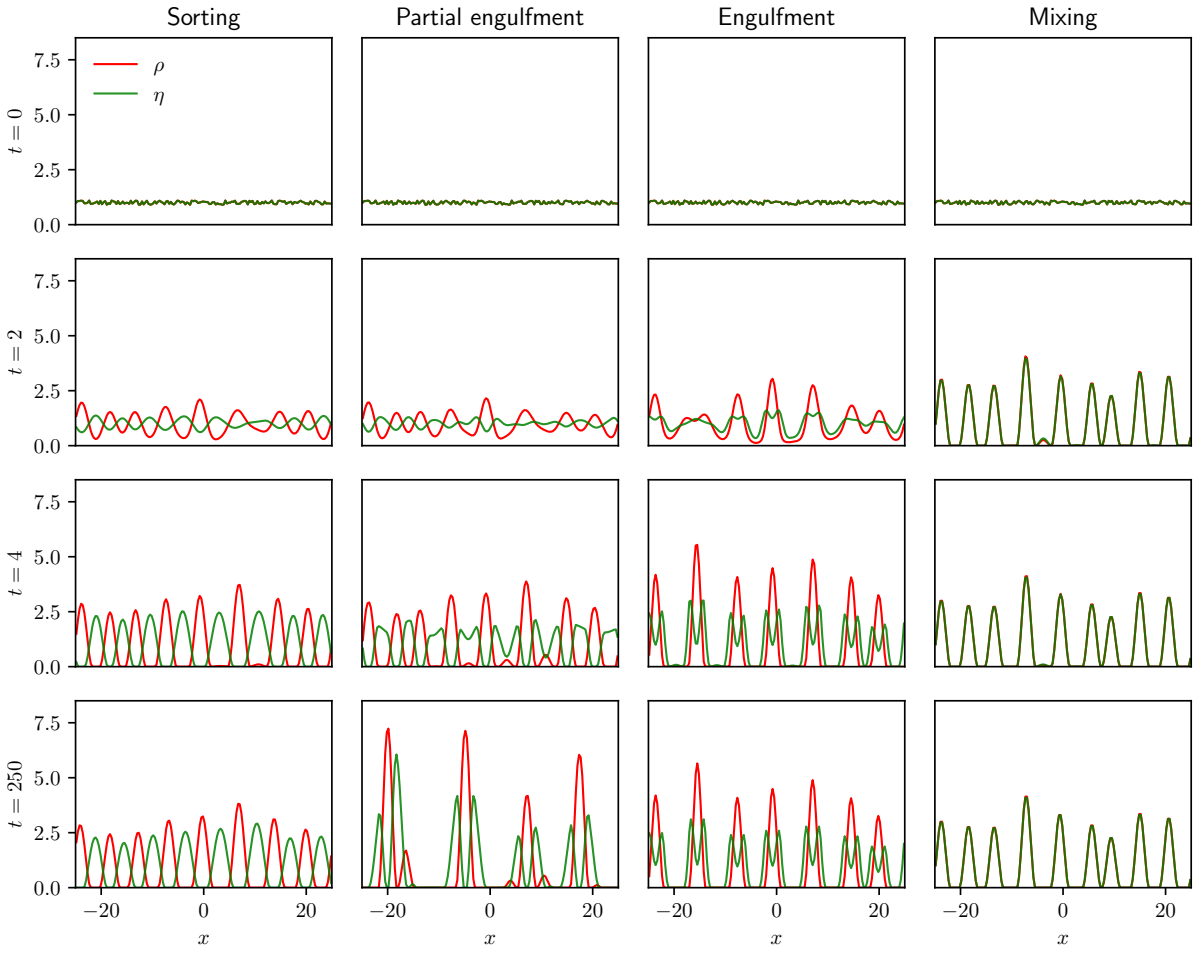


Figure 5.6: Solutions of the local model using model parameters related to the Steinberg experiments. Each column represents the solution with the same set of parameters and at different times. Mixing, $\alpha = 1.4$, $\omega = 6$; engulfment, $\alpha = 1.3$, $\omega = 2$; partial engulfment, $\alpha = 0.8$, $\omega = 0.2$; sorting, $\alpha = 0$, $\omega = -1$. In every case $\kappa = 2$ and $\mu = 4$ and also $L = 25$, $\delta x = 0.2$, $\delta t = 0.01$. See [91] for animated movies.

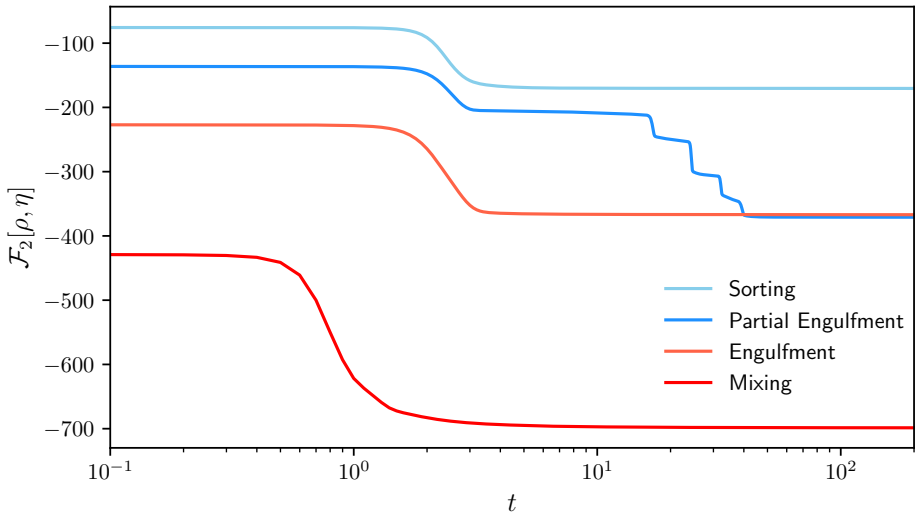


Figure 5.7: Energy decay given by Eq. (5.11) for the numerical solutions in Figure 5.6. Solutions corresponding to the weak and strong cross-adhesion regimes represented in blue and red, respectively. In general, the stronger the cross-adhesion, the faster the decay of $\mathcal{F}_2[\rho,\eta]$.

different coexistence regions. This kind of metastability (see Figure 5.7) has also been observed before in Cahn–Hilliard type systems [16, 59].

5.3.4 Stationary solutions

We now move our attention to the study of stationary solutions of the local system, Eqs (5.10). The computations in this section were performed with the help of the open-source software SageMath [203].

Motivated by the numerical simulations in the previous sections, we assume that such stationary solutions are compactly supported, and focus on one of the compact intervals shown in Figure 5.6 where we have nonzero solutions. In particular we assume that stationary solutions are of the form of the solutions corresponding to the engulfment/mixing patterns shown in Figure 5.5. Let us assume again that ρ is the more cohesive populations and thus, as discussed before, $\mu > \kappa$. Having this in mind, we consider

$$\operatorname{supp}(\rho) = [-b, b], \qquad \operatorname{supp}(\eta) = [-c, c],$$

with $0 < b < c$. Note that this defines a coexistence region for the two-species which is given by the support of ρ. For simplicity, we further assume that both species have the

same mass $\int_{-b}^{b} \rho \, \mathrm{d}x = \int_{-c}^{c} \eta \, \mathrm{d}x = m$. For different masses, a similar calculation shows that the stationary solutions depend on each individual mass, although we expect the same biological phenomena to be possible across parameter space. A similar behaviour is also observed for nonlocal models [53].

Stationary solutions of Eqs. (5.10) are characterised by

$$\kappa\rho'' + \alpha\eta'' + \mu\rho + \omega\eta = C_1; \tag{5.12a}$$

$$\alpha\rho'' + \eta'' + \omega\rho + \eta = C_2; \tag{5.12b}$$

where C_1 and C_2 are arbitrary constants to be determined. Note again that these two conditions mean that steady states of Eqs. (5.10) are critical points of the energy $\mathcal{F}_2$. Here, and motivated by our exploration of the one species system, we conjecture that these steady states correspond to the energy minimisers, which in the two-species case also satisfy the zero contact angle condition $\rho'(b) = \eta'(c) = 0$.

To see that, we follow the same argument, again motivated by [17], and consider $\bar{b} = b + \delta b$, $\bar{c} = c + \delta c$, $\bar{\rho} = \rho + \delta\rho$, and $\bar{\eta} = \eta + \delta\eta$. Calculating the perturbed energy up to first order we find

$$\begin{aligned}\mathcal{F}_2[\bar{\rho}, \bar{\eta}] = \mathcal{F}_2[\rho, \eta] &+ \int_{-b}^{b} \delta\rho \, \frac{\delta\mathcal{F}_2}{\delta\rho} \, \mathrm{d}x + \int_{-c}^{c} \delta\eta \, \frac{\delta\mathcal{F}_2}{\delta\eta} \, \mathrm{d}x \\ &+ \delta b \cdot \kappa\rho'(b)^2 + \delta c \cdot \eta'(c)^2 + \delta b \cdot 2\alpha\rho'(b)\eta'(b).\end{aligned}$$

Then, if (ρ, η) minimises the energy we need $\rho'(b) = \eta'(c) = 0$.

5.3.4.1 Outside the coexistence region

We thus look for symmetric solutions with zero contact angle. Using both assumptions and integrating Eqs. (5.12) on $[-b, b]$ we find

$$2b\, C_1 = 2\alpha\eta'(b) + m\,(\mu + \omega\delta)\,, \tag{5.13a}$$

$$2b\, C_2 = 2\eta'(b) + m\,(\omega + \delta)\,, \tag{5.13b}$$

where $\delta = \int_{-b}^{b} \eta \, \mathrm{d}x \,/ \int_{-c}^{c} \eta \, \mathrm{d}x$ is the fraction of mass of η in the coexistence region. Integrating the equation for η, Eq. (5.12b), on $[-c, c]$ we obtain

$$2c\, C_2 = m\,(\omega + 1)\,. \tag{5.14}$$

This last expression gives C_2 in terms of the model parameters and c. Note too that $\eta'(b)$ can be solved from Eqs. (5.13) and using Eq. (5.14), we can also find C_1 in terms of the model parameters, b, c and the mass fraction δ.

We solve first for η, outside the coexistence region. When $x \in [b, c]$ we have $\rho = 0$ and then

$$\eta'' + \eta = C_2.$$

General solutions read

$$\eta(x) = A_1 \cos(x) + B_1 \sin(x) + C_2, \tag{5.15}$$

with A_1, B_1 constants to be determined. Solutions on $[-c, -b]$ can be found via the substitution $B_1 \mapsto -B_1$, due to the symmetry assumption. Imposing $\eta(c) = \eta'(c) = 0$ and using Eq. (5.14) we find an explicit expression for η

$$\eta(x) = \frac{m\,(\omega+1)}{2c}\left(1 - \cos\left(c - \frac{x^2}{|x|}\right)\right), \qquad \text{for } |x| \in [b, c].$$

Note that c is still unknown. However, knowing η outside the coexistence region is enough to find also the mass fraction δ in terms of b and c and model parameters

$$\delta = \frac{\int_{-b}^{b} \eta \,\mathrm{d}x}{\int_{-c}^{c} \eta \,\mathrm{d}x} = 1 - \frac{2}{m}\int_{b}^{c} \eta \,\mathrm{d}x = 1 + (\mu + \omega)\left(\frac{b}{c} - 1 + \frac{\sin(c-b)}{c}\right).$$

This last expression allows us to write C_1, C_2 only in terms of b and c, and the model parameters.

5.3.4.2 Coexistence region

In order to find solutions on the coexistence region $[-b, b]$, we rewrite Eqs. (5.12) in more compact form

$$\mathbf{\Sigma}'' + M^{-1}N\mathbf{\Sigma} = M^{-1}\mathbf{C}, \tag{5.16}$$

where

$$\mathbf{\Sigma} = \begin{pmatrix}\rho\\ \eta\end{pmatrix}, \quad N = \begin{pmatrix}\mu & \omega\\ \omega & 1\end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix}C_1\\ C_2\end{pmatrix},$$

and M is defined by Eq. (5.3.1). With this, general solutions of Eq. (5.16) can be written as

$$\mathbf{\Sigma}(x) = A\mathbf{v_1}e^{i\lambda_1 x} + B\mathbf{v_1}e^{-i\lambda_1 x} + D\mathbf{v_2}e^{i\lambda_2 x} + E\mathbf{v_2}e^{-i\lambda_2 x} + N^{-1}\mathbf{C}, \tag{5.17}$$

where $\mathbf{v_1}, \mathbf{v_2}$ are eigenvectors of $M^{-1}N$ with eigenvalues λ_1^2, λ_2^2, respectively. For simplicity, we also set now

$$\begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = N^{-1} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}.$$

Note that both D_1 and D_2 can be written in terms of b, c and the model parameters. The eigenvalues and eigenvectors of $M^{-1}N$ can be found explicitly

$$\lambda_1^2 = \frac{\kappa + \mu - 2\alpha\omega + \sqrt{\Delta}}{2 \det M}, \qquad \lambda_2^2 = \frac{\kappa + \mu - 2\alpha\omega - \sqrt{\Delta}}{2 \det M};$$

$$\mathbf{v_1} = \begin{pmatrix} 2(\alpha - \omega) \\ \mu - \kappa - \sqrt{\Delta} \end{pmatrix}, \qquad \mathbf{v_2} = \begin{pmatrix} 2(\alpha - \omega) \\ \mu - \kappa + \sqrt{\Delta} \end{pmatrix};$$

with

$$\Delta = (\mu - \kappa)^2 + 4(\alpha\mu - \kappa\omega)(\alpha - \omega).$$

By looking at Δ as a quadratic polynomial in ω, we see that

$$\Delta \geq \left(1 - \frac{\alpha^2}{\kappa}\right)(\mu - \kappa)^2 > 0,$$

and hence λ_1^2, λ_2^2 are always real.

In general, λ_1^2 is always positive and λ_2^2 can be either positive or negative. To see this, write

$$\lambda_1^2 = \frac{1}{2}\left(\operatorname{tr}(M^{-1}N) + \sqrt{\operatorname{tr}(M^{-1}N)^2 - 4 \det M^{-1} \det N}\right).$$

If $\operatorname{tr}(M^{-1}N) < 0$, then $\kappa + \mu - 2\alpha\omega < 0$ and thus

$$\det N = \mu - \omega^2 < 2\alpha\omega - \omega^2 - \kappa < -(\omega - \alpha)^2 < 0\,.$$

Consequently λ_1^2 is always positive. However, for λ_2^2 we can write

$$\lambda_2^2 = \frac{1}{2}\left(\operatorname{tr}(M^{-1}N) - \sqrt{\operatorname{tr}(M^{-1}N)^2 - 4 \det M^{-1} \det N}\right),$$

and hence λ_2^2 will be negative whenever $\operatorname{tr}(M^{-1}N) < 0$ or $\operatorname{tr}(M^{-1}N) > 0$ and $\det N < 0$. In terms of ω, this happens whenever $\omega > \min\left(\sqrt{\mu}, (\kappa + \mu)/2\alpha\right) = \sqrt{\mu}$, where we used that $\mu > \kappa > \alpha^2$. We consider both cases separately now.

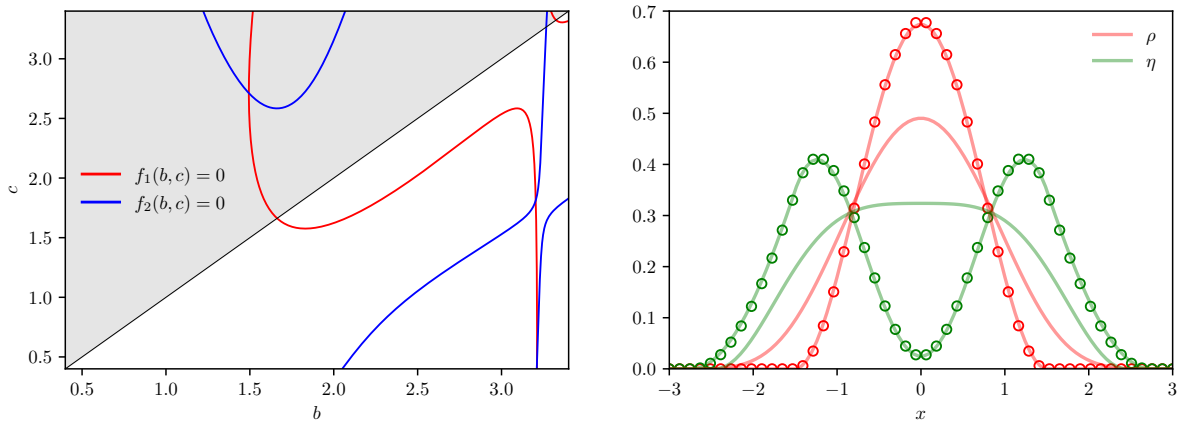


Figure 5.8: (left) Regularity of η determines the support. The shaded region represents the condition $b < c$, i.e. that ρ is the more cohesive population. (right) Numerical solution of the two-species model with initial condition $\rho(x,0) = \eta(x,0) = \chi_{|x|<1}/2$. Solutions are shown at $t = 1$ and $t = 25$ (solid line) and the corresponding analytical stationary solutions are also plotted (dots). The analytical and numerical stationary solutions agree perfectly. Simulation parameters: $\kappa = 2$, $\alpha = 1.3$, $\mu = 4$, $\omega = 1.8$, $m = 1$, $L = 3$, $\delta x = 0.1$, $\delta t = 10^{-2}$.

5.3.4.3 Two positive eigenvalues

As discussed, this happens whenever $\omega < \sqrt{\mu}$. In this case, and using the fact that the stationary solutions are symmetric, we can write general solutions of Eq. (5.17) as

$$\rho(x) = 2(\alpha-\omega)A_2\cos(\lambda_1 x) + 2(\alpha-\omega)B_2\cos(\lambda_2 x) + D_1\,;$$
$$\eta(x) = (\mu-\kappa-\sqrt{\Delta})A_2\cos(\lambda_1 x) + (\mu-\kappa+\sqrt{\Delta})B_2\cos(\lambda_2 x) + D_2\,.$$

The coefficients A_2, B_2 can be found in terms of b, c by imposing $\rho(b) = \rho'(b) = 0$. These conditions give

$$A_2 = -\frac{D_1\left(\lambda_1^{-1}\cot(\lambda_1 b) - \lambda_2^{-1}\cot(\lambda_2 b)\right)^{-1}}{2(\alpha-\omega)\lambda_1\sin(\lambda_1 b)}\,,$$
$$B_2 = \frac{D_1\left(\lambda_1^{-1}\cot(\lambda_1 b) - \lambda_2^{-1}\cot(\lambda_2 b)\right)^{-1}}{2(\alpha-\omega)\lambda_2\sin(\lambda_2 b)}\,.$$

Note that we have expressed ρ and η only in terms of the model parameters and b, c. In order to find these two parameters we only need to impose that η is continuously

differentiable on $x = b$. This condition gives two equations

$$f_1(b,c) = \lim_{x\to b^+} \eta'(x) - \lim_{x\to b^-} \eta'(x) = 0\,;$$
$$f_2(b,c) = \lim_{x\to b^+} \eta(x) - \lim_{x\to b^-} \eta(x) = 0\,;$$

which can be solved numerically to find b and c. This is shown in Figure 5.8, alongside the corresponding stationary solutions. We see that the numerical stationary state and the solution found in this section agree perfectly.

5.3.4.4 Positive and negative eigenvalues

Conversely, whenever $\omega > \sqrt{\mu}$ we have $\lambda_2^2 < 0$. Using again that the stationary states are symmetric, we can write general solutions Eq. (5.17) as

$$\rho(x) = 2(\alpha-\omega)A_2\cos(\lambda_1 x) + 2(\alpha-\omega)B_2\cosh(|\lambda_2|x) + D_1\,;$$
$$\eta(x) = (\mu-\kappa-\sqrt{\Delta})A_2\cos(\lambda_1 x) + (\mu-\kappa+\sqrt{\Delta})B_2\cosh(|\lambda_2|x) + D_2\,.$$

Imposing $\rho'(b) = \rho'(b) = 0$ we find

$$A_2 = -\frac{D_1\left(\lambda_1^{-1}\cot(\lambda_1 b) + |\lambda_2|^{-1}\coth(|\lambda_2|b)\right)^{-1}}{2(\alpha-\omega)\lambda_1\sin(\lambda_1 b)}\,,$$
$$B_2 = -\frac{D_1\left(\lambda_1^{-1}\cot(\lambda_1 b) + |\lambda_2|^{-1}\coth(|\lambda_2|b)\right)^{-1}}{2(\alpha-\omega)|\lambda_2|\sinh(|\lambda_2|b)}\,.$$

And again, we have the same two conditions on the regularity of η at $x = b$

$$f_1(b,c) = \lim_{x\to b^+} \eta'(x) - \lim_{x\to b^-} \eta'(x) = 0\,; \tag{5.23a}$$
$$f_2(b,c) = \lim_{x\to b^+} \eta(x) - \lim_{x\to b^-} \eta(x) = 0\,. \tag{5.23b}$$

Having explicit stationary solutions is useful for predicting transitions between the different patterns shown in Figure 5.6. We focus here on the engulfment and mixing patterns, obtained in the strong cross-adhesion regime. Although there is no sharp transition between these two patterns, it is instructive to understand how solutions vary when increasing the cross-adhesion strength. We plot the analytical solutions in Figure 5.9. Note that in general, this is not possible for nonlocal models. One could further ask whether these calculations can provide analytical insights into the transitions

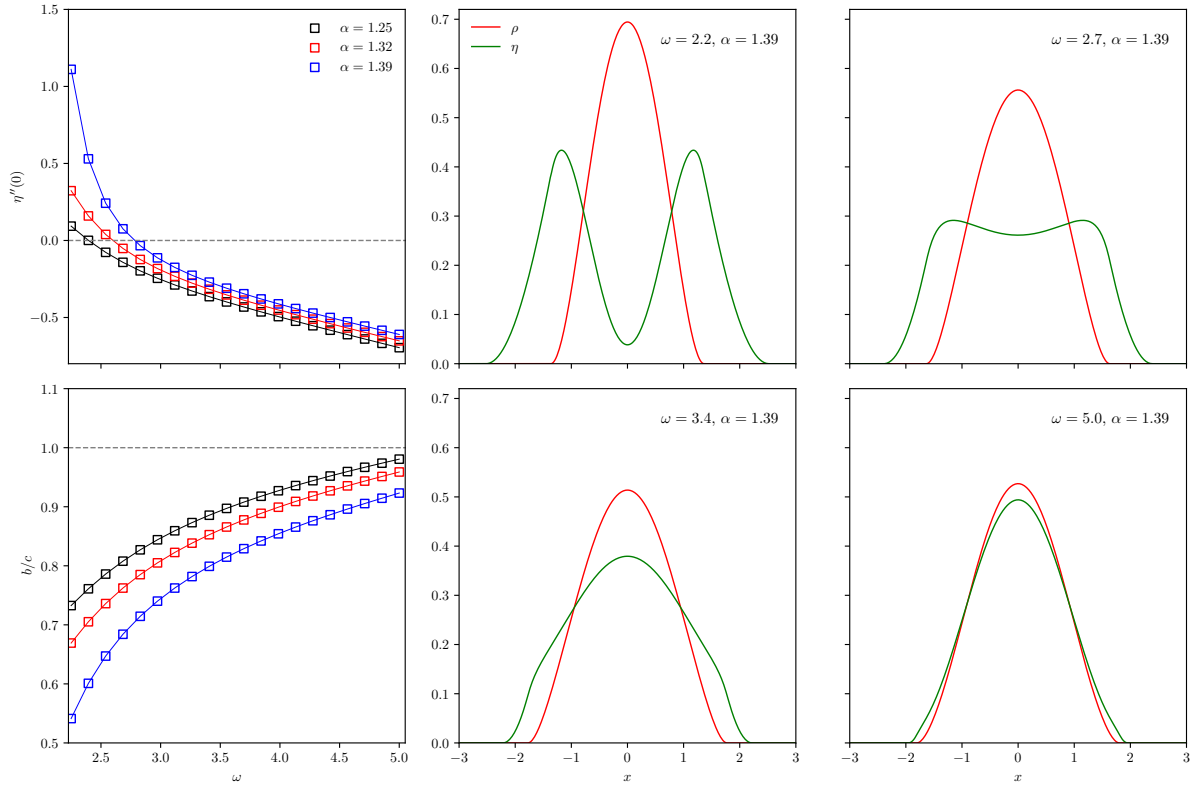


Figure 5.9: Engulfment-mixing transition from the analytical stationary solutions. In every case $\kappa = 2$, $\mu = 4$. On the left we plot the second derivative of the engulfing species at the origin $\eta''(0)$ and the quotient b/c as a function of ω and for different values of α. The support length relation is found by solving numerically Eqs. (5.23). On the right, we plot different solutions, showing the transition from one pattern to the other.

between different patterns. However, the equations to determine the support of the densities (b, c) are complicated and need to be solved numerically. as we show in Figure 5.8. In future work, it would be interesting to examine this question under a slightly simpler setting — for instance, assuming identical self-adhesion interactions [53].

One measure to quantify when one of the populations is trapped inside the other is given by the second derivative of the engulfing species, $\eta''(0)$. This quantity is positive for engulfment and negative when the two-species are mixed. Another possibility is to look at the quotient b/c, which should approach unity as we increase the strength of cross-adhesion between the two-species. Using the expressions we found for the steady states, we plot how these quantities vary with ω and α in Figure 5.9. Again we confirm our intuition, since increasing the cross-adhesion strength — and hence ω —

yields the expected behaviour, decreasing $\eta''(0)$ and a quotient b/c that approaches unity.

5.3.5 Numerical simulations for Steinberg experiments in two dimensions

In two spatial dimensions, the explicit calculations performed in the previous section involve Bessel functions, and imposing boundary and regularity conditions becomes a very cumbersome task. Here instead we explore the model in two dimensions numerically, performing the same type of experiments as in the one-dimensional case, which we show in Figure 5.10.

By choosing appropriate parameters, as explained in the previous sections, we can again recover the four patterns seen in the Steinberg experiments. The model dynamics are in general similar to the one-dimensional case, with the strong cross-adhesion regime showing a faster decay to the stationary solution. Note that although the final configurations are very different for the chosen parameters, in the weak cross-adhesion regime solutions show very similar patterns for early times. In this case, the model shows two separated timescales with solutions showing large differences only after the first one, as it already happened in the one-dimensional case (see Figure 5.7). In the limit of vanishing cross-adhesion we recover the typical cell sorting pattern with sharp segregation of the two-species, which is only observed in models that account for population pressure. Note also that this pattern is accentuated with respect to the one-dimensional case.

5.4 Discussion and outlook

We have presented a local continuum model of aggregation–diffusion in the form of a system of thin-film equations, and used this framework to describe adhesion-based pattern formation. The idea of describing tissues using fluid-like properties has been recurrent. In fact, Steinberg first employed this analogy in developing the differential adhesion hypothesis. However, the model proposed here differs from other descriptions based on the fluid analogy [2], as it can be directly related to aggregation–diffusion equations. The local model has physically interpretable parameters and suc-

cessfully explains the patterns observed in experiments, as well as those predicted by the differential adhesion hypothesis. This had only been achieved in the continuum setting with nonlocal models [87].

There are also more modern views on the differential adhesion hypothesis, such

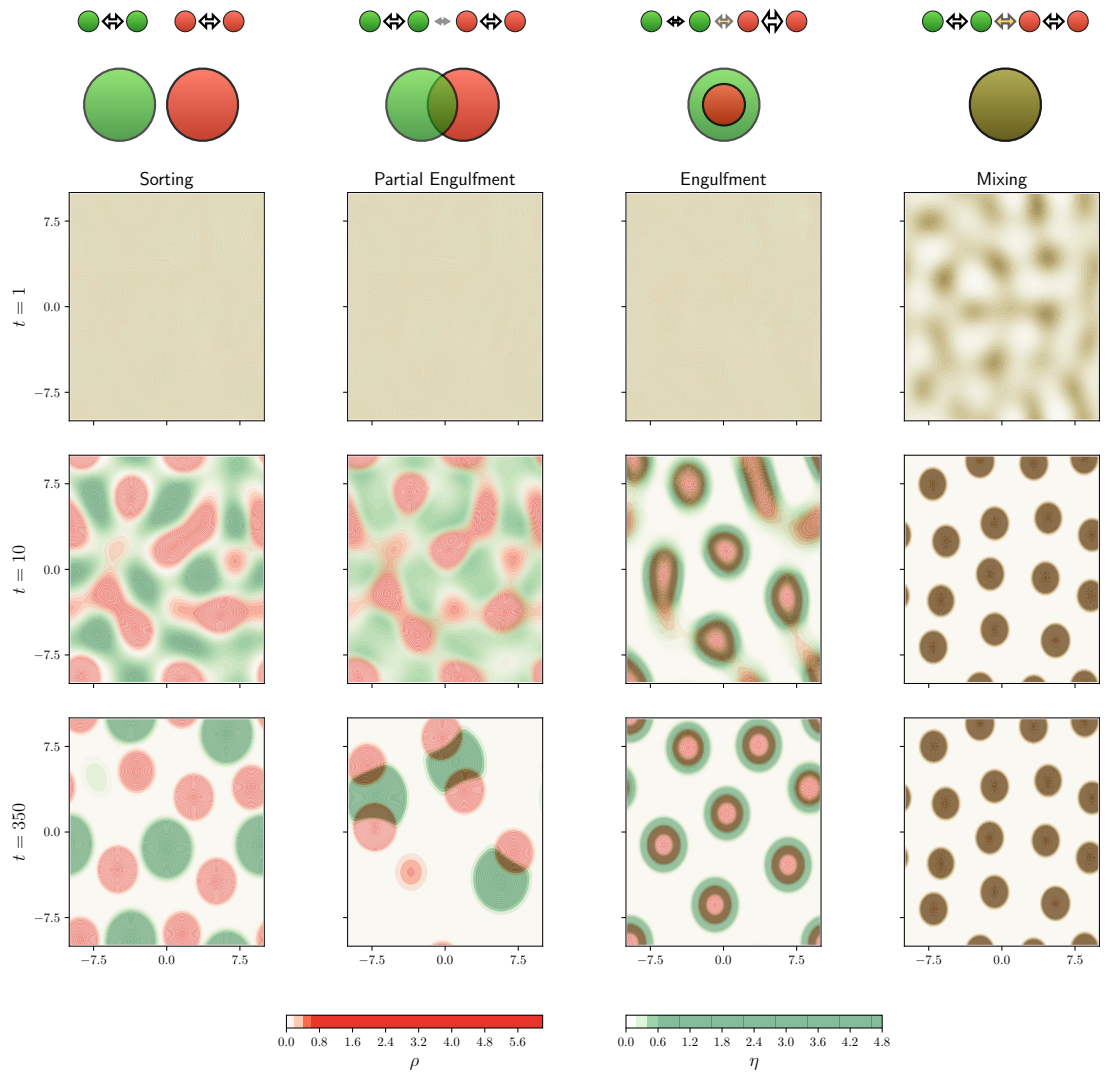


Figure 5.10: Numerical solutions of the local model using model parameters related to Steinberg experiments. Each column represents the solution with the same set of parameters and at different times. Mixing, $\alpha = 1.4$, $\omega = 8$; engulfment, $\alpha = 1.3$, $\omega = 2$; partial engulfment, $\alpha = 0.5$, $\omega = -0.02$; sorting, $\alpha = 0$, $\omega = -1$. In every case $\kappa = 2$ and $\mu = 4$ and also $L = 10$, $\delta x, \delta y = 0.2$, $\delta t = 0.001$. The initial condition is the same for every experiment, $\rho(x,0)$, $\eta(x,0) = 0.3$ plus a small perturbation. See [91] for animated movies.

as the differential interfacial tension hypothesis [26, 208]. This hypothesis suggests that tissue surface tension is determined not only by adhesion bonds between cells but also by cortical tension [4, 222]. Some modelling efforts have aimed to incorporate these factors, showing that when cell–cell adhesion is the dominant interaction, the differential adhesion hypothesis successfully predicts tissue behavior [147]. However, when cortical tension is stronger, the differential adhesion hypothesis may no longer be sufficient, indicating that in this regime cell shape plays an important role. While this is an important consideration, our model here focuses on the adhesion-based regime, where both the differential adhesion hypothesis and the particle-based approximation hold.

The approach taken here is based on previous studies of aggregation–diffusion models, where the nonlocal terms are approximated by a series of terms including higher-order derivatives of the densities [17, 71] — these consider a porous-medium type repulsion with exponent three, instead of the exponent two considered here. Following a similar approach, energy minimisers and linear stability for multi-species systems have also been explored [83]. *Bernoff and Topaz* [17] show that energy minimisers of the local and nonlocal models are in good agreement and share similar qualitative properties in the limit of large populations, far from aggregation boundaries. In our case, the choice of diffusion is different so this result may not directly apply. However, recent work [49, 82] demonstrates that, in a similar setting with unit volume and compactly supported potential, the nonlocal model converges to its local approximation as the scaling parameter tends to zero ($a \to 0$). Recent numerical investigations by *Buttenschön et al.* [39] suggest more significant differences between the local and nonlocal models when the interaction potential has non-compact support.

In recent work, we established the existence theory for the local model in both the one-species and two-species cases, as well as for a more general second-order aggregation term [51, 87]. However, several analytical challenges remain unresolved. Uniqueness is an open problem, as the functionals involved are not convex. The proof of existence of free energy minimisers in the entire space is also missing, as we do not have a method to control the escape of mass at infinity. Additionally, long-time asymptotics remain an open question in all global existence cases. We refer to [51] for further

discussion on the analytical open problems.

The local model is, in principle, less complex and more analytically tractable than its nonlocal counterpart. Thus, we believe it could offer significant advantages for applications. Although solving numerically fourth-order equations can be challenging, there is a simplification in the numerical scheme complexity when one approximates convolutions with local operators [12, 45]. From an applications perspective, it still needs to be tested whether the model can fully capture the dynamics of co-culture experiments involving multiple cell populations with different adhesive properties. While we have confirmed that the model can capture steady states that are consistent with the differential adhesion hypothesis, further exploration is needed to assess the extent to which it can be applied to these settings. To this end, one could estimate the different parameters in the model using computational inference techniques [89]. This is now more feasible with the local model, as we only need to infer the parameters directly, rather than the interaction potentials required in the nonlocal model. We thus expect that the use of local models will open up further developments in the inference of aggregation–diffusion phenomena across various fields.

Chapter 6

Modelling orbiting in multicellular spheroids via differential adhesion

In this final research chapter, we use differential-adhesion models to study the emergence of collective orbiting in small multicellular spheroids. This work builds on the experimental–theoretical study

"J. Kim, H. Jeong, C. Falcó, A. M. Hruska, W. D. Martinson, A. Marzoratti, M. Araiza, H. Yang, C. Franck, J. A. Carrillo, M. Guo, & I. Y. Wong. *Collective transitions from orbiting to matrix invasion in 3D multicellular spheroids.* bioRxiv preprint, 2025" [128].

To capture the dynamics at low cell numbers, we constructed an agent-based model based on the framework introduced in Chapter 1. My main contributions involve formulating the model, implementing the numerical code (available on Github), analysing the model dynamics, and writing the modelling sections of the manuscript.

6.1 Introduction

Cell–cell and cell–matrix adhesions play a crucial role in tissue morphogenesis by driving cell sorting, boundary formation, and other collective rearrangements that shape developing tissues and organs [56, 207]. These adhesive interactions also maintain tissue homeostasis in adult organisms and, when dysregulated, can lead to pathological processes such as cancer invasion. In previous chapters we have focused on continuum frameworks for cell–cell interactions, focusing on how differential adhesion gives rise to sorting phenomena. Yet in most *in vitro* and *in vivo* contexts, cells interact not only with one another but also with the extracellular matrix, dynamically remodelling

the matrix as it guides their movement. Understanding this two-way interactions remains a central challenge in the field of collective cell migration, and more generally in quantitative biology [69, 151].

In many developmental processes, the interplay between extracellular matrix interactions and geometric constraints, such as curvature, can lead to coordinated migration. Yet how this behaviour emerges from individual cell dynamics and environmental interactions is not well understood. For instance, coherent rotational motions drive lumen formation in mammary and lung acini [202, 63] and orchestrate global egg-chamber rotation during *Drosophila* oogenesis [110, 61]. These processes establish tissue symmetry, direct extracellular matrix alignment, and sites of symmetry breaking that underlie branching morphogenesis in organs such as kidney and salivary gland [179]. Conversely, aberrant transitions from collective motion to invasion can result in carcinoma dissemination, where epithelial cells breach the basement membrane through localised protrusions [93].

In this chapter we turn to a model experimental system, multicellular epithelial spheroids embedded in three-dimensional collagen gels [24], to study how cells balance adhesion to one another with adhesion to the surrounding matrix [40, 60]. The emergence of orbiting in multicellular spheroids provides a minimal, tunable experimental model for understanding how curvature and dynamic cell–matrix reciprocity govern collective migration in development and disease. Recent work by *Kim et al.* [128], showed that small spheroids first adopt a coherent, circumferential orbiting motion along the curved spheroid–matrix interface, and then break symmetry to launch radially oriented invasive strands — see Figure 6.1. The transition from coordinated rotation to localised invasion implicates both the build-up of contractile tractions at cell–matrix contacts and the modulation of intercellular adhesions, yet the mechanistic underpinnings of this switch are not well understood.

Here, we develop and analyse an agent-based model in which individual cells are represented as self-propelled particles [1, 24, 79] with tunable cell–cell [46] and cell–matrix adhesive interactions. By systematically varying the relative strengths of these two interactions, as well as the curvature of the confining matrix boundary, we can

reproduce collective orbiting observed experimentally, and predict the number and geometry of invasive strands that lead to disrupted orbiting.

Beyond recapitulating experimental behaviour, this modelling framework provides a quantitative approach for mechanistic hypothesis testing. For example, we predict that reducing matrix adhesion below a critical threshold will halt orbiting in the presence of strong cell–cell cohesion. Moreover, by introducing heterogeneity in adhesion parameters, we can investigate how epithelial–mesenchymal transitions modulate the balance between collective and invasive phenotypes, which we validate experimentally [128].

The rest of this chapter is organised as follows. First, we present the experimental setup and give a concise summary of the main findings together with the key questions to be analysed by the model. Next, we describe the agent-based model in detail, including the form of the intercellular and cell–matrix potentials and the numerical integration scheme. We then present simulation results across a range of adhesion and curvature parameters, highlighting the emergent orbiting phase diagram. Finally, we discuss the implications of our findings for broader questions of tissue morphogenesis and suggest future extensions that incorporate matrix remodelling and feedbacks.

6.2 Experimental setup and summary of main findings

Multicellular spheroids of mammary epithelial cells (MCF-10A) were embedded in a collagen hydrogel (type I) by using a hanging drop culture, yielding aggregates of approximately 500 cells with a spheroid diameter of approximately 150 μm[1]. Cell nuclei were fluorescently labelled (H2B-mCherry) to enable single-cell tracking in the equatorial plane and imaged using spinning disk confocal microscopy in an environmentally controlled chamber. To quantify collective motion, we extracted trajectories from maximum intensity projections.

Spheroids displayed three distinct phases over the timespan of the experiment (24 h) — see Figure 6.1. From 0-6 h post-embedding, cells remained largely stationary.

[1]The experimental work was carried out in the Wong Lab (Brown University). Ian Y. Wong and Jiwon Kim designed the experiments, which were performed by Ian Y. Wong, Jiwon Kim, and Alex M. Hruska. Collective migration data were analysed by Ian Y. Wong, Jiwon Kim, Hyuntae Jeong, Christian Franck, and Mauricio Araiza.

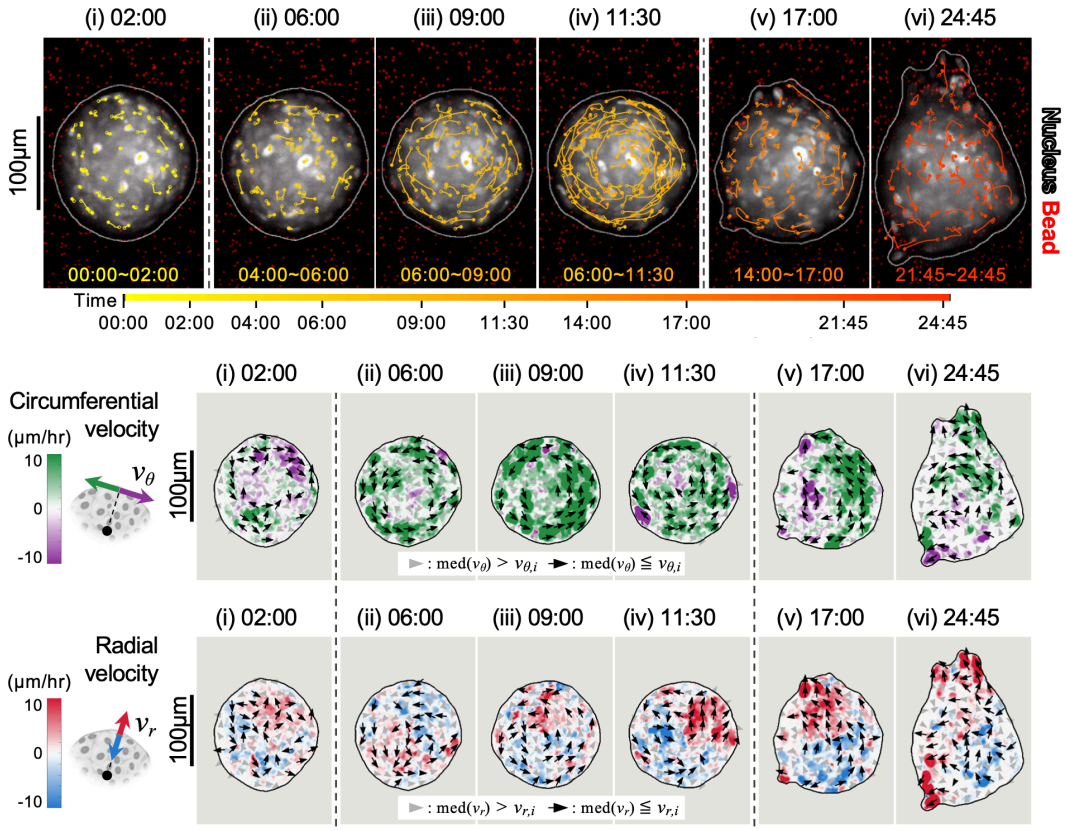


Figure 6.1: Multicellular spheroids transition from circumferential orbiting to radial invasion in 3D matrix. (top) Representative images of an MCF-10A spheroid embedded in a 3D collagen matrix, overlaid with selected trajectories of tracked nuclei across the equatorial plane. Trajectories are colour-coded from yellow to red to indicate different starting times over 24 h. (middle) Circumferential velocity profiles extracted via optical flow. Angular-velocity magnitudes are colour-coded from purple to green; black arrows denote velocities above the median, and gray arrowheads denote velocities below the median. (bottom) Radial velocity profiles extracted via optical flow.

Between 6-12 h, the outermost cell layer entered a highly coordinated orbiting phase. After 12 h, angular coordination broke down as two diametrically opposed hotspots of elevated radial velocity emerged, leading to multicellular strand invasion after 24 h. Morphologically, spheroids were initially slightly elongated, then rounded up during orbiting before developing sharper protrusions during invasion. Moreover, fibre tracking revealed that inward matrix tractions intensified at the poles throughout pre-orbiting and orbiting, then increased further during invasion; confocal imaging showed collagen fibres aligning along these high-traction axes — see *Kim et al.* for more

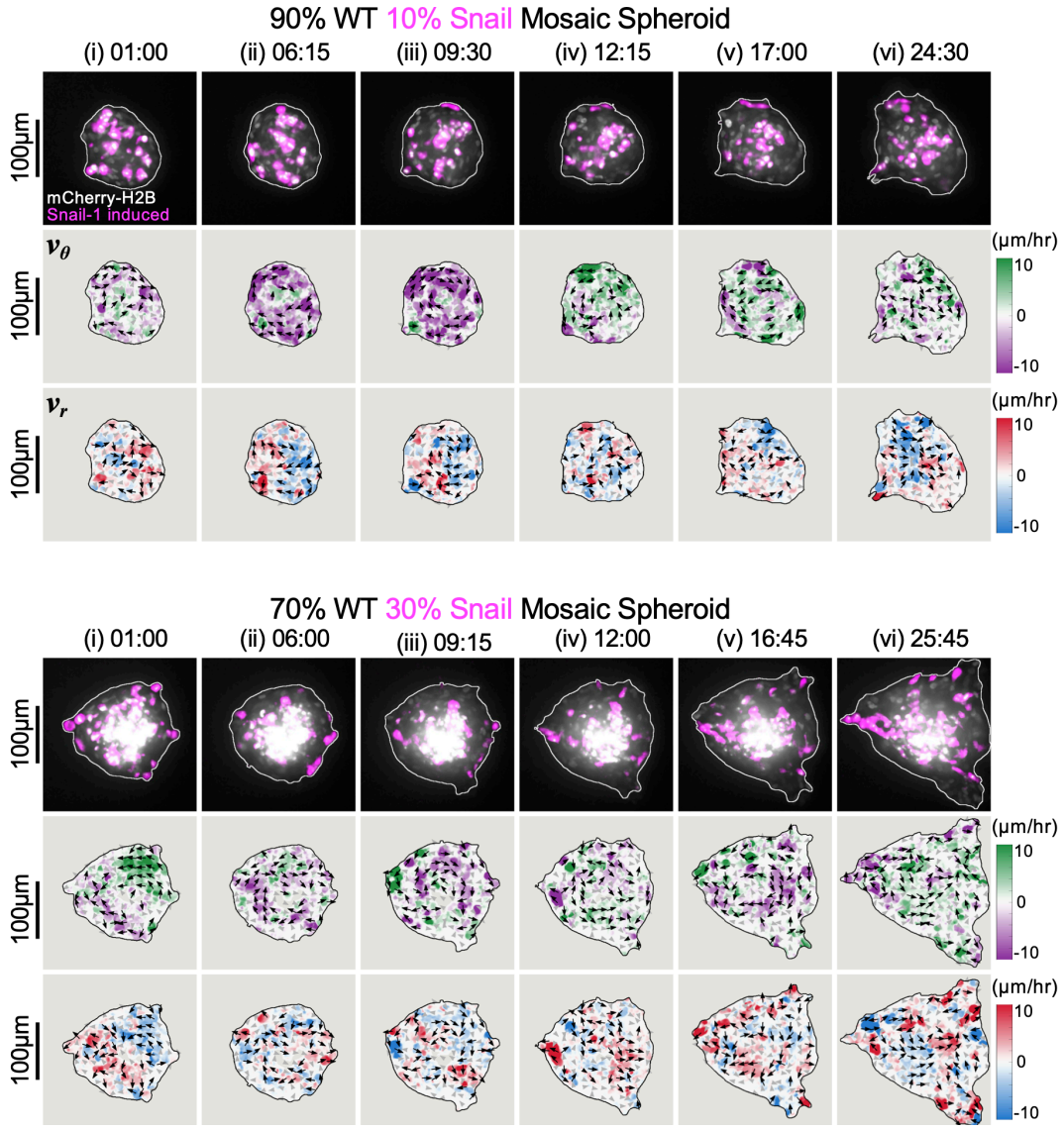


Figure 6.2: Coordinated migration is disrupted as the proportion of Snail-induced cells (pink) increases within mosaic spheroids of Snail-induced and wild-type (WT) cells. Representative snapshots at each Snail:WT ratio are shown with circumferential velocity profiles to highlight orbiting behaviour and radial velocity profiles to highlight invasive behaviour.

details [128].

To probe the impact of adhesion heterogeneity on both coordinated orbiting and subsequent invasion, we generated *mosaic* spheroids by mixing fluorescently labelled MCF-10A cells expressing Snail1 — which downregulates E-cadherin, weakening cell–

cell junctions, and simultaneously upregulates vimentin, increasing cytoskeletal plasticity and traction force generation, and promoting invasive motility — with wild-type (WT) cells at a constant total cell number — see Figure 6.2. With only 10% of the Snail cells, spheroids behaved nearly identically to pure WT aggregates: they elongated slightly in the first hours, then entered a robust orbiting phase that carried them into a rounded configuration by 18 h, and finally initiated radial invasion at three well-defined hotspots by 24 h. Importantly, this peripheral orbiting was driven by the WT majority.

By contrast, increasing the Snail1+ fraction to 30% disrupted the collective rotation. These spheroids retained an elongated shape throughout and never achieved coherent circumferential motion — pockets of transient rotation appeared briefly around 9 h but failed to span the entire aggregate. Nevertheless, invasion still emerged at three fixed peripheral sites, with Snail1 cells frequently occupying the leading tips of the protruding strands. Unlike the abrupt surge in matrix deformation seen in the 10% mosaics, the 30% spheroids displayed a continuous, gradual increase in matrix displacement, reflecting the loss of a synchronised contractile build-up and the persistence of locally driven invasion despite disrupted orbiting.

6.3 Mathematical model

First, we describe the mathematical model, followed by a discussion of biologically valid parameter ranges.

6.3.1 Model description

To further elucidate how collective orbiting was mediated by interactions between cells and the matrix boundary, we develop a mathematical model with tunable interactions and analyse how perturbations to this system affected the stability of coordinated migration. We model cells as discrete, off-lattice particles moving in a two-dimensional domain. We chose to work in two spatial dimensions because interlayer movement within spheroids is minimal, and the experimental cell migration analysis is performed across the equatorial plane. Note, however, the same framework can be extended to

three spatial dimensions. Cells are treated as active particles subject to drag (frictional) forces. We incorporate two key terms to describe their dynamics: a self-propulsion term proportional to the cell velocity, and a nonlinear friction term which acts against the cell velocity. This formulation, commonly used in models of self-propelled particles as a first approximation for active movement [1, 18, 79], naturally predicts an equilibrium speed at which cells travel where friction balances self-propulsion such that the speed of cell i is given by $|\mathbf{v}_i| = \sqrt{\alpha/\beta} =: u$ in the absence of other forces.

We model the matrix boundary as a set of discrete equally spaced particles to simplify numerical calculations of cell–matrix forces and facilitate the introduction of boundary perturbations. For a circular boundary of radius R made up of M points, the position of the l^{th} matrix point is given by $\mathbf{y}_\ell = (R\cos(2\pi\ell/M), R\sin(2\pi\ell/M))$ for $\ell = 1, \ldots, M$. Furthermore, we assume that these points are static so that we may focus on how the boundary geometry, cell–cell, and cell–matrix forces influence the stability of collective orbiting in a controlled setting.

We assume that cells can interact with each other and with their matrix boundary via cell–cell, $\mathbf{F}^{cc}$, and cell–matrix, $\mathbf{F}^{cm}$, forces which each encode soft repulsion at short distances and adhesion-based attraction at longer distances. The strength and magnitude of the cell–cell (*resp.* cell–matrix) force is assumed to depend on the relative displacement between the cell and its neighbour (*resp.* neighbouring matrix point). The cell–matrix force additionally depends on the direction in which the cell travels, as we will later introduce the assumption that cells can only adhere to matrix points within a specified angle of this direction. The specific functions that yield the cell–cell and cell–matrix forces are described in more detail below.

The position and velocity of cell i, $\mathbf{x}_i$ and $\mathbf{v}_i$, are given by Newton's second law, accounting for active, friction, cell–cell, and cell–matrix forces. The system of differential equations is given by

$$\frac{\mathrm{d}\mathbf{x}_i}{\mathrm{d}t} = \mathbf{v}_i\,, \tag{6.1}$$

$$\frac{\mathrm{d}\mathbf{v}_i}{\mathrm{d}t} = \left(\alpha - \beta|\mathbf{v}_i|^2\right)\mathbf{v}_i + \sum_{j\neq i,\, j=1}^{N} \mathbf{F}^{cc}(\mathbf{x}_j - \mathbf{x}_i) + \sum_{k=1}^{M} \mathbf{F}^{cm}(\mathbf{y}_k - \mathbf{x}_i, \mathbf{v}_i)\,, \tag{6.2}$$

$$\frac{\mathrm{d}\mathbf{y}_k}{\mathrm{d}t} = 0\,, \tag{6.3}$$

In contrast to the more common overdamped (first-order) descriptions — often supplemented by an explicit equation for polarity dynamics that effectively encodes directional persistence [30] — Eqs. (6.1)–(6.3) retain a second-order structure without implying that inertial mass effects are important in cell motility. We adopt this formulation for two reasons. First, the overdamped limit in the presence of active-drag is subtle and not well-characterised, whereas many active Brownian model reductions add such forces phenomenologically after eliminating acceleration. Second, closely related second-order swarm models have a well-developed literature, studying milling and flocking states [1], providing a useful baseline for analysis and comparison with our model and experiments.

The magnitudes of both cell–cell and cell–matrix forces are taken to depend only on the relative position $\mathbf{r}$ between a cell and its neighbour or between a cell and the matrix. This is captured by using two radially symmetric kernels, $K^{cc}(|\mathbf{r}|)$ and $K^{cm}(|\mathbf{r}|)$, which are related to the forces via

$$\mathbf{F}^{cc}(\mathbf{r}) = K^{cc}(|\mathbf{r}|)\frac{\mathbf{r}}{|\mathbf{r}|}, \tag{6.4}$$

$$\mathbf{F}^{cm}(\mathbf{r},\mathbf{v}) = \tilde{H}(|\mathbf{r}|,\, \hat{\mathbf{r}}\cdot\hat{\mathbf{v}} - \cos\theta_{\max})\, K^{cm}(|\mathbf{r}|)\frac{\mathbf{r}}{|\mathbf{r}|}\,. \tag{6.5}$$

The term $\tilde{H}(\cdot,\cdot)$ is a modified Heaviside function that arises from an assumption that cells can only adhere to matrix points that fall within an angle $\theta_{\max} \in [0,\pi]$ of their unit velocity vector $\hat{\mathbf{v}}$. Repulsive forces are not assumed to be affected by the modified Heaviside function, as they are assumed steric in origin and therefore isotropic. Assuming anisotropic cell–matrix adhesion — which reflects the front–rear polarity of migrating cells [2] — accelerates the emergence of collective orbiting. The modified Heaviside function is defined as

$$\tilde{H}(r,\,\lambda) = \begin{cases} 1, & \text{if } 0 < r < d_r; \\ H(\lambda), & \text{if } r \geq d_r\,; \end{cases}$$

where $H(\cdot)$ is the standard Heaviside function.

In all numerical simulations, we fix the angular sensing region such that the cells do not adhere to matrix points which lie outside an angle $\theta_{\max} = \pi/3$ from its current velocity vector. In Figure 6.10 we quantify the impact of changes in this parameter,

showing that results are robust under similar values. Note also that cell–matrix adhesion changes the equilibrium velocity of cells, such that the long-term cell speed will be slightly larger than $u = \sqrt{\alpha/\beta}$.

For simplicity we assume that repulsion (f_r) is identical for cell–cell and cell–matrix interactions, while adhesion depends on two parameters, f_a^{cc} and f_a^{cm}, that correspond to the respective kernel. While there are multiple possibilities for the functional forms of the repulsive and attractive kernels, for example power-law or exponentially decaying functions [1, 79], we opt instead to use kernels which ensure that cells do not experience forces from objects which lie beyond a certain maximum distance. The cell–cell and cell–matrix kernels used in this manuscript are based on [46], which have compact support and have been well characterised. They are given by

$$K^{\ell}(r) = \begin{cases} -f_r \left(\dfrac{d_r}{2r}\right)^{3-2s}, & \text{if } 0 < r < \frac{d_r}{2}; \\ \dfrac{2f_r(r-d_r)}{d_r}, & \text{if } \frac{d_r}{2} \leq r < d_r; \\ -\dfrac{4f_a^{\ell}(r-d_a)(r-d_r)}{(d_a-d_r)^2}, & \text{if } d_r \leq r < d_a; \\ 0, & \text{if } r \geq d_a, \end{cases} \tag{6.6}$$

where $\ell \in \{cc, cm\}$, $s \in (0, 2]$ is a parameter that captures the behaviour of the kernel near the origin and is related to the cell nucleus stiffness, d_r is a typical cell diameter, and d_a gives the range of adhesive interactions. For the simulations in this chapter, we set $s = 1.25$, which corresponds to a singular kernel at the origin. In the sections below, we show that a different choice of the exponent s affects the results only qualitatively. This is consistent with previous work using the same family of kernels [46]. Figure 6.3 shows a plot of this kernel.

Finally, we solve the system of differential equations given by Eqs. (6.1)-(6.3) in Python, using an explicit fourth order Runge-Kutta scheme with time step $\delta t = 0.1$h. Because the simulations are deterministic, we verified numerical convergence by repeating runs with identical initial conditions and smaller time steps, and confirmed that the resulting cell trajectories showed no appreciable differences. Initial cell positions were sampled uniformly at random within a circle of radius $R = 100$ μm and with a minimum interparticle distance of 10 μm. Initial velocities were similarly sam-

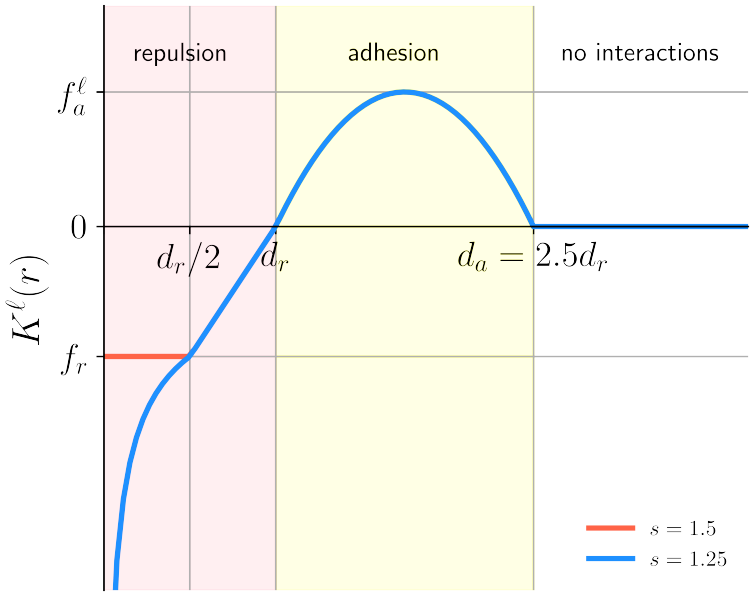


Figure 6.3: Representation of the interaction kernels $K^\ell(r)$. Parameters are not to scale.

pled uniformly at random, with each component of the velocity sampled uniformly on the interval $[-\sqrt{10}, \sqrt{10}]$ μm/h.

6.3.2 H-stability and adhesion parameter ranges

We use the concept of H-stability from statistical mechanics to constrain the range of admissible model parameters [46]. This constraint ensures that cells do not collapse to a single point as more cells are added to the system. In biological terms, this means that there is a sufficient balance between repulsive and attractive forces so that cells do not form clusters with a distance less than a cell diameter. If the kernels are not H-stable, then unphysical solutions can arise in which a very large number of particles occupies a small, finite region of space.

Mathematically, a kernel or potential is H-stable if the total potential energy of the system is bounded below by a constant that is independent of the number of particles. This concept has been applied in the context of interacting models of cell migration and is well-characterised for the kernel in Eq. (6.6) [46][2]. In particular, given an interaction

[2]This is another motivating reason for why we have used the kernels described above.

potential $u_\ell(r)$, such that $K^\ell(r) = u'_\ell(r)$, K^ℓ is called H-stable if

$$\int_0^{+\infty} u_\ell(r) r \, \mathrm{d}r > 0 \, .$$

If the integral is negative, then the kernel is not H-stable (also called *catastrophic*). For the kernel in Eq. (6.6), the integral above yields the following constraint on the strength of repulsive and attractive forces:

$$\frac{f_r}{f_a^\ell} > \frac{32s(d_a - d_r)(3d_a^2 + 4d_a d_r + 3d_r^2)}{5(11s + 6)d_r^3} := F_* \, . \tag{6.7}$$

Following these ideas, in numerical simulations we vary the cell–cell and cell–matrix adhesion strengths in $f_a^\ell \in [0, F_c]$, with $F_c := f_r / F_*$ so that the potential is always H-stable.

To illustrate the impact of H-stability, we present snapshots of numerical simulations in which the kernels are not H-stable — see Figure 6.4. Depending on the H-stability of the cell–cell and cell–matrix interactions, the system either collapses to the centre of the spheroid or to its boundary.

6.3.3 Model parameters

We set the radius of the circular matrix boundary, $R = 100$ μm, to have a comparable length scale to the experimental spheroids. For simplicity, the length scales of repulsive and adhesive forces, d_r and d_a, and the strength of repulsive forces f_r, are assumed to be identical for cell–cell and cell–matrix interactions. The parameter d_r can be estimated from cell–cell spacing observed in experiments, while d_a can be similarly estimated by noting that MCF-10A cells do not appear to interact with neighbours beyond 2-3 cell diameters. This leads us to fix $d_r = 15$ μm, $d_a = 2.5 d_r = 37.5$ μm. The experimental data also provide sufficient information to fix the asymptotic cell speed, which is roughly equal to $u = \sqrt{\alpha/\beta} = 3$ μm/h. The average cell velocity in numerical simulations, however, is higher than u — around 10 μm/h, agreeing with experimental observations — due to the additional tangential acceleration provided by cell–matrix adhesion[3]. Fixing the cell speed sets the ratio α/β but not their absolute magnitudes.

[3]Note that the additional tangential acceleration arises from the velocity dependence of the cell–matrix forces.

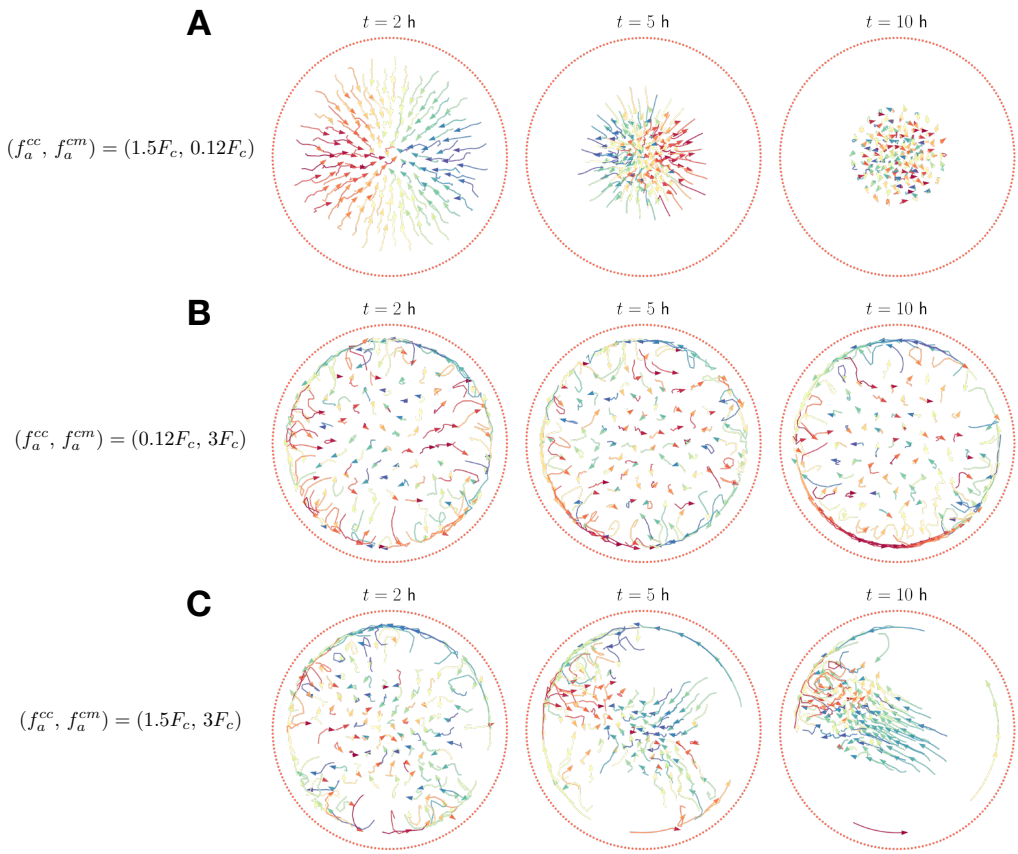


Figure 6.4: Interaction kernels that are not H-stable lead to cell collapse and unphysical behaviour. Snapshots of numerical simulations where: (A) K^{cc} is not H-stable and K^{cm} is H-stable; (B) K^{cc} is H-stable and K^{cm} is not H-stable; and (C) neither is H-stable. Arrows represent each cell's velocity vector direction. With the exception of adhesion strengths, all model parameters are listed in Table 6.1.

Hence the value of u alone does not uniquely determine α and β. This choice is discussed later.

The number of discrete matrix particles, M, is chosen to be sufficiently large to avoid cell escape from the spheroid, yet also small enough to ensure that the net cell–matrix force on a given cell is no more than an order of magnitude larger than the cell–cell repulsion force. We have found that, for the parameters used here, $M = 200$ molecules is sufficient to achieve both goals. For a circular matrix boundary of radius $R = 100$ μm, this corresponds to a one-dimensional matrix density of 0.32 points/μm, which is approximately five times larger than the one-dimensional density corresponding to the external layer of orbiting cells. This corresponds to a mean nearest neighbour distance of roughly 1-2 μm, on the same order as the $\sim$ 0.7 μm fiber spacing in the collagen

hydrogel used in experiments [128].

The number of cells, N, is similarly chosen to maximise the number of individuals in the spheroid — in order to avoid the formation of empty voids — while simultaneously limiting the degree of overlap between them, as otherwise repulsive forces overly dominate cell movement and uncoordinated behaviour occurs. For cells of radius $d_r/2 = 7.5\ \mu$m packed in a larger circle of radius $R = 100\ \mu$m, this density we found when $N >$ 130-150 cells, the spheroid dynamics are dominated by repulsive forces, with adhesion playing a negligible role. Consequently, we set $N = 130$ cells for the simulations.

We have thus far fixed all but four parameters in the mathematical model, with α, f_r, f_a^{cc}, and f_a^{cm} remaining undetermined. The adhesion-related parameter values are constrained by Eq. (6.7), ensuring that the cell–cell and cell–matrix potentials are H-stable. As it is unclear what are reasonable values for these parameters, we sweep over various values to determine how they influence the long-term system behaviour. Meanwhile, the parameter f_r sets the timescale to reach the equilibrium orbiting configuration. All simulations have $f_r = 100\ \mu\text{m/h}^2$ as we observe in numerical simulations that this corresponds to an equilibrium timescale of around $10 - 20$ h. Finally, α^{-1} represents the timescale at which cells readjust their velocities as a consequence of active and drag forces. To prevent cells from escaping the spheroid, the acceleration caused by cell–matrix repulsion must dominate over the acceleration arising from active–drag forces, which drives changes in the velocity $\delta v \sim u$ on timescales approximately equal to α^{-1}. Consequently, the condition $f_r \gg u\alpha$ is fulfilled, and we set $\alpha = 0.01, 0.1, 1\,\text{h}^{-1}$, all of which satisfy this requirement.

We summarise in Table 6.1 the model parameters and their chosen values. We also study in the next sections the impact of varying different model parameters.

6.3.4 Quantifying collective orbiting: angular momentum

We quantify orbiting using the normalised angular momentum in the z direction about the spheroid centre, $\hat{L}$, as an order parameter. We define it as

$$\hat{L} = \frac{1}{N}\left|\sum_{i=1}^{N} \hat{\mathbf{x}}_i \times \hat{\mathbf{v}}_i\right|, \tag{6.8}$$

Table 6.1: Summary of model and simulation parameters and their values used in numerical simulations

Parameter	Values	Units	Meaning
α	10^{-2}, 10^{-1} , 10^{0}	h^{-1}	self-propulsion coefficient
u	3	$\mu\mathrm{m/h}$	equilibrium cell speed in the absence of cell–cell and cell–matrix forces
β	α/u^2	$\mathrm{h}/\mu\mathrm{m}^2$	friction coefficient
d_r	15	$\mu\mathrm{m}$	repulsion interaction range
d_a	$2.5d_r$	$\mu\mathrm{m}$	adhesion interaction range
f_r	100	$\mu\mathrm{m/h}^2$	cell–cell and cell–matrix repulsion strength
F_c	5.18	$\mu\mathrm{m/h}^2$	critical adhesion strength for H-stability (Eq. (6.7))
f_a^{cc}	$[0, 0.5F_c]$, $1.5F_c$ (Fig. 6.4)	$\mu\mathrm{m/h}^2$	cell–cell adhesion strength
f_a^{cm}	$[0, 0.5F_c]$, $3F_c$ (Fig. 6.4)	$\mu\mathrm{m/h}^2$	cell–matrix adhesion strength
$\theta_{\max}$	$\pi/3$, $[0, \pi]$ (Fig. 6.10)	rad	maximum cell–matrix alignment angle (Eq. (6.5))
s	1.25	—	exponent controlling the repulsive force behaviour at the origin
ρ	0.004	$\mathrm{cells}/\mu\mathrm{m}^2$	cell density
N	130 (circular boundary) $+ n_p\rho\Delta A$ (boundary perturbations, Eq. (6.13))	cells	number of cells
M	200	points	number of matrix points
R	100	$\mu\mathrm{m}$	spheroid radius
h_p	$[0, 40]$	$\mu\mathrm{m}$	height of boundary perturbation
k_p	$[0, 80]$	$\mu\mathrm{m}$	width of boundary perturbation
n_p	$1, 2, 3$	—	number of boundary perturbations

where $\hat{\mathbf{x}}_i$ denotes the unit position vector of cell i and $\hat{\mathbf{v}}_i$ its unit velocity vector. This order parameter satisfies $\hat{L} \geq 0$ and also

$$\hat{L} \leq \frac{1}{N} \sum_{i=0}^{N} |\hat{\mathbf{x}}_i \times \hat{\mathbf{v}}_i| \leq 1\,,$$

with equality when each position vector is orthogonal to its corresponding velocity vector. This occurs in a perfectly coordinated state where all cells orbit either clockwise or anticlockwise. Hence, values of the angular momentum close to unity indicate global orbiting. By contrast, order parameter values close to zero are indicative that most cells do not exhibit collective orbiting — although they could still exhibit coordinated motion, e.g. flocking.

Moreover, to quantify the time to reach the orbiting state, τ, we first average $\hat{L}$ across the simulation replicates to obtain $\langle \hat{L} \rangle_i$. We then fit $\langle \hat{L} \rangle_i$ as a function of time, to a sigmoid function using non-linear least squares. The sigmoid function, defined as

$$L_0 + \frac{L_{\text{eq}}}{1 + \exp\left(-(t-\tau)/b\right)}, \tag{6.9}$$

depends on time, t, and on four parameters: the initial angular momentum, L_0, which we expect to be close to zero given the random initial conditions; the equilibrium angular momentum, L_{eq}; the time at which $\hat{L} = L_0 + L_{\text{eq}}/2$, τ; and a parameter that quantifies the steepness of the sigmoid, b.

6.4 Collective orbiting in multicellular spheroids

Here we present the main results from numerically solving the model, including a detailed analysis of the orbiting regimes and their stability under boundary perturbations.

6.4.1 Circular spheroids: orbiting and phase diagram

First, we systematically varied parameters associated with cell–cell and cell–matrix adhesion forces acting on cells confined within a perfect circular boundary to evaluate whether these interactions can stabilise collective orbiting over long time periods. We vary the strengths of cell–cell and cell–matrix adhesion, f_a^{cc} and f_a^{cm}, respectively,

within the interval $[0, 0.5F_c]$, as predicted by the H-stability analysis. These values are chosen because, for strong cell–cell adhesion ($f_a^{cc} > 0.5F_c$), we observed parameter combinations that do not provide sufficient coverage of the spheroid. The space $[0, 0.5F_c] \times [0, 0.5F_c]$ is discretised into a 40×40 grid, for each parameter set, we simulate the model fifty times to a final time of 300 h. All other parameters are kept fixed, and numerical simulations differ only in their initial conditions, which are generated randomly. We then average the order parameter over the last 100 h of the numerical simulation and across the fifty different simulations. We denote by $\langle\cdot\rangle_i$ averages with respect to the simulation replicates, and by $\langle\cdot\rangle_t$ the averages with respect to time.

We found migration was arrested in the limit of weak cell–matrix adhesion and strong cell–cell adhesion — see Figure 6.5. Globally coordinated orbiting occurred in an intermediate regime with slightly stronger cell–matrix adhesion and slightly weaker cell–cell adhesion. Further increases in the cell–matrix adhesion strength parameter resulted in peripheral orbiting of the outermost layers but uncoordinated migration of interior cells — see also Figure 6.6D-F for snapshots at different times of the three behaviours predicted by the model.

Numerically, we observe that cell–matrix interactions drive the emergence of collective orbiting within the first few hours of the simulation. When cell–matrix adhesion is weak, collective orbiting does not appear within the first $\sim 100 - 200$ h. However, in a small fraction of simulations, global orbiting eventually emerges over much longer timescales — see Figure 6.6A,C. In these cases, the average normalised angular momentum, $\langle \hat{L} \rangle_{i,t}$, takes up to $\sim 10^3$ h to reach equilibrium, which we confirmed with additional simulations. Here, collective orbiting arises purely due to geometric boundary constraints, as cell–matrix adhesion is almost negligible, leading to a significantly longer coordination timescale[4]. To identify such cases, we check whether the increase in the order parameter during the last 100 h is greater than 0.1 — corresponding to a slope of more than 0.001 in a linear regression fit. Simulations meeting this criterion are labelled as *unclassified* and appear white in the phase diagram.

[4]These cases are of limited biological interest for multicellular spheroids, since the coordination timescales are orders of magnitude longer than in the experiments.

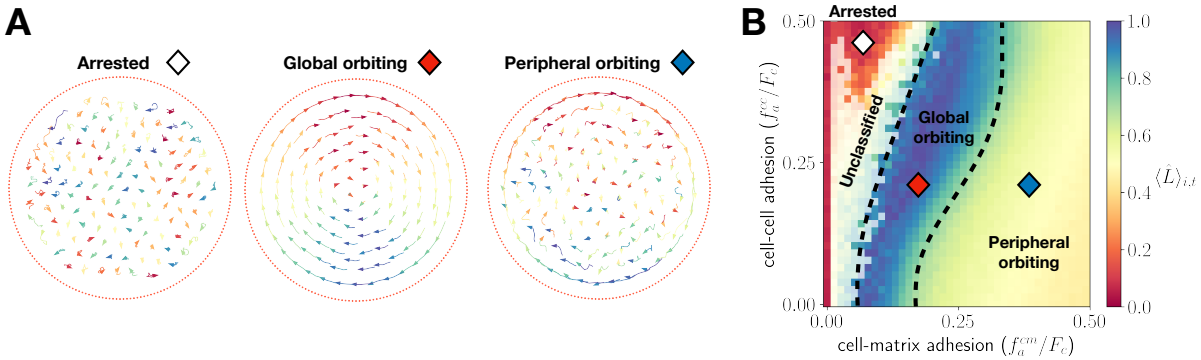


Figure 6.5: Elucidating the impact of cell–cell, cell–matrix adhesion, and boundary geometry in collective orbiting. With the exception of adhesion strengths, all model parameters are listed in Table 6.1 (with $\alpha = 10^{-1}$ h^{-1}). (A) Representative simulations of uncoordinated (white diamond), global orbiting (red diamond) and peripheral orbiting (blue diamond) states, and (B) the corresponding phase diagram of the angular order parameter $\langle L \rangle_{i,t}$ over the last 100 hours of the simulation across various cell–cell and cell–matrix adhesion strength parameters. Upper bounds on adhesion strengths are set to prevent void formation within the spheroid. The dashed lines on the phase diagram are hand-drawn and approximate the boundaries of the three states. Adhesive parameters $(f_a^{cc}/F_c,\ f_a^{cm}/F_c)$, from left to right, correspond to $(0.05, 0.45)$, $(0.2, 0.2)$, and $(0.2, 0.4)$.

6.4.2 Scaling with number of orbiting layers

In experiments and numerical simulations we often observe states in which only cells within the external layers of the spheroid orbit, while cells near the centre remain uncoordinated. To quantify this, we study how the angular momentum scales with the number or orbiting layers, n_{orb}. Given the typical distance between cells, a, and the outer radius of the spheroid, R, n_{orb} is a non-negative integer between 0 and R/a. Next, we observe that the number of orbiting cells in a layer of radius r_j is approximately $2\pi r_j/a$, where the allowed values for r_j are given by $r_j = R - ja$, with $j = 1, \ldots, n_{\text{orb}}$. By assuming that uncoordinated cells negligibly influence the total angular momentum of the system and all cells within orbiting layers are perfectly aligned, we can approximate Eq. (6.8) as

$$\hat{L} \approx \frac{2\pi}{N} \frac{(R-a) + (R-2a) + \ldots + (R - n_{\text{orb}}a)}{a} = \frac{2\pi n_{\text{orb}}}{N} \left(\frac{R}{a} - \frac{n_{\text{orb}} + 1}{2} \right) . \quad (6.10)$$

When the maximum number of possible layers $n_{\text{orb}} = R/a \gg 1$, this approximation further reduces to

$$\hat{L} \approx \frac{2\pi R}{Na} \left(\frac{R}{2a} - \frac{1}{2} \right) \approx \frac{\pi R^2}{Na^2} = 1 ,$$

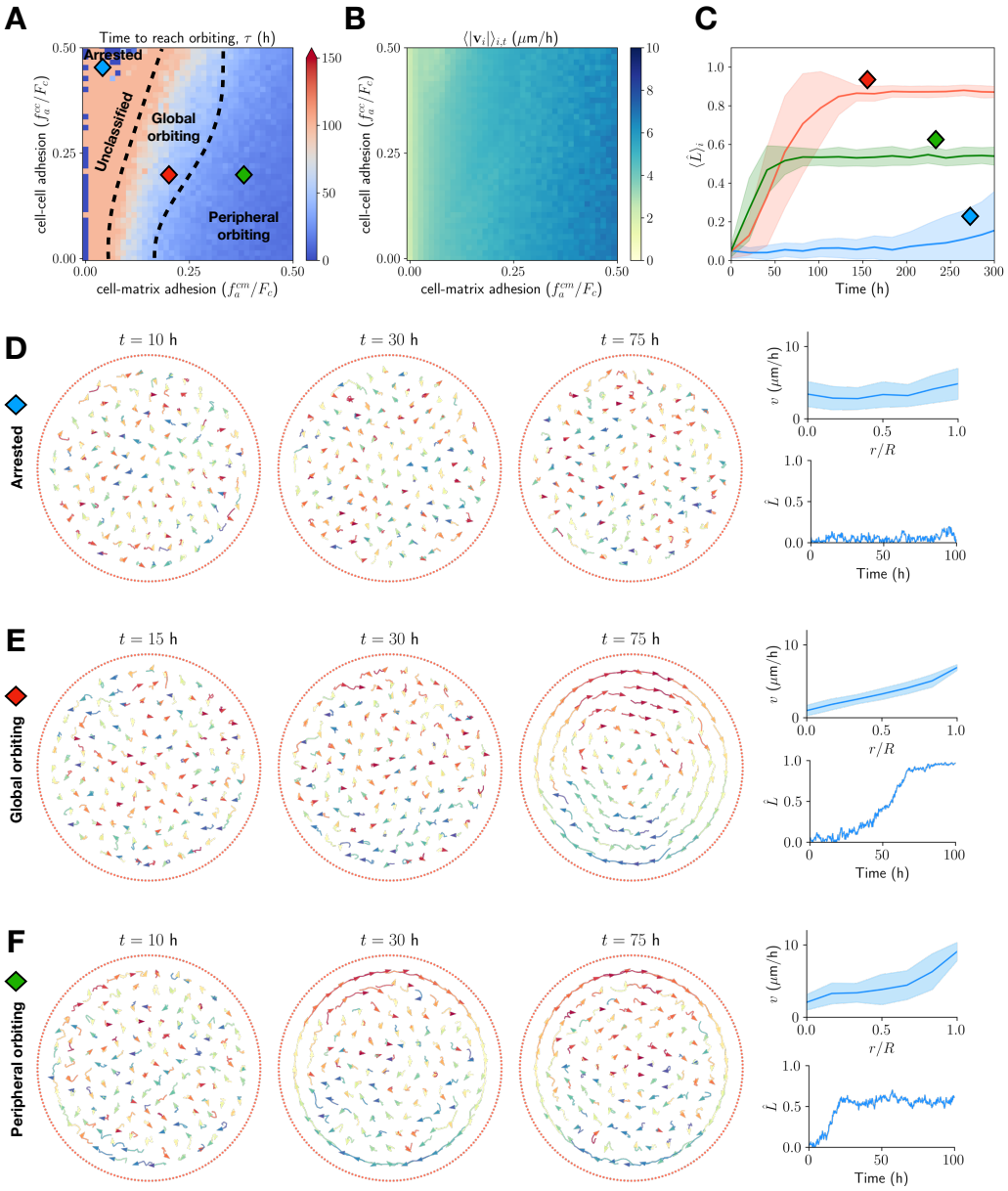


Figure 6.6: (A) Phase diagram of the timescale τ required to reach orbiting, as a function of the relative strength of cell–cell and cell–matrix adhesion. The dashed lines on the phase diagram are hand-drawn and approximate the boundaries of the three regimes. (B) Phase diagram of mean cell speed at equilibrium, which depends on active–drag and cell–matrix forces. (C) Angular momentum order parameter dynamics in three distinct regions of the phase diagram: arrested motion (blue), global orbiting (red), and peripheral orbiting (green). (D-F) Representative snapshots of the three behaviours predicted by the model (arrested motion, global orbiting, and peripheral orbiting) at $t = 10$, 30, 75 h, along with mean cell speed as a function of spheroid radius and angular momentum order parameter dynamics. With the exception of adhesion strengths, all model parameters are listed in Table 6.1 (with $\alpha = 10^{-1}$ h^{-1}).

which is consistent with the expected angular momentum value for a global orbiting state.

Identifying the typical distance between cells as $a \sim \rho^{-1/2} = R\sqrt{\pi/N}$, Eq. (6.10) becomes independent of the geometry of the spheroid giving

$$\hat{L} \approx \frac{2\pi n_{\text{orb}}}{N}\left(\sqrt{\frac{N}{\pi}} - \frac{n_{\text{orb}}+1}{2}\right). \tag{6.11}$$

In particular, for our numerical simulations with $N = 130$, the order parameter should correspond to $\hat{L} \sim 0.26, 0.48, 0.64, 0.76$ for one, two, three, and four layers, respectively.

Studying how the order parameter $\hat{L}$ scales with the number of orbiting layers allows us to partition the observed behaviour into three distinct regimes, as shown in the discretised phase diagram (Fig. 6.7): an uncoordinated (arrested) regime and two qualitatively different orbiting modes, global and peripheral, distinguished by thresholds in $\langle \hat{L} \rangle_{i,t}$. Only the peripheral orbiting mode, in which coordination is confined to outer layers, is observed experimentally, which enables us to restrict the biologically relevant region of parameter space. This scaling analysis thus provides a quantitative basis for identifying relevant collective behaviours in the model.

6.4.3 Boundary perturbations

All numerical simulations in the previous sections used a circular matrix boundary, given by the position of discrete particles $\mathbf{y}_l = (R\cos 2\pi l/M, R\sin(2\pi l/M))$ for $l = 1, \ldots, M$. To investigate how the onset of invasive branches formed by groups of cells disrupts experimental orbiting, we introduce a boundary perturbation with height h_p, and width k_p. The perturbed matrix boundary can be parametrised by $\theta \in [0, 2\pi]$, with the peak of the perturbation found at $\theta = \pi$:

$$\mathbf{y}(\theta) = R_p(\theta)\,(\cos\theta, \sin\theta)\,, \quad R_p(\theta) = R + h_p \exp\{-4R^2(\theta-\pi)^2/k_p^2\}\,.$$

The case of boundary perturbations at n_p different locations can be considered analogously. In numerical simulations, we generate matrix points with a fixed density ($200/2\pi R \approx 0.32$ cells/μm).

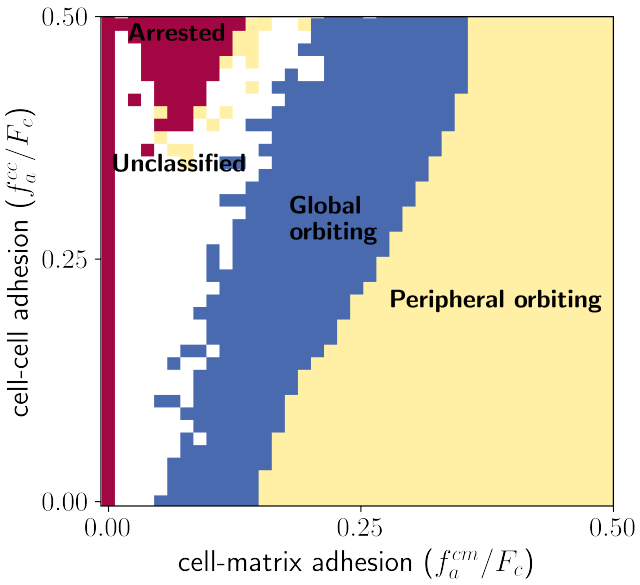


Figure 6.7: The mathematical model predicts three regimes based on collective orbiting behaviour. States are classified according to the number of orbiting layers, as determined by the normalised angular momentum: arrested — $\langle \hat{L} \rangle_{i,t} < 0.24$, no layers —, peripheral orbiting — $0.24 \leq \langle \hat{L} \rangle_{i,t} < 0.76$, one to three layers —, and global orbiting — $\langle \hat{L} \rangle_{i,t} \geq 0.76$, four or more layers. Unclassified states reach equilibrium only after very long times. Upper bounds on adhesion strengths are set to prevent void formation within the spheroid. With the exception of adhesion strengths, all model parameters are listed in Table 6.1 (with $\alpha = 10^{-1}\ \mathrm{h}^{-1}$)

Moreover, to keep constant cell density inside the spheroid, we calculate the change in area due to one boundary perturbation

$$\begin{aligned}
\Delta A &= \frac{1}{2}\int_0^{2\pi} R_p(\theta)^2\,\mathrm{d}\theta - \pi R^2 \\
&= R r_p \int_0^{2\pi} e^{-4R^2(\theta-\pi)^2/k_p^2}\,\mathrm{d}\theta + \frac{r_p^2}{2}\int_0^{2\pi} e^{-8R^2(\theta-\pi)^2/k_p^2}\,\mathrm{d}\theta \\
&\approx \sqrt{\pi} r_p \frac{k_p}{2} + \sqrt{\pi} r_p^2 \frac{k_p}{4\sqrt{2}R} \qquad (6.12) \\
&= \sqrt{\pi} r_p \frac{k_p}{2}\left(1 + \frac{r_p}{2\sqrt{2}R}\right). \qquad (6.13)
\end{aligned}$$

In numerical simulations we consider a number of cells given by $N = \rho(\pi R^2 + n_p \Delta A)$.

We explored orbiting stability by varying the boundary perturbation parameters over the ranges

$$0 \leq h_p \leq 40\,\mu\mathrm{m} \quad \text{and} \quad 0 \leq k_p \leq 80\,\mu\mathrm{m},$$

averaging angular momentum over the final 100 h of each simulation. We extended

each simulation run to 500 h to ensure convergence, with all other model parameters held constant, and adhesive forces fixed at $(f_a^{cc}, f_a^{lm}) = (0.2F_c, 0.2F_c)$. All remaining parameters match those used in the circular-boundary case (see Table 6.1).

In simulations with a single outward perturbation, collective orbiting remained robust — cells followed the circular boundary and simply detoured around the misaligned region (Figure 6.8A,B). In contrast, two sharp features — each taller than one cell diameter ($h_p > d_r$) and wider than two cell diameters ($k_p > 2d_r$) — sufficed to disrupt orbiting (Figure 6.8C,D). Under these conditions, cells encountered multiple obstacles and lacked continuous curved boundary segments to follow. Notably, the number and dimensions of these perturbations align with the experimental observations [128] (Figure 6.8E,F). We can summarise these findings as follows: once curvature defects exceed a critical size and number they disrupt the matrix guidance, preventing the maintenance of coherent orbiting.

6.4.4 Mosaic spheroids

Finally, we analysed collective orbiting in a mixture of two cell types with distinct adhesion phenotypes, as in the mosaic spheroid experiments of *Kim et al.* [128] (Figure 6.2). We thus extend the model to consider two cell species. To model Snail1 induced effects on MCF-10A cells, in the model, Snail cells had reduced cell–cell adhesion (mimicking E-cadherin downregulation), increased cell–matrix adhesion (mimicking vimentin upregulation), and isotropic matrix sensing (promoting invasive motility and higher traction forces), in contrast to wildtype (WT) cells. We consider a version of the mathematical model with two cell species. Both WT (1) and Snail cells (2) follow the same

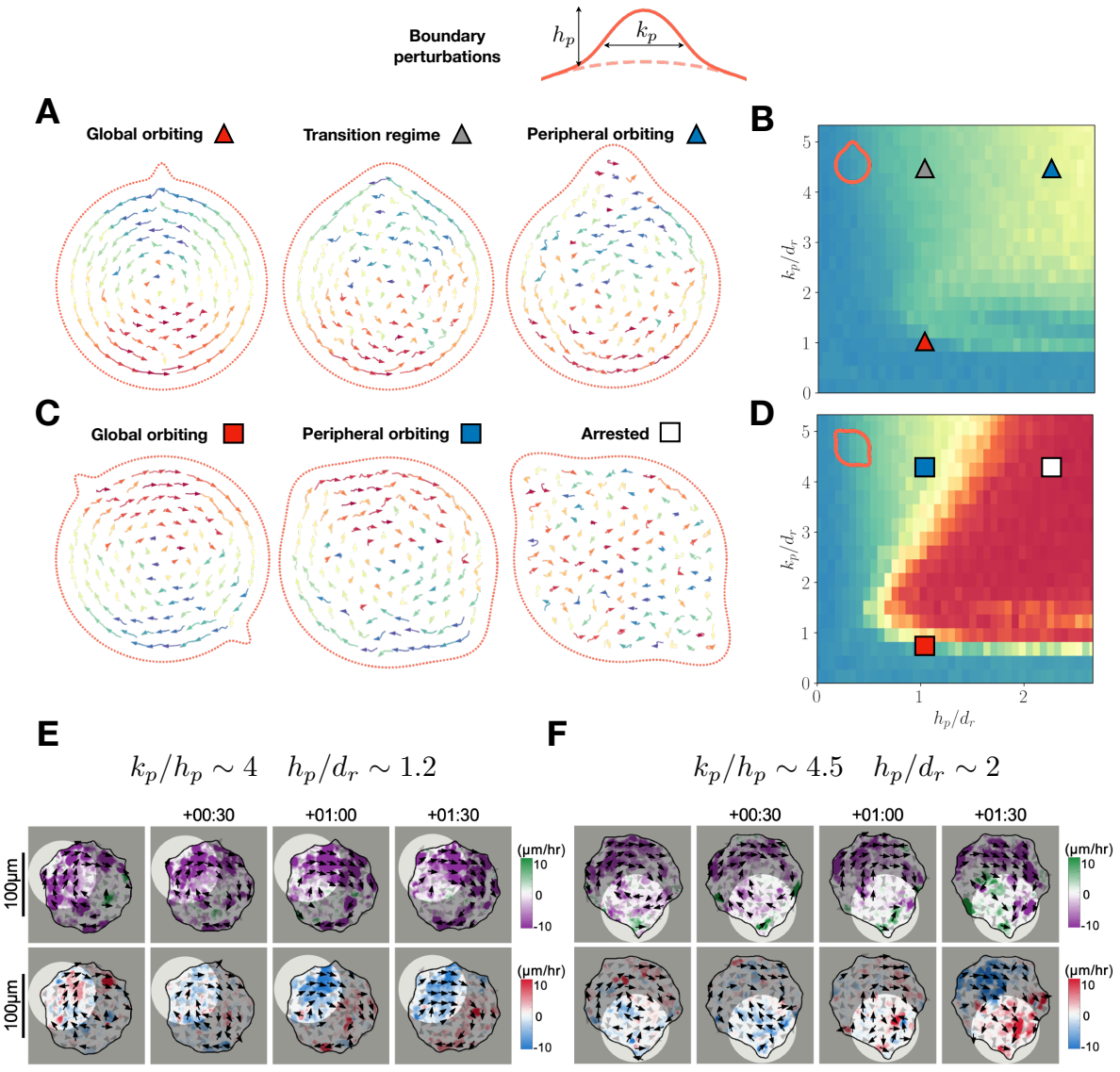


Figure 6.8: Stability of orbiting to matrix boundary perturbations. (A,B) Representative snapshots of global (red triangle), transitory (grey triangle) and peripheral orbiting (blue triangle) associated with a single boundary perturbation, and the corresponding phase diagram of the average order parameter. Boundary perturbation parameters, $(h_p/d_r, k_p/d_r)$, from left to right, given by $(1,1)$, $(1,4)$, and $(2.4,4)$. The colourbar is the same as in Figure 6.5. (C, D) Representative snapshots of global orbiting (red square), peripheral orbiting (blue square) and uncoordinated migration (white square) with two oppositely-oriented boundary perturbations, and the corresponding phase diagram. Boundary perturbation parameters, from left to right, given by $(1,1)$, $(1,4)$, and $(2.4,4)$. With the exception of adhesion strengths, all model parameters are listed in Table 6.1 (with $\alpha = 10^{-1}\,\mathrm{h}^{-1}$). (E,F) Velocity profiles from experimental data of spheroids with a comparable perturbation size. White circles represent invasive branches locations.

equations, but with different model parameters,

$$\frac{\mathrm{d}\mathbf{x}_i^{(1)}}{\mathrm{d}t} = \mathbf{v}_i^{(1)},$$
$$\frac{\mathrm{d}\mathbf{v}_i^{(1)}}{\mathrm{d}t} = \left(\alpha - \beta\left|\mathbf{v}_i^{(1)}\right|^2\right)\mathbf{v}_i^{(1)} + \sum_{l\neq i,\, j=1}^{N_1} \mathbf{F}_{11}^{cc}(\mathbf{x}_l^{(1)} - \mathbf{x}_i^{(1)}) + \sum_{l=1}^{N_2} \mathbf{F}_{12}^{cc}(\mathbf{x}_l^{(2)} - \mathbf{x}_i^{(1)}) + \sum_{k=1}^{M} \mathbf{F}_1^{cm}(\mathbf{y}_k - \mathbf{x}_i^{(1)}, \mathbf{v}_i^{(1)}),$$
$$\frac{\mathrm{d}\mathbf{x}_j^{(2)}}{\mathrm{d}t} = \mathbf{v}_j^{(2)},$$
$$\frac{\mathrm{d}\mathbf{v}_j^{(2)}}{\mathrm{d}t} = \left(\alpha - \beta\left|\mathbf{v}_j^{(2)}\right|^2\right)\mathbf{v}_j^{(2)} + \sum_{l=1}^{N_1} \mathbf{F}_{21}^{cc}(\mathbf{x}_l^{(1)} - \mathbf{x}_j^{(2)}) + \sum_{l\neq j,\, l=1}^{N_2} \mathbf{F}_{22}^{cc}(\mathbf{x}_l^{(2)} - \mathbf{x}_j^{(2)}) + \sum_{k=1}^{M} \mathbf{F}_2^{cm}(\mathbf{y}_k - \mathbf{x}_i^{(2)}, \mathbf{v}_i^{(2)}),$$
$$\frac{\mathrm{d}\mathbf{y}_k}{\mathrm{d}t} = \mathbf{0},$$

where N_1 and N_2 are the number of WT and Snail cells respectively, $\mathbf{F}_{11}^{cc}$, $\mathbf{F}_{22}^{cc}$ represent adhesive-repulsive forces between the same type of cell; and $\mathbf{F}_{21}^{cc}$, $\mathbf{F}_{12}^{cc}$ represent cross-interaction forces.

Table 6.2: Summary of two-species model and simulation parameters and their values used in numerical simulations. The remaining model parameters (cell–cell repulsion, active–drag forces, and matrix geometry) are the same as in the one-species case (Table 6.1).

Parameter	Values	Units	Meaning
N_1	$\{85, 86, \ldots, 130\}$	cells	number of WT cells
N_2	$N - N_1$	cells	number of Snail cells
$f_{a_{1,1}}^{cc}$	$0.2F_c$	μm/h^2	cell–cell adhesion strength (WT-WT)
$f_{a_{1,2}}^{cc}$, $f_{a_{2,1}}^{cc}$	0	μm/h^2	cell–cell adhesion strength (WT-Snail)
$f_{a_{2,2}}^{cc}$	0	μm/h^2	cell–cell adhesion strength (Snail-Snail)
$f_{a_1}^{cm}$	$0.2F_c$, $0.4F_c$	μm/h^2	cell–matrix adhesion strength (WT-matrix)
$f_{a_2}^{cm}$	$0.8F_c$	μm/h^2	cell–matrix adhesion strength (Snail-matrix)
$\theta_{\max}^{(1)}$	$\pi/3$	rad	maximum cell–matrix alignment angle (WT)
$\theta_{\max}^{(2)}$	π , $\pi/3$ (Fig. 6.13)	rad	maximum cell–matrix alignment angle (Snail)

The forces, $\mathbf{F}_1^{cm}$ and $\mathbf{F}_2^{cm}$, model cell–matrix interactions for each cell type, similarly to the one-species case. These forces follow the expressions in Eqs. (6.4)-(6.6). We also note that active and drag forces are identical for both cell types. We summarise the chosen model parameters in Table 6.2. As in the previous section, we solve numerically

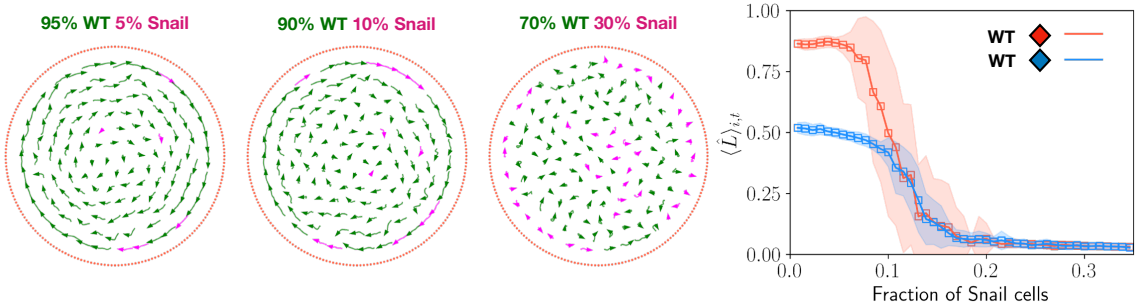


Figure 6.9: Numerical simulations of mosaic spheroids at increasing Snail cell fractions. Snapshots show that Snail cells — cell type 2 in the model, and with reduced cell–cell adhesion and increased cell–matrix attraction versus WT — sort to the spheroid boundary. Diamonds mark the WT adhesion parameters in the phase diagram of Figure 6.11. The adjacent plot of the angular order parameter versus Snail percentage reveals that, as more Snail cells occupy the edge, global orbiting breaks down and collective migration is disrupted. Model parameters can be found in Table 6.2.

the model until a final time of 500 h, and record angular momentum using the last 100 h of the simulation. Cells' positions and velocities are initialised as in the one-species case, and then assigned a type randomly.

We observed that collective orbiting remained robust for Snail cell fractions up to 10%, since WT cells at the periphery could migrate around isolated Snail cells (Figure 6.9). However, initialising simulations with higher Snail fractions — which have a stronger cell–matrix adhesion — led Snail cells to fully occupy the cell–matrix interface, which blocked WT cells from migrating along the boundary. These results demonstrate that boundary curvature supports coordinated orbiting under comparable cell–cell and cell–matrix adhesion — suggesting that this is likely so for the experimental spheroids and potentially in a range of developmental and pathological contexts — and that an increasing fraction of invasive cells at the edge can destabilise collective orbiting.

6.5 Additional simulations: impact of model parameters

In this section, we present additional numerical simulations of the model, where we test the impact of varying different model parameters.

6.5.1 Cell–matrix alignment angle

We begin by exploring the impact of the cell–matrix alignment angle $\theta_{\max}$ (see Figure 6.10), which sets the maximum angle between the cell velocity and the relative position to the spheroid boundary (Eq. (6.5)). As expected intuitively, we find that for small values of $\theta_{\max}$, orbiting is disrupted easily. The angular momentum parameter peaks at a value close to 1 (for comparable cell–cell and cell–matrix adhesion parameters), and for values close to those used in the previous figures ($\theta_{\max} = \pi/3$). Numerical results remain consistent with the rest of this chapter for larger values of the maximum cell–matrix alignment angle ($\theta_{\max} \sim 0.3\pi - 0.8\pi$). Interestingly, when $\theta_{\max}$ is close to π, meaning that cells can adhere to the matrix from nearly any direction except directly behind them, global orbiting appears to be more stable. This may be due to the increased range over which cells can detect and interact with the matrix, enhancing their ability to align with the boundary. However, in the limit of isotropic interactions, orbiting is completely disrupted ($\langle \hat{L} \rangle_{i,t} \sim 0$), highlighting the crucial role of anisotropy in cell–matrix interactions for the emergence of orbiting.

6.5.2 Active–drag forces

All the previous figures use a self-propulsion parameter of $\alpha = 0.1$ h^{-1}, setting a timescale for speed readjustments due to active–drag forces at approximately 10 hours. This choice yields realistic cell speeds. To explore the effect of this parameter on orbiting behaviour (Figure 6.11), we vary α by an order of magnitude ($\alpha = 10^{-2}$, 10^{0} h^{-1}) while adjusting the friction coefficient, β, to keep the equilibrium speed in the absence of cell–matrix interactions, $\sqrt{\alpha/\beta} = u$, constant.

In both cases, the phase diagram reveals regions of global orbiting, peripheral orbiting, and arrested motion (Figure 6.11). The timescales to reach these states remain largely unchanged, supporting our numerical findings that they are primarily governed by the strength of the interaction kernel. Additionally, Figure 6.11 suggests that cell speeds decrease with increasing self-propulsion strength, consistent with equilibrium speed being determined by the balance between active–drag and cell–matrix forces.

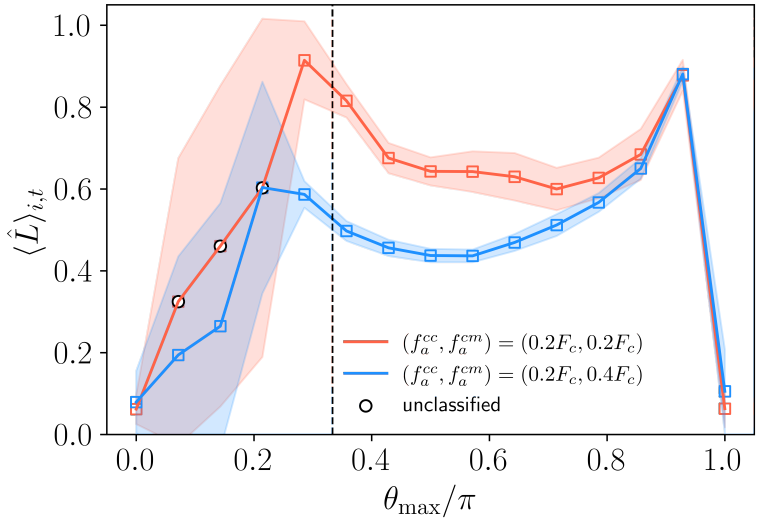


Figure 6.10: Impact of cell–matrix alignment on collective orbiting. All parameters are fixed except for $\theta_{\max}$ in Eq. (6.5), which is varied. Parameter values are taken from Table 6.1, with cell–matrix and cell–cell adhesion parameters chosen to represent the global (red) and peripheral orbiting (blue) regimes. Angular momentum is recorded over $t \in [200, 300]$ h and averaged across thirty simulations. Shaded regions indicate one standard deviation. The black dashed line denotes $\theta_{\max} = \pi/3$, which is the value used in the rest of figures and results.

While all three cases exhibit similar behaviour, when $\alpha = 10^{-2}\,\mathrm{h}^{-1}$, a larger portion of the phase diagram corresponds to higher values of the angular momentum order parameter, indicating a global orbiting state. This may result from cell–matrix interactions dominating over active–drag forces, allowing cells to synchronise within the spheroid. Conversely, stronger self-propulsion accelerates the onset of arrest in the strong cell–cell adhesion regime.

6.5.3 Three boundary perturbations

We further examined the effect of three outward-oriented Gaussian perturbations of the matrix and found a similar destabilising effect on collective orbiting as in the case of two perturbations. Beyond a critical perturbation size — approximately one cell diameter in height and two cell diameters in width — orbiting was disrupted (Figure 6.12). The phase diagram closely resembles the two–perturbation case, reinforcing that multiple sharp boundary features make coordinated migration more difficult by reducing the length of continuously curved boundary available for collective motion.

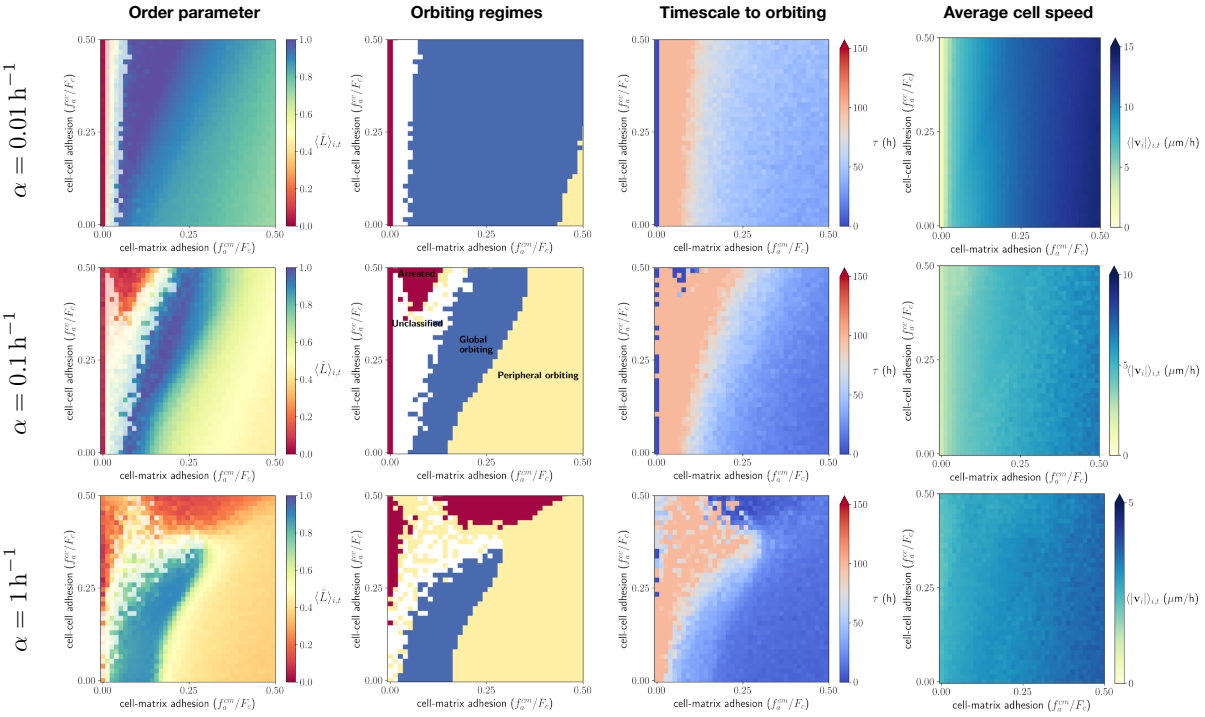


Figure 6.11: Impact of active–drag forces on collective orbiting. All parameters are fixed except for α (and $\beta = \alpha/u^2$ to keep u constant), which is varied. Parameter values are taken from Section 1. Angular momentum is recorded over $t \in [200, 300]$ h and averaged across fifty simulations. The columns, from left to right, represent the normalised angular momentum phase diagram, the discretized phase diagram of orbiting regimes, the timescale to orbiting (τ, Eq. (6.9)), and the mean cell velocity.

6.5.4 Differential adhesion and orbiting in mosaic spheroids

We also investigated the case where Snail cells exhibit directionality when sensing the matrix, implemented by setting the same maximum cell–matrix alignment angle as WT cells ($\theta_{\max}^{(2)} = \theta_{\max}^{(1)} = \pi/3$). In this scenario, orbiting remains stable in mixed cell populations, as both cell types are capable of orbiting independently (Figure 6.13). Interestingly, these simulations reveal classical sorting patterns driven by differential adhesion, considering both cell–cell and cell–matrix interactions. Since Snail cells exhibit weaker cell–cell adhesion but stronger cell–matrix adhesion, they preferentially localise at the boundary while maintaining peripheral orbiting. These findings highlight that disrupting orbiting requires a change in the directionality of cell–matrix interactions, as demonstrated in previous sections.

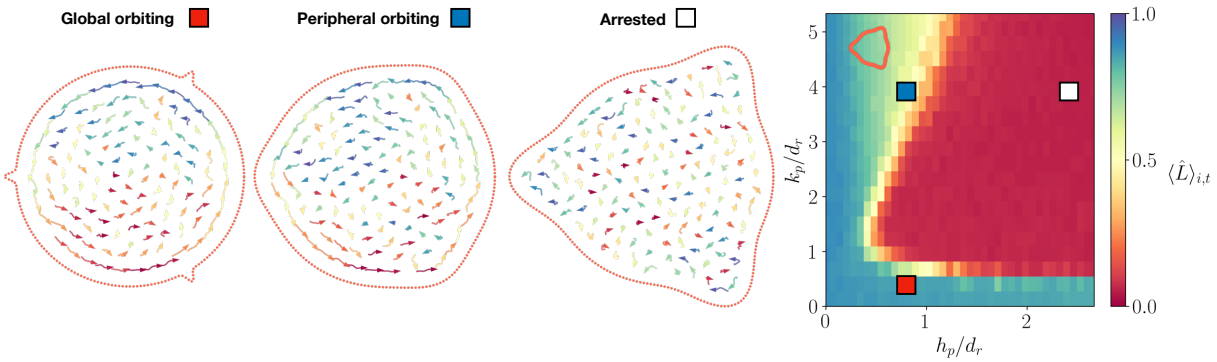


Figure 6.12: Orbiting is disrupted by three evenly spaced boundary perturbations. Representative snapshots of global orbiting (red square), peripheral orbiting (blue square) and arrested migration (white square) with three boundary perturbations, and the corresponding phase diagram depicting how the angular order parameter phase diagram responds to different perturbation heights and widths.

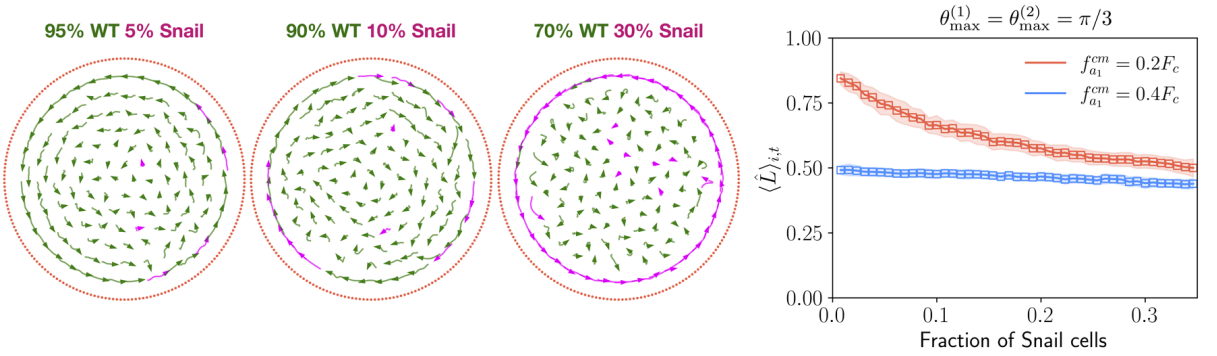


Figure 6.13: Orbiting and differential adhesion in two-species spheroids. Representative snapshots of numerical simulations with increasing percentages of Snail populations are shown. The graph on the right depicts how the average normalised angular momentum over the last 100 hours of the simulation varies with the fraction of snail cells. Red squares denote simulations in which the cell–matrix adhesion parameter for WT cells, $f^{cm}_{a_1}$, is equal to $0.2F_c$, while blue squares indicate simulations in which the same parameter is set equal to $0.4F_c$ (see Table 6.2 for more details). Shaded regions indicate one standard deviation.

6.6 Discussion and outlook

In this chapter, we have shown how differential adhesion between cells and between cells and a confining matrix boundary can drive a rich of collective behaviours — from coherent orbiting motion to the onset of localised invasive branching — in multicellular spheroids. The transition from spherical symmetry to elongated or branched architectures underlies many morphogenetic processes, from embryonic development to pathological invasion in cancer [61, 110]. Our results demonstrate that coordinated orbiting along a curved boundary stabilises tissue shape by producing coherent motion across cells, whereas perturbations to the boundary can create regions where collective motion is disrupted. In developmental contexts such as kidney or salivary gland branching, similar symmetry breaking events localise budding at defined positions [179]. Conversely, in carcinoma dissemination, focal protrusive invasion recapitulates an aberrant form of such developmental branching [93]. Dissecting how cell–cell cohesion and cell–matrix traction conspire to trigger these symmetry breaks is therefore central to understanding both normal organogenesis and tumour progression.

While our agent-based model treats the matrix boundary as static, experimental evidence (Figure 6.1) shows that cells both sense and remodel their surrounding matrix. Incorporating this dynamic reciprocity, in which matrix fibres realign and orient in response to cell generated forces, constitutes a critical next step. One promising approach is to endow matrix particles with simple remodelling rules [151], coupled to the existing cell dynamics. Such an extended model could capture feedbacks by which collective migration both shapes and is shaped by matrix microstructure, shedding light on how emergent tissue architectures arise from purely mechanical interactions [69].

Our numerical phase diagrams uncover clear regimes of global versus peripheral orbiting and identify how boundary perturbations can disrupt coordination — see Figure 6.11. To move beyond simulation and obtain analytical insight, it would be helpful to perform a linear stability analysis of the homogeneous orbiting state. Building on the work of *Albi et al.* [1], one could derive conditions on the interaction potentials under which small perturbations grow or decay. In the presence of periodic boundary modulations, such an analysis might predict preferred wavelengths or critical amplitudes for the breakdown of orbiting. These analytical results would complement the

computational phase diagrams (Figure 6.8) and offer predictive design principles for engineered tissues or biomaterials.

Agent-based models provide detailed insight into individual cell behaviours, but analytical tractability can be limited when scaling to larger cell numbers or more complex geometries. Recent advances in continuum modelling [32, 35, 48, 142, 152] could be adapted to account for anisotropic cell–matrix adhesion and polarity-driven migration. Continuum models would allow us to explore, for example, the role of domain boundary curvature in establishing orbiting motion, and could reveal parameter regimes for the onset of invasion across different spheroid sizes and adhesion regimes. Ultimately, such quantitative models may guide the engineering of organoids with designed morphogenetic programs or inform therapeutic strategies to arrest invasive transitions.

Chapter 7

Discussion and conclusions

7.1 Summary

The overarching aim of this thesis has been to develop, analyse and apply mechanistic continuum and discrete models that capture interactions between multiple cell populations. These models are then used to elucidate the mechanisms driving collective cell behaviour and inform practical applications in tissue engineering and developmental biology.

In Chapter 2 we analysed parameter identifiability in classical tissue-spreading models [89]. By combining both the Fisher–KPP and Porous–Fisher equations within a Bayesian inference framework and fitting to high-resolution epithelial expansion data [112], we obtained well-defined posterior distributions for all model parameters and demonstrated that crowding-driven motility (Porous–Fisher model) can be calibrated just as robustly as the classical linear diffusion Fisher–KPP model. This work complements recent studies on practical identifiability in similar partial differential equation models and highlights some of the trade-offs between model complexity and experimental data quality [144, 188]

Building on that foundation, Chapter 3 incorporated explicit cell cycle regulation into a minimal two-stage continuum model. Using spatiotemporal FUCCI measurements of epithelial densities, we fitted density-dependent cell cycle progression rates via Bayesian inference [90]. Our results quantitatively confirm that G1-to-S and G2 progression are delayed by crowding constraints with mechanical checkpoints that have been observed experimentally [75, 199]. Additionally, through travelling wave analy-

sis, we reveal how tissue–level proliferation patterns emerge from these local regulatory rules.

In Chapter 4 we turned to tissue–tissue interactions by formulating two–species versions of Fisher–KPP and Porous–Fisher equations for the study of monolayer collisions [89, 114]. Only the Porous–Fisher model with its density-dependent population pressure reproduces the sharply defined interfaces and interface-speed dynamic observed across a variety of collision geometries. After validating early-time dynamics with parameters inferred from single-tissue assays, we fully calibrated the collision assays, derived explicit scaling laws for interface motion both with and without proliferation, and provided a minimal yet predictive continuum framework for designing and interpreting tissue self-assembly experiments.

Chapter 5 addresses adhesion-driven patterning via a thin-film model of cell–cell adhesion. By taking a short-range limit of established nonlocal aggregation–diffusion equations [6, 56], we obtained a local fourth-order equation whose parameters admit a clear interpretation as relative surface tensions. Numerical simulations demonstrate that this local model recovers all classical sorting and engulfment patterns of the differential adhesion hypothesis, while remaining far more analytically and computationally tractable than its nonlocal predecessor. With this work, we take a significant step toward bringing cell–cell adhesion models, and more generally aggregation–diffusion systems, into close alignment with experimental data.

Finally, in Chapter 6 we explored collective migration in confined geometries using agent-based models that include both cell–cell and cell–matrix interactions. We showed how curvature and adhesion parameters drive a transition from coordinated orbiting to localised invasion strands. Our phase diagrams reveal regimes of global versus peripheral rotation, and suggest analytical avenues such as linear stability of the orbiting state, to predict symmetry-breaking thresholds in developmental and pathological contexts.

Taken together, these studies illustrate a coherent methodology: we derive or select mechanistic models, calibrate them rigorously against data, analyse their qualitative and quantitative behaviour, and then revisit the modelling assumptions in light of experimental observations. Looking forward, these advances set the stage for the next

generation of quantitative biology research, in which mechanistic modelling will be coupled with data-driven methods to deliver predictive and experimentally grounded frameworks.

7.2 Additional work

During the course of this thesis, additional projects and collaborations have emerged that, while related to its core themes, lie beyond its main scope. For the sake of completeness, we briefly summarise them below.

7.2.1 Nonlinear diffusion on networks

Motivated by our work on density-dependent diffusion in Chapters 2 and 4, we asked whether an analogous porous-medium type equation could be derived and studied on a discrete spatial domain, namely a network. Starting from the edge-centric nonlinear random walk on a graph, as proposed by [8, 43], we showed that in the continuum limit this discrete process converges to a continuum porous-medium type diffusion equation. Building on that continuum limit, we then turned back to finite networks to explore how the interplay between nonlinearity and topology shapes key dynamical features.

On infinite lattices and homogeneous trees, the emergent equation exhibits finite-speed propagation and self-similar spreading analogous to classical Barenblatt solutions [212]. On finite, connected graphs, the same macroscopic density-dependent diffusion coefficient governs both the long-time equilibrium, and the timescale over which the system relaxes to that equilibrium. On modular networks, we showed that the spectral gap of the graph Laplacian induces a clear separation of timescales — fast diffusion within each module and slow mixing between modules — just as in the linear case. Finally, we coupled our density-dependent diffusion law to logistic growth (a Fisher–KPP reaction term) and found that the nonlinear diffusion coefficient itself can drive a temporal transition: at early times proliferation dominates given the low densities and small diffusion, but once the local density exceeds a critical threshold (set by the nonlinearity), diffusion takes over and controls the subsequent spread.

Across all of these settings, the continuum limit provides not only a unified macroscopic description but also a diagnostic tool that separates the effects of nonlinear diffusion from those of the underlying network structure. This single-author project is published as

"C. Falcó. *From random walks on networks to nonlinear diffusion.* Physical Review E, 106(5):054103, 2022" [86].

7.2.2 Existence theory for local aggregation–diffusion equations

Upon deriving our local thin-film–type cell–cell adhesion model in Chapter 5, we noticed that no general existence theory was available for this fourth-order equation. A rigorous 2-Wasserstein gradient-flow theory for fourth-order equations was established in [143, 155], but the latter omits the backward-diffusion term, and the former focuses on more general density-dependent mobilities. To address this gap, we formulated and studied the more general equation on $\mathbb{R}^d$

$$\partial_t \rho \;=\; -\nabla\cdot\left(\rho\,\nabla\Delta\rho\right) \;-\; \chi\,\Delta\rho^m, \qquad m \geq 1,\ \chi > 0, \tag{7.1}$$

which recovers our adhesion model as the special case $m = 2$.

We exploit the 2-Wasserstein gradient-flow structure,

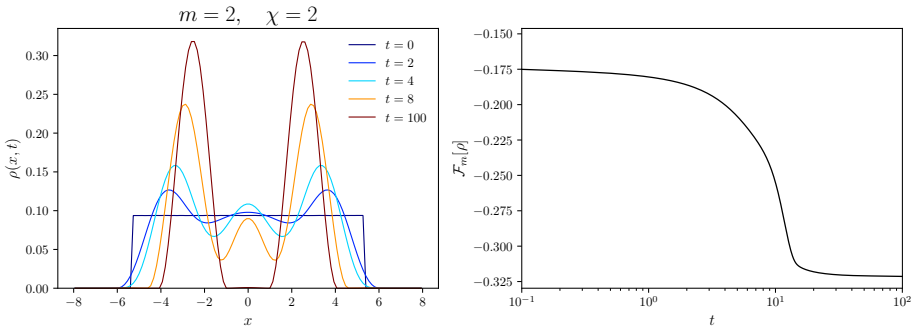
$$\partial_t \rho = \nabla\cdot\Big(\rho\,\nabla\frac{\delta\mathcal{F}}{\delta\rho}\Big), \qquad \mathcal{F}[\rho] = \frac{1}{2}\int_{\mathbb{R}^d} |\nabla\rho|^2\,\mathrm{d}\mathbf{x} - \frac{\chi}{m-1}\int_{\mathbb{R}^d}\rho^{m-1}\,\mathrm{d}\mathbf{x},$$

to prove global-in-time existence of weak solutions whenever $m < m_c$ or $m = m_c$ and $\chi < \chi_c$.

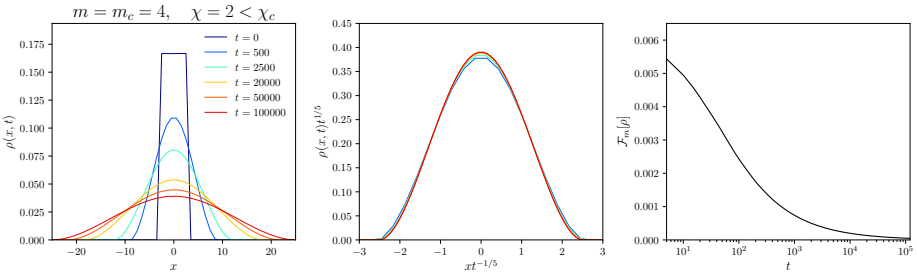
The *critical* exponent, m_c, arises by observing that under the mass-preserving dilation, $\rho_\lambda(x) = \lambda^d\,\rho(\lambda x)$,

$$\mathcal{F}[\rho_\lambda] = \frac{\lambda^{d+2}}{2}\int_{\mathbb{R}^d} |\nabla\rho|^2\,\mathrm{d}\mathbf{x} - \frac{\chi\lambda^{d(m-1)}}{m-1}\int_{\mathbb{R}^d}\rho^m\,\mathrm{d}\mathbf{x}\,.$$

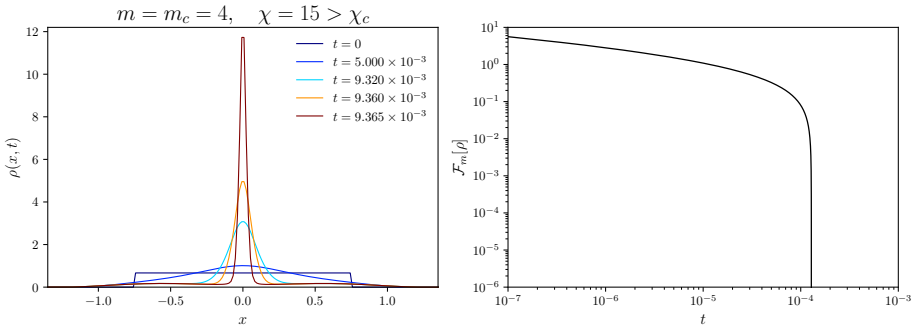
Hence the two competing terms scale identically precisely when $m = m_c := 2 + 2/d$, while for $m > m_c$ one shows $\mathcal{F}$ is unbounded below by letting $\lambda \to \infty$ — where the aggregation term dominates: $d(m-1) > d+2$. In the *subcritical* exponent regime $m < m_c$, a Gagliardo–Nirenberg inequality gives a uniform lower bound on $\mathcal{F}$. In the



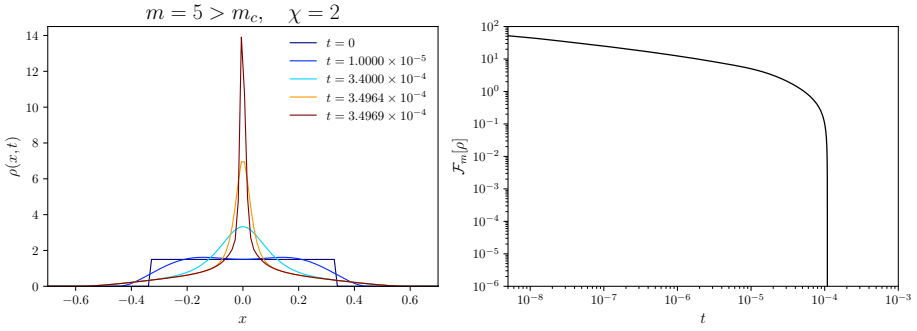
(a) Subcritical exponent $m < m_c$.



(b) Critical exponent $m = m_c$, subcritical mass $\chi < \chi_c$.



(c) Critical exponent $m = m_c$, supercritical mass $\chi > \chi_c$.



(d) Supercritical exponent $m > m_c$.

Figure 7.1: Numerical solutions to Eq. (7.1) in one spatial dimension for different values of m and χ, and decay of the free energy $\mathcal{F}[\rho]$ as a function of time.

critical exponent case $m = m_c$, the sharp inequality yields boundedness of $\mathcal{F}$ as long as $\chi \leq \chi_c$, which plays the role of a critical mass — see rescaling argument in [51]. The results on the free energy, together with snapshots of numerical simulations of the local aggregation–diffusion equation are shown in Figure 7.1. This threshold behaviour closely mirrors the well-known critical mass phenomenon in the Keller–Segel chemotaxis model with a power-law interaction potential [23, 47, 54].

With a bounded from below free energy, we use the JKO or minimising-movement scheme — developed by *Jordan, Kinderlherer and Otto* [123] — to construct global-in-time weak solutions whenever $m < m_c$. We exploit the *flow interchange* technique developed in [155] to obtain a H^2 estimate for the constructed solution. Finally, exploiting the same variational structure and symmetry in the cross-diffusion terms, we extended these results to the two-species system derived in Chapter 5 showing global existence under the natural positivity and definiteness conditions on the fourth-order diffusion matrix ($\kappa - \alpha^2 > 0$).

This work thus fills a crucial analytical gap for local aggregation–diffusion equations and reveals a deep analogy with Keller–Segel critical mass phenomena. It was published as

"J. A. Carrillo, A. Esposito, C. Falcó & A. Fernández-Jiménez *Competing effects in fourth-order aggregation–diffusion equations.* Proceedings of the London Mathematical Society, 129:e12623, 2024" [50].

7.2.3 Travelling waves in heterogeneous populations: go-or-grow

The *go-or-grow* hypothesis states that adherent cells undergo reversible phenotypic switching between a migratory (*go*) state and a proliferative (*grow*) state, with migration capacity maximised when proliferative processes are downregulated. Motivated by our travelling wave analysis of heterogeneous populations in Chapter 3 and by experimental investigations of this dichotomy in cell invasion [117, 193, 214], we decided to study invasion dynamics in the simplest possible go–or–grow reaction–diffusion model. We formulated a pair of coupled equations for the densities of migratory and a proliferative population, with linear diffusion acting on the first, logistic growth on the second, and general density-dependent switching rates between the two states.

By taking a fast-switching limit (i.e. *fast-reaction*), we showed that the total density satisfies a single reaction–diffusion equation with both a density-dependent diffusion coefficient and a density-dependent proliferation rate. This reduction reveals an interesting connection between multi-phenotype go-or-grow systems and single-equation models featuring degenerate diffusion.

Focusing on one-dimensional travelling waves, we derived a dispersion relation for the wave speed and proved that this speed never exceeds the classical Fisher–KPP speed. In the biologically relevant case where the switching rate into proliferation vanishes at zero density, nonlinear wave speed selection pushes the actual wave speed strictly above the linear prediction, and explicit formulas for the minimal speed cannot be obtained via the standard linear stability analysis. These results explain previously observed discrepancies between analytic estimates and numerical simulations in similar reaction–diffusion systems [68, 70]. In future work, we plan to address the problem of nonlinear wave speed selection in systems of reaction–diffusion equations.

This work is published as

"C. Falcó, R. M. Crossley & R. E. Baker. *Travelling waves in a minimal go-or- grow model of cell invasion.* Applied Mathematics Letters, 158:109209, 2024" [92].

7.3 Perspectives and future work

Throughout this thesis, we have uncovered more questions than we have been able to answer. Below, we summarise some of these ideas and outline future perspectives for mathematical, and more broadly quantitative biology.

7.3.1 Consistent continuum theory of cellular adhesion

The local thin-film–type adhesion model presented in Chapter 5 is obtained by taking a formal short-range limit of nonlocal aggregation–diffusion equations, yet those nonlocal models themselves rest on mean-field approximations for long-range interactions. This two-step approach leaves open the question of whether the intermediate nonlocal description is necessary at all, since cells interact through finite-range adhesive and repulsive forces.

A more direct route would begin with an agent-based system of (potentially) hard spheres endowed with finite-range attractive potentials, and then perform a continuum limit as both particle volume and interaction range vanish [33, 34]. It would be interesting if such an analysis could directly yield a local fourth-order aggregation–diffusion equation of the same general form as our thin-film model. Clarifying this connection would help reveal how specific microscopic adhesion rules give rise to the macroscopic equations used to describe tissue-scale patterning.

7.3.2 Mechanics, cell polarity, and interactions

Adhesive interactions alone are not sufficient to reproduce many of the collective behaviours observed in migrating cell populations. The models considered in this thesis — whether based on an effective diffusion law or on simple attraction–repulsion kernels — neglect the intrinsic polarity of individual cells and the active forces they generate [2]. In experimental systems, such as the vortex-forming assays of *Heinrich et al.* [112], the emergence of coherent rotational flows cannot be captured by purely diffusive or adhesion-driven models. Instead, these dynamics require velocity-alignment mechanisms and active terms that allow cells to expend metabolic energy on directed motion, rather than relying solely on diffusive fluctuations.

Accordingly, the diffusion-based models developed here should be understood as an effective macroscopic description valid on long timescales and in regimes where polarity and activity average out. To move beyond this limitation, it would be highly beneficial to develop continuum models that couple density-dependent adhesion with explicit polarity fields and active force generation [32]. Such models could, for example, track a polarity order parameter alongside cell density and include polar advection, alignment interactions, and active stress contributions. By deriving these terms directly from underlying cell motility rules — rather than prescribing them phenomenologically — one would obtain a more mechanistic description than is typical in the active fluids literature, clarifying how microscopic polarity and force generation give rise to tissue-scale flows [2, 112].

By incorporating these mechanisms, one could gain analytical insight into phenomena like those reported by *Kim et al.* [128], where the interplay of adhesion and active

forces gives rise to complex spatiotemporal patterns. A unified continuum framework for adhesion, polarity, and active stresses would therefore provide a powerful tool for interpreting, and ultimately predicting, the rich collective behaviours observed in migrating tissues.

7.3.3 Cell–substrate and environmental interactions

Accounting for cell–substrate interactions becomes crucial in environments with spatial heterogeneities. For example, in the spheroid assays of *Kim et al.* (see Chapter 6), curvature itself creates differential cues that influence both adhesion and collective migration [24, 69, 128]. More generally, understanding how surface curvature couples to cell–cell and cell–substrate forces, particularly in mixtures of distinct cell types, could shed light on morphogenetic processes such as neural-crest migration, where tissue geometry and local adhesion heterogeneity guide cohesive streaks of migrating cells [158].

At the same time, there is growing interest in actively steering cell migration to construct tissue-like materials or to promote regenerative growth. External stimuli such as electric fields [67, 148], optogenetic activation [180, 201], or substrates with spatially varying adhesion ligands [133] have been shown to guide cell polarity, giving rise to directed motion, and patterns such as lanes and more elaborate self-organised architectures. Developing continuum models that couple density-dependent cellular and mechanical interactions with explicit cell–substrate couplings would provide a mechanistic framework for these phenomena. Such theories could predict how substrate curvature, and external fields combine to give rise to long-term self-organisation of cells. Ultimately, these models would establish design principles for tissue engineering and regenerative medicine, linking material patterning and design to desired multicellular outcomes.

7.3.4 Learning models from data

As we have seen, deriving mechanistic continuum models from agent-based rules can be challenging and become technically demanding. Once we have candidate, classical parameter estimation techniques allow us to calibrate unknown coefficients against

experimental measurements. However, the unprecedented volume and resolution of modern experimental data open the door to learning models (or parts of them) directly from observations [14].

One promising class of methods is based on *sparse regression* [36, 181, 159], in which time derivatives and spatial gradients are estimated from data and then matched against a library of candidate terms to identify the best fitting differential equation. Alternatively, neural network based techniques can be used to infer models: *universal differential equations* embed a trainable network inside a differential equation model [171, 174], while *physics-informed neural networks* (PINNs) impose known model structure as a soft constraint on a network that learns solution outcomes [135, 175].

All of the continuum models developed in this thesis are amenable to these data-driven frameworks [150, 159, 160]. A systematic investigation of how different methods perform on our models would not only allow robust parameter estimation, but could also reveal missing mechanisms or suggest entirely new model terms. In the coming years, we anticipate that combining first-principles model derivations with data-driven discovery will become the standard workflow in mathematical biology, yielding hybrid models that leverage both biological knowledge and the full richness of experimental datasets.

Appendix A

Finite volume numerical schemes

This appendix summarises the finite volume methods used throughout this thesis to solve the different continuum models. We present a unified framework for the two models solved on a Cartesian grid — nonlinear diffusion with proliferation (Chapters 2 and 4, and the local cell–cell adhesion model (Chapter 5) — highlighting their common structure and differences, and then describe the scheme in polar coordinates applied to the cell cycle model (Chapter 3). All schemes are based on the finite volume schemes developed in [11, 12, 45].

A.1 Finite volume schemes in Cartesian coordinates

We present the finite volume schemes in one spatial dimension for simplicity. The extension to two spatial dimensions is straightforward. All simulations approximate the cell density $\rho(x,t)$ on a uniform grid of discrete volumes (cells) and evolve the average densities via ordinary differential equations obtained from integrating the governing equation over each volume. Denote the discrete cells by $C_i = [x_{i-1/2}, x_{i+1/2}]$ of width δx, centered at $x_i = i\delta x$, and let

$$\rho_i(t) \;=\; \frac{1}{\delta x}\int_{C_i} \rho(x,t)\,\mathrm{d}x.$$

Integrating

$$\partial_t \rho = -\partial_x\big(\rho\, v\big) \;+\; (\text{source terms})\,,$$

over C_i and applying the divergence theorem gives

$$\frac{\mathrm{d}\rho_i}{\mathrm{d}t} = -\frac{F_{i+1/2} - F_{i-1/2}}{\delta x} \;+\; (\text{source})\,,$$

where $F_{i\pm1/2}$ approximates the flux $\rho\, v$ at $x_{i\pm1/2}$. In each model we compute

$$v^+_{i+1/2} = \max(v_{i+1/2}, 0), \quad v^-_{i+1/2} = \min(v_{i+1/2}, 0), \quad F_{i+1/2} = v^+_{i+1/2}\,\rho_i + v^-_{i+1/2}\,\rho_{i+1}.$$

Time stepping is carried out explicitly with an explicit fourth–order Runge–Kutta method (via `scipy.integrate.ode`).

A.1.1 Nonlinear diffusion with proliferation

For Chapters 2 and 4, we study nonlinear reaction-diffusion equations,

$$\partial_t \rho = \partial_x \Big(\rho\, \partial_x P(\rho) \Big) + r\,\rho \Big(1 - \frac{\rho}{K} \Big),$$

with no–flux boundary conditions. Here $P(\rho) \sim \rho$ is the pressure. Neglecting the growth term for flux discretisation, define

$$v_{i+1/2} = -\frac{P(\rho_{i+1}) - P(\rho_i)}{\delta x}.$$

Then the numerical fluxes are

$$J_{i+1/2} = (v_{i+1/2})^+\,\rho_i + (v_{i+1/2})^-\,\rho_{i+1},$$

and the logistic growth term $r\,\rho_i(1-\rho_i/K)$ is added explicitly to the differential equation right–hand side. No–flux boundaries are enforced by setting $J_{1/2} = J_{N+1/2} = 0$.

A.1.2 Local cell–cell adhesion model

In Chapter 5 we consider the Cahn–Hilliard–type adhesion model

$$\partial_t \rho = -\partial_x \Big[\rho\, \partial_x (\partial_x^2 \rho + \mu^2 \rho) \Big],$$

with periodic boundary conditions. We use the standard finite difference approximation of the Laplacian with

$$(\Delta\rho)_i = \frac{\rho_{i+1} - 2\rho_i + \rho_{i-1}}{\delta x^2}, \quad \xi_i = (\Delta\rho)_i + \mu^2 \rho_i, \quad v_{i+1/2} = -\frac{\xi_{i+1} - \xi_i}{\delta x}.$$

The upwind flux is

$$F_{i+1/2} = (v_{i+1/2})^+\,\rho_i + (v_{i+1/2})^-\,\rho_{i+1}.$$

This scheme preserves energy decay at the semidiscrete (spatially discrete) level [11].

In both models, uniform mesh sizes δx and explicit fourth-order Runge–Kutta timesteps δt are chosen to satisfy stability constraints.

A.2 Finite volume scheme in polar coordinates for the cell cycle model

For radial tissue expansion, we assume radial symmetry and rewrite the Laplacian operator in polar form:

$$\Delta\rho = \partial_r^2\rho + \frac{1}{r}\,\partial_r\rho.$$

The domain is partitioned into a central disk C_0 of radius $r_{1/2} = \delta r/2$ and concentric annuli C_i of width δr, for $i = 1, \ldots, N$. Denoting the average densities ρ_k^i in each volume, integration yields

$$\frac{\mathrm{d}\rho_k^i}{\mathrm{d}t} = \frac{2\pi D}{|C_i|}\left[r\,\partial_r\rho_k\right]_{r_{i-1/2}}^{r_{i+1/2}} + \text{reaction terms},$$

where $|C_0| = \pi r_{1/2}^2$ and $|C_i| = \pi(r_{i+1/2}^2 - r_{i-1/2}^2)$. Radial derivatives at annuli interfaces are approximated by

$$\left(\partial_r\rho_k\right)_{r_{i+1/2}} \approx \frac{\rho_k^{i+1} - \rho_k^i}{\delta r}.$$

Reaction terms coupling the two cell cycle compartments are incorporated explicitly. The resulting differential equation system is again solved with a fourth–order Runge–Kutta integrator.

References

[1] G. Albi, D. Balague, J. A. Carrillo, and J. von Brecht. Stability analysis of flock and mill rings for second order models in swarming. *SIAM Journal on Applied Mathematics*, 74(3):794, 2014.

[2] R. Alert and X. Trepat. Physical models of collective cell migration. *Annual Review of Condensed Matter Physics*, 11:77–101, 2020.

[3] S. Alt, P. Ganguly, and G. Salbreux. Vertex models: From cell mechanics to tissue morphogenesis. *Philosophical Transactions of the Royal Society B*, 372:20150520, 2017.

[4] J. D. Amack and M. L. Manning. Knowing the boundaries: extending the differential adhesion hypothesis in embryonic cell sorting. *Science*, 338(6104):212–215, 2012.

[5] K. Anguige and C. Schmeiser. A one-dimensional model of cell diffusion and aggregation, incorporating volume filling and cell-to-cell adhesion. *Journal of Mathematical Biology*, 58:395–427, 2008.

[6] N. J. Armstrong, K. J. Painter, and J. A. Sherratt. A continuum approach to modelling cell–cell adhesion. *Journal of Theoretical Biology*, 243(1):98–113, 2006.

[7] P. B. Armstrong. Light and electron microscope studies of cell sorting in combinations of chick embryo neural retina and retinal pigment epithelium. *W. Roux's Archiv für Entwicklungsmechanik*, 168:125–141, 1971.

[8] M. Asllani, T. Carletti, F. Di Patti, D. Fanelli, and F. Piazza. Hopping in the crowd to unveil network topology. *Physical Review Letteres*, 120:158301, 2018.

[9] R. Bailo, J. A. Carrillo, and P. Degond. Pedestrian models based on rational behaviour. In *Crowd Dynamics, Volume 1: Theory, Models, and Safety Problems*, pages 259–292, 2018.

[10] R. Bailo, J. A. Carrillo, and D. Gómez-Castro. Aggregation-diffusion equations for collective behavior in the sciences. *SIAM News*, 2024.

[11] R. Bailo, J. A. Carrillo, S. Kalliadasis, and S. P. Pérez. Unconditional bound-preserving and energy-dissipating finite-volume schemes for the Cahn–Hilliard equation. *Communications in Computational Physics*, 34(3), 2023.

[12] R. Bailo, J. A. Carrillo, H. Murakawa, and M. Schmidtchen. Convergence of a fully discrete and energy-dissipating finite-volume scheme for aggregation-diffusion equations. *Mathematical Models and Methods in Applied Sciences*, 30(13):2487–2522, 2020.

[13] B. T. Bajar, A. J. Lam, R. K. Badiee, Y.-H. Oh, J. Chu, X. X. Zhou, N. Kim, B. B. Kim, M. Chung, A. L. Yablonovitch, et al. Fluorescent indicators for simultaneous reporting of all four cell cycle phases. *Nature Methods*, 13(12):993–996, 2016.

[14] 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.

[15] R. E. Baker and M. J. Simpson. Models of collective cell motion for cell populations with different aspect ratio: diffusion, proliferation and travelling waves. *Physica A: Statistical Mechanics and its Applications*, 391(14):3729–3750, 2012.

[16] J. W. Barrett, J. F. Blowey, and H. Garcke. On fully practical finite element approximations of degenerate Cahn–Hilliard systems. *ESAIM: Mathematical Modelling and Numerical Analysis*, 35(4):713–748, 2001.

[17] A. Bernoff and C. Topaz. Biological aggregation driven by social and environmental factors: a nonlocal nodel and its degenerate Cahn–Hilliard approximation. *SIAM Journal on Applied Dynamical Systems*, 15, 2015.

[18] A. L. Bertozzi, T. Kolokolnikov, H. Sun, D. Uminsky, and J. v. Brecht. Ring patterns and their bifurcations in a nonlocal model of biological swarms. *Communications in Mathematical Sciences*, 13(4):955–985, 2015.

[19] A. L. Bertozzi and M. Pugh. The lubrication approximation for thin viscous films: regularity and long-time behavior of weak solutions. *Communications on Pure and Applied Mathematics*, 49(2):85–123, 1996.

[20] A. L. Bertozzi and M. C. Pugh. Long-wave instabilities and saturation in thin film equations. *Communications on Pure and Applied Mathematics*, 51(6):625–661, 1998.

[21] M. Bertsch, R. Dal Passo, and M. Mimura. A free boundary problem arising in a simplified tumour growth model of contact inhibition. *Interfaces and Free Boundaries*, 12(2):235–250, 2010.

[22] C. Blanch-Mercader, R. Vincent, E. Bazellières, X. Serra-Picamal, X. Trepat, and J. Casademunt. Effective viscosity and dynamics of spreading epithelia: a solvable model. *Soft Matter*, 13(6):1235–1243, 2017.

[23] A. Blanchet, J. A. Carrillo, and P. Laurençot. Critical mass for a Patlak–Keller–Segel model with degenerate diffusion in higher dimensions. *Calculus of Variations and Partial Differential Equations*, 35(2):133–168, 2009.

[24] T. Brandstätter, D. B. Brückner, Y. L. Han, R. Alert, M. Guo, and C. P. Broedersz. Curvature induces active velocity waves in rotating spherical tissues. *Nature Communications*, 14(1):1643, 2023.

[25] S. E. Brantley and S. Di Talia. Cell cycle control during early embryogenesis. *Development*, 148(13):dev193128, 2021.

[26] G. Brodland. The Differential Interfacial Tension Hypothesis (DITH): A comprehensive theory for the self-rearrangement of embryonic cells and tissues. *Journal of Biomechanical Engineering*, 124:188–97, 2002.

[27] 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.

[28] A. P. Browning, O. J. Maclaren, P. R. Buenzli, M. Lanaro, M. C. Allenby, M. A. Woodruff, and M. J. Simpson. Model-based data analysis of tissue growth in thin 3D printed scaffolds. *Journal of Theoretical Biology*, 528:110852, 2021.

[29] A. P. Browning, M. Taşcă, C. Falcó, and R. E. Baker. Structural identifiability analysis of linear reaction–advection–diffusion processes in mathematical biology. *Proceedings of the Royal Society A*, 480(2286):20230911, 2024.

[30] D. Brückner and C. P. Broedersz. Learning dynamical models of single and collective cell migration: a review. *Reports on Progress in Physics*, 2024.

[31] D. B. Brückner, P. Ronceray, and C. P. Broedersz. Inferring the dynamics of underdamped stochastic systems. *Physical Review Letters*, 125(5):058103, 2020.

[32] M. Bruna, M. Burger, A. Esposito, and S. M. Schulz. Phase separation in systems of interacting active Brownian particles. *SIAM Journal on Applied Mathematics*, 82(4):1635–1660, 2022.

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

[34] M. Bruna, S. J. Chapman, and M. Robinson. Diffusion of particles with short-range interactions. *SIAM Journal on Applied Mathematics*, 77(6):2294–2316, 2017.

[35] M. Bruna, S. J. Chapman, and M. Schmidtchen. Derivation of a macroscopic model for Brownian hard needles. *Proceedings of the Royal Society A*, 479(2274):20230076, 2023.

[36] S. L. Brunton, J. L. Proctor, and J. N. Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings on the National Academy of Sciences*, 113(15):3932–3937, 2016.

[37] M. Burger and A. Esposito. Porous medium equation and cross-diffusion systems as limit of nonlocal interaction. *Nonlinear Analysis*, 235:113347, 2023.

[38] M. Burger, M. Francesco, and M. Franek. Stationary states of quadratic diffusion equations with long-range attraction. *Communications in Mathematical Sciences*, 11:709–738, 2011.

[39] A. Buttenschön, S. Sinclair, and L. Edelstein-Keshet. How cells stay together: a mechanism for maintenance of a robust cluster explored by local and non-local continuum models. *Bulletin of Mathematical Biology*, 86(11):129, 2024.

[40] G. Cai, X. Li, S. Lin, S. Chen, N. Rodgers, K. Koning, D. Bi, and A. Liu. Matrix confinement modulates 3D spheroid sorting and burst-like collective migration. *Acta Biomater*, 179:192–206, 2024.

[41] V. Calvez and J. A. Carrillo. Volume effects in the Keller–Segel model: energy estimates preventing blow-up. *Journal de Mathématiques Pures et Appliquées*, 86(2):155–175, 2006.

[42] J. Canosa. On a nonlinear diffusion equation describing population growth. *IBM Journal of Research and Development*, 17(4):307–313, 1973.

[43] T. Carletti, M. Asllani, D. Fanelli, and V. Latora. Nonlinear walkers and efficient exploration of congested networks. *Physical Review Research*, 2:033012, 2020.

[44] L. C. Carpenter, F. Pérez-Verdugo, and S. Banerjee. Mechanical control of cell proliferation patterns in growing epithelial monolayers. *Biophysical Journal*, 123(7):909–919, 2024.

[45] J. A. Carrillo, A. Chertock, and Y. Huang. A finite-volume method for nonlinear nonlocal equations with a gradient flow structure. *Communications in Computational Physics*, 17(1):233–258, 2014.

[46] J. A. Carrillo, A. Colombi, and M. Scianna. Adhesion and volume constraints via nonlocal interactions determine cell organisation and migration profiles. *Journal of Theoretical Biology*, 445:75–91, 2018.

[47] J. A. Carrillo, K. Craig, and Y. Yao. Aggregation-diffusion equations: dynamics, asymptotics, and singular limits. In *Active particles. Vol. 2. Advances in theory, models, and applications*, Model. Simul. Sci. Eng. Technol., pages 65–108. Birkhäuser/Springer, Cham, 2019.

[48] J. A. Carrillo, M. R. D'Orsogna, and V. Panferov. Double milling in self-propelled swarms from kinetic theory. *Kinetic and Related Models*, 2(2):363–378, 2009.

[49] J. A. Carrillo, C. Elbar, and J. Skrzeczkowski. Degenerate Cahn–Hilliard systems: From nonlocal to local. *Communications in Contemporary Mathematics*, 27(06):2450041, 2025.

[50] J. A. Carrillo, A. Esposito, C. Falcó, and A. Fernández-Jiménez. Competing effects in fourth-order aggregation-diffusion equations. *arXiv preprint arXiv:2307.14706*, 2023.

[51] J. A. Carrillo, A. Esposito, C. Falcó, and A. Fernández-Jiménez. Competing effects in fourth-order aggregation–diffusion equations. *Proceedings of the London Mathematical Society*, 129:e12623, 2024.

[52] J. A. Carrillo, S. Fagioli, F. Santambrogio, and M. Schmidtchen. Splitting schemes and segregation in reaction cross-diffusion systems. *SIAM Journal on Mathematical Analysis*, 50(5):5695–5718, 2018.

[53] J. A. Carrillo, Y. Huang, and M. Schmidtchen. Zoology of a nonlocal cross-diffusion model for two species. *SIAM Journal on Applied Mathematics*, 78(2):1078–1104, 2018.

[54] J. A. Carrillo and K. Lin. Sharp conditions on global existence and blow-up in a degenerate two-species and cross-attraction system. *Adv. Nonlinear Anal.*, 11(1):1–39, 2022.

[55] J. A. Carrillo, R. J. McCann, and C. Villani. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Revista Matematica Iberoamericana*, 19(3):971–1018, 2003.

[56] J. A. Carrillo, H. Murakawa, M. Sato, H. Togashi, and O. Trush. A population dynamics model of cell-cell adhesion incorporating population pressure and density saturation. *Journal of Theoretical Biology*, 474:14–24, 2019.

[57] J. A. Carrillo and G. Toscani. Asymptotic L^1-decay of solutions of the porous medium equation to self-similarity. *Indiana University Mathematics Journal*, 49(1):113–142, 2000.

[58] G. L. Celora, S. B. Bader, E. M. Hammond, P. K. Maini, J. M. Pitt-Francis, and H. M. Byrne. A DNA-structured mathematical model of cell-cycle progression in cyclic hypoxia. *Journal of Theoretical Biology*, 545:111104, 2022.

[59] G. L. Celora, M. G. Hennessy, A. Muench, B. Wagner, and S. Waters. The dynamics of a collapsing polyelectrolyte gel. *SIAM Journal on Applied Mathematics*, 83(3):1146–1171, 2023.

[60] A. Cerchiari, J. Garbe, N. Jee, M. Todhunter, K. Broaders, D. Peehl, T. Desai, M. LaBarge, M. Thomson, and Z. Gartner. A strategy for tissue self-organization that is robust to cellular heterogeneity and plasticity. *Proceedings of the National Academy of Sciences*, 112(7):2287–2292, 2015.

[61] M. Cetera, G. Ramirez-San Juan, P. Oakes, L. Lewellyn, M. Fairchild, G. Tanentzapf, M. Gardel, and S. Horne-Badovinac. Epithelial rotation promotes the global alignment of contractile actin bundles during Drosophila egg chamber elongation. *Nature Communications*, 5:5511, 2014.

[62] L. Chen, K. Painter, C. Surulescu, and A. Zhigun. Mathematical models for cell migration: a non-local perspective. *Philosophical Transactions of the Royal Society B*, 375(1807):20190379, 2020.

[63] A. Chin, K. Worley, P. Ray, G. Kaur, J. Fan, and L. Wan. Epithelial cell chirality revealed by three-dimensional spontaneous rotation. *Proceedings of the National Academy of Sciences*, 115(48):12188–12193, 2018.

[64] O.-T. Chis, J. R. Banga, and E. Balsa-Canto. Structural identifiability of systems biology models: a critical comparison of methods. *PLoS One*, 6(11):e27755, 2011.

[65] M. Chuai, D. Hughes, and C. J Weijer. Collective epithelial and mesenchymal cell migration during gastrulation. *Current Genomics*, 13(4):267–277, 2012.

[66] C. Cobelli and J. J. Distefano III. Parameter and structural identifiability concepts and ambiguities: a critical review and analysis. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 239(1):R7–R24, 1980.

[67] D. J. Cohen, W. James Nelson, and M. M. Maharbiz. Galvanotactic control of collective cell migration in epithelial monolayers. *Nature Materials*, 13(4):409–417, 2014.

[68] 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 A*, 477(2256):20210593, 2021.

[69] 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.

[70] 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.

[71] M. G. Delgadino. Convergence of a one-dimensional Cahn–Hilliard equation with degenerate mobility. *SIAM Journal on Mathematical Analysis*, 50(4):4457–4482, 2018.

[72] M. Di Francesco, A. Esposito, and S. Fagioli. Nonlinear degenerate cross-diffusion systems with nonlocal interaction. *Nonlinear Analysis*, 169:94–117, 2018.

[73] M. Di Francesco and S. Fagioli. Measure solutions for non-local interaction PDEs with two species. *Nonlinearity*, 26(10):2777, 2013.

[74] P. Domschke, D. Trucu, A. Gerisch, and M. A. J. Chaplain. Mathematical modelling of cancer invasion: Implications of cell adhesion variability for tumour infiltrative growth patterns. *Journal of Theoretical Biology*, 361:41–60, 2014.

[75] L. Donker, R. Houtekamer, M. Vliem, F. Sipieter, H. Canever, M. Gómez-González, M. Bosch-Padrós, W.-J. Pannekoek, X. Trepat, N. Borghi, and M. Gloerich. A mechanical G2 checkpoint controls epithelial cell division through E-cadherin-mediated regulation of Wee1-Cdk1. *Cell Reports*, 41(2):111475, 2022.

[76] D. Duguay, R. A. Foty, and M. S. Steinberg. Cadherin-mediated cell adhesion and tissue segregation: qualitative and quantitative determinants. *Developmental Biology*, 253(2):309–323, 2003.

[77] L. Dyson and R. E. Baker. The importance of volume exclusion in modelling cellular migration. *Journal of Mathematical Biology*, 71:691–711, 2014.

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

[79] M. R. D'Orsogna, Y.-L. Chuang, A. L. Bertozzi, and L. S. Chayes. Self-propelled particles with soft-core interactions: patterns, stability, and collapse. *Physical Review Letters*, 96(10):104302, 2006.

[80] 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.

[81] M. El-Hachem, S. W. McCue, and M. J. Simpson. A continuum mathematical model of substrate-mediated tissue growth. *Bulletin of Mathematical Biology*, 84(4):1–27, 2022.

[82] C. Elbar and J. Skrzeczkowski. Degenerate Cahn–Hilliard equation: From nonlocal to local. *Journal of Differential Equations*, 364:576–611, 2023.

[83] E. Ellefsen and N. Rodriguez. On equilibrium solutions to nonlocal mechanistic models in ecology. *Journal of Applied Analysis & Computation*, 11(6):2664–2686, 2021.

[84] C. M. Elliott. The Cahn–Hilliard model for the kinetics of phase separation. In *Mathematical Models for Phase Change Problems*, pages 35–73. Springer, 1989.

[85] C. M. Elliott and H. Garcke. On the Cahn–Hilliard equation with degenerate mobility. *SIAM Journal on Mathematical Analysis*, 27(2):404–423, 1996.

[86] C. Falcó. From random walks on networks to nonlinear diffusion. *Physical Review E*, 106(5):054103, 2022.

[87] C. Falcó, R. E. Baker, and J. A. Carrillo. A local continuum model of cell–cell adhesion. *SIAM Journal on Applied Mathematics*, 84(3):S17–S42, 2024.

[88] C. Falcó, R. E. Baker, and J. A. Carrillo. A nonlocal-to-local approach to aggregation-diffusion equations. *SIAM Review*, 67(2):353–372, 2025.

[89] C. Falcó, D. J. Cohen, J. A. Carrillo, and R. E. Baker. Quantifying tissue growth, shape and collision via continuum models and Bayesian inference. *Journal of the Royal Society Interface*, 20(204):20230184, 2023.

[90] C. Falcó, D. J. Cohen, J. A. Carrillo, and R. E. Baker. Quantifying cell cycle regulation by tissue crowding. *Biophysical Journal*, 124(6):923–932, 2025.

[91] C. Falcó, R. E. Baker, and J. A. Carrillo. A local continuum model of cell-cell adhesion: figshare media, 2022. `https://figshare.com/projects/A_local_continuum_model_of_cell-cell_adhesion/142427`.

[92] 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.

[93] T. Fessenden, Y. Beckham, M. Perez-Neut, G. Ramirez-San Juan, A. Chourasia, K. Macleod, P. Oakes, and M. Gardel. Dia1-dependent adhesions are required by epithelial tissues to initiate invasion. *Journal of Cell Biology*, 217(4):1485–1502, 2018.

[94] F. Flandoli and M. Leimbach. Mean field limit with proliferation. *Discrete and Continuous Dynamical Systems - B*, 21(9):3029–3052, 2016.

[95] R. A. Foty and M. S. Steinberg. Cadherin-mediated cell-cell adhesion and tissue segregation in relation to malignancy. *International Journal of Developmental Biology*, 48(5-6):397–409, 2004.

[96] R. A. Foty and M. S. Steinberg. The differential adhesion hypothesis: a direct evaluation. *Developmental Biology*, 278(1):255–263, 2005.

[97] M. Francesco, A. Esposito, and S. Fagioli. Nonlinear degenerate cross-diffusion systems with nonlocal interaction. *Nonlinear Analysis*, 169, 2017.

[98] A. Gamba, D. Ambrosi, A. Coniglio, A. de Candia, S. Di Talia, E. Giraudo, G. Serini, L. Preziosi, and F. Bussolino. Percolation, morphogenesis, and Burgers dynamics in blood vessels formation. *Physical Review Letters*, 90(11):118101, 2003.

[99] E. Gauquelin, S. Tlili, C. Gay, G. Peyret, R.-M. Mège, M.-A. Fardin, and B. Ladoux. Influence of proliferation on the motions of epithelial monolayers invading adherent strips. *Soft Matter*, 15(13):2798–2810, 2019.

[100] E. Gavagnin, M. J. Ford, R. L. Mort, T. Rogers, and C. A. Yates. The invasion speed of cell migration models with realistic cell cycle time distributions. *Journal of Theoretical Biology*, 481:91–99, 2019.

[101] A. Gerisch and M. 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.

[102] B. G. Godard and C.-P. Heisenberg. Cell division and tissue mechanics. *Current Opinion in Cell Biology*, 60:114–120, 2019.

[103] R. E. Goldstein, A. I. Pesci, and M. J. Shelley. Topology transitions and singularities in viscous flows. *Physical Review Letters*, 70:3043–3046, 1993.

[104] R. E. Goldstein, A. I. Pesci, and M. J. Shelley. Instabilities and singularities in Hele–Shaw flow. *Physics of Fluids*, 10(11):2701–2723, 1998.

[105] D. Gómez-Castro. Beginner's guide to aggregation-diffusion equations. *SeMA Journal*, 81:531–587, 2024.

[106] N. Grodzinski, R. L. Jack, and M. E. Cates. Hydrodynamic theory of wetting by active particles. *arXiv preprint arXiv:2506.14559*, 2025.

[107] V. K. Gupta and O. Chaudhuri. Mechanical regulation of cell-cycle progression and division. *Trends in Cell Biology*, 32(9):773–785, 2022.

[108] M. E. Gurtin and R. C. MacCamy. On the diffusion of biological populations. *Mathematical Biosciences*, 33(1-2):35–49, 1977.

[109] E. J. Hackett-Jones, K. A. Landman, D. F. Newgreen, and D. Zhang. On the role of differential adhesion in gangliogenesis in the enteric nervous system. *Journal of Theoretical Biology*, 287:148–159, 2011.

[110] S. L. Haigo and D. Bilder. Global tissue revolutions in a morphogenetic movement controlling elongation. *Science*, 331(6020):1071–1074, 2011.

[111] A. Hashimoto, A. Nagao, and S. Okuda. Topological graph description of multicellular dynamics based on vertex model. *Journal of Theoretical Biology*, 437, 2017.

[112] M. A. Heinrich, R. Alert, J. M. LaChance, T. J. Zajdel, A. Košmrlj, and D. J. Cohen. Size-dependent patterns of cell proliferation and migration in freely-expanding epithelia. *eLife*, 9:e58945, 2020.

[113] M. A. Heinrich, R. Alert, J. M. LaChance, T. J. Zajdel, A. Kosmrlj, and D. J. Cohen. Size-dependent patterns of cell proliferation and migration in freely-expanding epithelia (Version 1) [Data set]. Zenodo, 2020. `https://doi.org/10.5281/zenodo.3858845`.

[114] M. A. Heinrich, R. Alert, A. E. Wolf, A. Košmrlj, and D. J. Cohen. Self-assembly of tessellated tissue sheets by expansion and collision. *Nature Communications*, 13(1):1–10, 2022.

[115] K. E. Hines, T. R. Middendorf, and R. W. Aldrich. Determination of parameter identifiability in nonlinear biophysical models: A Bayesian approach. *The Journal of General Physiology*, 143(3):401–416, 2014.

[116] T. Hirashima, E. G. Rens, and R. M. Merks. Cellular Potts modeling of complex multicellular behaviors in tissue morphogenesis. *Development, Growth & Differentiation*, 59(5):329–339, 2017.

[117] K. S. Hoek, O. M. Eichhoff, N. C. Schlegel, U. Döbbeling, N. Kobert, L. Schaerer, S. Hemmi, and R. Dummer. In vivo switching of human melanoma cells between proliferative and invasive states. *Cancer Research*, 68(3):650–656, 2008.

[118] K. Höllring, L. Nuić, L. Rogić, S. Kaliman, S. Gehrer, C. Wollnik, F. Rehfeldt, M. Hubert, and A.-S. Smith. Capturing the mechanosensitivity of cell proliferation in models of epithelium. *Proceedings of the National Academy of Sciences*, 121(45):e2308126121, 2024.

[119] J. Holtfreter. Properties and functions of the surface coat in amphibian embryos. *Journal of Experimental Zoology*, 93(2):251–323, 1943.

[120] D. L. Janzén, L. Bergenholm, M. Jirstrand, J. Parkinson, J. Yates, N. D. Evans, and M. J. Chappell. Parameter identifiability of fundamental pharmacodynamic models. *Frontiers in Physiology*, 7:590, 2016.

[121] W. Jin, E. T. Shah, C. J. Penington, S. W. McCue, L. K. Chopin, and M. J. Simpson. Reproducibility of scratch assays is affected by the initial degree of confluence: Experiments, modelling and model selection. *Journal of Theoretical Biology*, 390:136–145, 2016.

[122] W. Jin, E. T. Shah, C. J. Penington, S. W. McCue, P. K. Maini, and M. J. Simpson. Logistic proliferation of cells in scratch assays is delayed. *Bulletin of Mathematical Biology*, 79(5):1028–1050, 2017.

[123] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.

[124] P. Jorgensen and M. Tyers. How cells coordinate growth and division. *Current Biology*, 14(23):R1014–R1027, 2004.

[125] S. Katsunuma, H. Honda, T. Shinoda, Y. Ishimoto, T. Miyata, H. Kiyonari, T. Abe, K. Nibu, Y. Takai, and H. Togashi. Synergistic action of nectins and cadherins generates the mosaic cellular pattern of the olfactory epithelium. *Journal of Cell Biology*, 212(5):561–575, 2016.

[126] E. Khain and J. Straetmans. Dynamics of an expanding cell monolayer. *Journal of Statistical Physics*, 184(2):1–13, 2021.

[127] E. Khain and L. S. Tsimring. Effective pressure and cell area distribution in a confined monolayer. *Fluid Dynamics Research*, 50(5):051413, 2018.

[128] J. Kim, H. Jeong, C. Falcó, A. M. Hruska, W. D. Martinson, A. Marzoratti, M. Araiza, H. Yang, C. Franck, J. A. Carrillo, M. Guo, and I. Y. Wong. Collective transitions from orbiting to matrix invasion in 3D multicellular spheroids. *bioRxiv*, 2025. preprint.

[129] J. J. Klowss, A. P. Browning, R. J. Murphy, E. J. Carr, M. J. Plank, G. Gunasingh, N. K. Haass, and M. J. Simpson. A stochastic mathematical model of 4D tumour spheroids with real-time fluorescent cell cycle labelling. *Journal of the Royal Society Interface*, 19(189):20210903, 2022.

[130] M. Krieg, Y. Arboleda-Estudillo, P.-H. Puech, J. Käfer, F. Graner, D. Müller, and C.-P. Heisenberg. Tensile forces govern germ-layer organization in zebrafish. *Nature Cell Biology*, 10(4):429–436, 2008.

[131] J. C. Kynaston, C. Guiver, and C. A. Yates. Equivalence framework for an age-structured multistage representation of the cell cycle. *Physical Review E*, 105(6):064411, 2022.

[132] J. LaChance and D. J. Cohen. Practical fluorescence reconstruction microscopy for large samples and low-magnification imaging. *PLoS Computational Biology*, 16(12):e1008443, 2020.

[133] M. Lacroix, B. Smeets, C. Blanch-Mercader, S. Bell, C. Giuglaris, H.-Y. Chen, J. Prost, and P. Silberzan. Emergence of bidirectional cell laning from collective contact guidance. *Nature Physics*, 20(8):1324–1331, 2024.

[134] B. Ladoux and R.-M. Mège. Mechanobiology of collective cell behaviours. *Nature Reviews Molecular Cell Biology*, 18(12):743–757, 2017.

[135] 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.

[136] B. Lambert, C. L. Lei, M. Robinson, M. Clerx, R. Creswell, S. Ghosh, S. Tavener, and D. J. Gavaghan. Autocorrelated measurement processes and inference for ordinary differential equation models of biological systems. *Journal of the Royal Society Interface*, 20(199):20220725, 2023.

[137] R. Laugesen and M. Pugh. Linear stability of steady states for thin film and Cahn–Hilliard type equations. *Archive for Rational Mechanics and Analysis*, 154:3–51, 2000.

[138] R. Laugesen and M. Pugh. Properties of steady states for thin film equations. *European Journal of Applied Mathematics*, 11:293 – 351, 2000.

[139] R. Laugesen and M. Pugh. Energy levels of steady states for thin-film-type equations. *Journal of Differential Equations*, 182(2):377–415, 2002.

[140] H. Levine, W.-J. Rappel, and I. Cohen. Self-organization in systems of self-propelled particles. *Physical Review E*, 63:017101, 2000.

[141] M. A. Lewis, S. V. Petrovskii, and J. R. Potts. *The Mathematics Behind Biological Invasions*, volume 44 of *Interdisciplinary Applied Mathematics*. Springer, 2016.

[142] Y. li Chuang, M. R. D'Orsogna, D. Marthaler, A. L. Bertozzi, and L. S. Chayes. State transitions and the continuum limit for a 2D interacting, self-propelled particle system. *Physica D: Nonlinear Phenomena*, 232(1):33–47, 2007.

[143] S. Lisini, D. Matthes, and G. Savaré. Cahn-Hilliard and thin film equations with nonlinear mobility as gradient flows in weighted-Wasserstein metrics. *J. Differential Equations*, 253(2):814–850, 2012.

[144] Y. Liu, K. Suh, P. K. Maini, D. J. Cohen, and R. E. Baker. Parameter identifiability and model selection for partial differential equation models of cell invasion. *Journal of the Royal Society Interface*, 21(212):20230607, 2024.

[145] Y. Liu, D. J. Warne, and M. J. Simpson. Likelihood-based inference, identifiability, and prediction using count data from lattice-based random walk models. *Physical Review E*, 110:044405, 2024.

[146] P. K. Maini, D. 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.

[147] M. L. Manning, R. A. Foty, M. S. Steinberg, and E.-M. Schoetz. Coaction of intercellular adhesion and cortical tension specifies tissue surface tension. *Proceedings of the National Academy of Sciences*, 107(28):12517–12522, 2010.

[148] S. F. Martina-Perez, I. B. Breinyn, D. J. Cohen, and R. E. Baker. Optimal control of collective electrotaxis in epithelial monolayers. *Bulletin of Mathematical Biology*, 86(8):95, 2024.

[149] S. F. Martina Perez, H. Sailem, and R. E. Baker. Efficient Bayesian inference for mechanistic modelling with high-throughput data. *PLoS Computational Biology*, 18(6):e1010191, 2022.

[150] S. F. Martina-Perez, M. J. Simpson, and R. E. Baker. Bayesian uncertainty quantification for data-driven equation learning. *Proceedings of the Royal Society A*, 477(2254):20210426, 2021.

[151] 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.

[152] J. Mason, C. Erignoux, R. L. Jack, and M. Bruna. Exact hydrodynamics and onset of phase separation for an active exclusion process. *Proceedings of the Royal Society A*, 479(2279):20230524, 2023.

[153] J. Massagué. G1 cell-cycle control and cancer. *Nature*, 432(7015):298–306, 2004.

[154] O. M. Matsiaka, C. J. Penington, R. E. Baker, and M. J. Simpson. Discrete and continuum approximations for collective cell migration in a scratch assay with cell size dynamics. *Bulletin of Mathematical Biology*, 80(4):738–757, 2018.

[155] D. Matthes, R. J. McCann, and G. Savaré. A family of nonlinear fourth order equations of gradient flow type. *Communications in Partial Differential Equations*, 34(11):1352–1397, 2009.

[156] A. I. McClatchey and A. S. Yap. Contact inhibition (of proliferation) redux. *Current Opinion in Cell Biology*, 24(5):685–694, 2012.

[157] C. R. McKeown and H. T. Cline. Nutrient restriction causes reversible G2 arrest in Xenopus neural progenitors. *Development*, 146(20):dev178871, 2019.

[158] 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.

[159] D. A. Messenger and D. M. Bortz. Weak SINDy for partial differential equations. *Journal of Computational Physics*, 443:110525, 2021.

[160] D. A. Messenger and D. M. Bortz. Weak SINDy: Galerkin-based data-driven model selection. *Multiscale Modeling & Simulation*, 19(3):1474–1497, 2021.

[161] D. A. Messenger, G. E. Wheeler, X. Liu, and D. M. Bortz. Learning anisotropic interaction rules from individual trajectories in a heterogeneous cellular population. *Journal of The Royal Society Interface*, 19(195):20220412, 2022.

[162] H. Murakawa and H. Togashi. Continuous models for cell–cell adhesion. *Journal of Theoretical Biology*, 374:1–12, 2015.

[163] J. D. Murray. *Mathematical Biology I: An Introduction*. Springer New York, 2001.

[164] T. G. Myers. Thin films with high surface tension. *SIAM Review*, 40(3):441–462, 1998.

[165] C. M. Nowak, T. Quarton, and L. Bleris. Impact of variability in cell cycle periodicity on cell population dynamics. *PLoS Computational Biology*, 19(6):e1011080, 2023.

[166] K. Oelschläger. Large systems of interacting particles and the porous medium equation. *Journal of Differential Equations*, 88:294–346, 1990.

[167] T. Otto and P. Sicinski. Cell cycle proteins as promising targets in cancer therapy. *Nature Reviews Cancer*, 17(2):93–115, 2017.

[168] P. H. O'Farrell. Growing an embryo from a single cell: a hurdle in animal life. *Cold Spring Harbor Perspectives in Biology*, 7(11):a019042, 2015.

[169] K. Painter and T. Hillen. Volume-filling and quorum-sensing in models for chemosensitive movement. *Canadian Applied Mathematics Quarterly*, 10(4):501–544, 2002.

[170] A. B. Pardee. G1 events and regulation of cell proliferation. *Science*, 246(4930):603–608, 1989.

[171] M. Philipps, N. Schmid, and J. Hasenauer. Universal differential equations for systems biology: Current state and open problems. *bioRxiv*, 2024. preprint.

[172] 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.

[173] A. Puliafito, L. Hufnagel, P. Neveu, S. Streichan, A. Sigal, D. K. Fygenson, and B. I. Shraiman. Collective and single cell behavior in epithelial contact inhibition. *Proceedings of the National Academy of Sciences*, 109(3):739–744, 2012.

[174] 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.

[175] 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.

[176] A. Raue, J. Karlsson, M. P. Saccomani, M. Jirstrand, and J. Timmer. Comparison of approaches for parameter identifiability analysis of biological systems. *Bioinformatics*, 30(10):1440–1448, 2014.

[177] M. Renardy, D. Kirschner, and M. Eisenberg. Structural identifiability analysis of age-structured PDE epidemic models. *Journal of Mathematical Biology*, 84(1):1–30, 2022.

[178] D. A. Ridenour, M. C. McKinney, C. M. Bailey, and P. M. Kulesa. Cycletrak: a novel system for the semi-automated analysis of cell cycle dynamics. *Developmental Biology*, 365(1):189–195, 2012.

[179] P. Rørth. Fellow travellers: emergent properties of collective cell migration. *EMBO reports*, 13(11):984–991, 2012.

[180] L. Rossetti, S. Grosser, J. F. Abenza, L. Valon, P. Roca-Cusachs, R. Alert, and X. Trepat. Optogenetic generation of leader cells reveals a force–velocity relation for collective cell migration. *Nature Physics*, 20(10):1659–1669, 2024.

[181] 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.

[182] A. Sakaue-Sawano, H. Kurokawa, T. Morimura, A. Hanyu, H. Hama, H. Osawa, S. Kashiwagi, K. Fukami, T. Miyata, H. Miyoshi, et al. Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. *Cell*, 132(3):487–498, 2008.

[183] F. Santambrogio. Optimal Transport for Applied Mathematicians. *Birkäuser, NY*, 55(58-63):94, 2015.

[184] Y. Schälte, F. Fröhlich, P. J. Jost, J. Vanhoefer, D. Pathirana, P. Stapor, P. Lakrisenko, D. Wang, E. Raimúndez, S. Merkt, L. Schmiester, P. Städter, S. Grein, E. Dudkin, D. Doresic, D. Weindl, and J. Hasenauer. pyPESTO: a modular and scalable tool for parameter estimation for dynamic models. *Bioinformatics*, 39(11):btad711, 2023.

[185] B. G. Sengers, C. P. Please, and R. O. 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.

[186] J. A. Sherratt and J. D. Murray. Models of epidermal wound healing. *Proceedings of the Royal Society B*, 241(1300):29–36, 1990.

[187] I. Siekmann, J. Sneyd, and E. J. Crampin. MCMC can detect nonidentifiable models. *Biophysical Journal*, 103(11):2275–2286, 2012.

[188] M. J. Simpson, R. E. Baker, S. T. Vittadello, and O. J. Maclaren. Practical parameter identifiability for spatio-temporal models of cell invasion. *Journal of the Royal Society Interface*, 17(164):20200055, 2020.

[189] 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.

[190] M. J. Simpson, R. J. Murphy, and O. J. Maclaren. Modelling count data with partial differential equation models in biology. *Journal of Theoretical Biology*, 580:111732, 2024.

[191] D. Slepčev. Linear stability of selfsimilar solutions of unstable thin-film equations. *Interfaces and Free Boundaries*, 11(3):375–398, 2009.

[192] J. Smith and L. Martin. Do cells cycle? *Proceedings of the National Academy of Sciences*, 70(4):1263–1267, 1973.

[193] A. M. Stein, T. Demuth, D. Mobley, M. Berens, and L. M. Sander. A mathematical model of glioblastoma tumor spheroid invasion in a three-dimensional in vitro experiment. *Biophysical Journal*, 92(1):356–365, 2007.

[194] M. S. Steinberg. Mechanism of tissue reconstruction by dissociated cells, II: Time-course of events. *Science*, 137(3532):762–763, 1962.

[195] M. S. Steinberg. On the mechanism of tissue reconstruction by dissociated cells, I. Population kinetics, differential adhesiveness, and the absence of directed migration. *Proceedings of the National Academy of Sciences*, 48(9):1577, 1962.

[196] M. S. Steinberg. On the mechanism of tissue reconstruction by dissociated cells, III. Free energy relations and the reorganization of fused, heteronomic tissue fragments. *Proceedings of the National Academy of Sciences*, 48(10):1769, 1962.

[197] M. S. Steinberg. Reconstruction of tissues by dissociated cells: some morphogenetic tissue movements and the sorting out of embryonic cells may have a common explanation. *Science*, 141(3579):401–408, 1963.

[198] M. S. Steinberg. Differential adhesion in morphogenesis: a modern view. *Current Opinion in Genetics & Development*, 17(4):281–286, 2007.

[199] S. J. Streichan, C. R. Hoerner, T. Schneidt, D. Holzer, and L. Hufnagel. Spatial constraints control cell proliferation in tissues. *Proceedings of the National Academy of Sciences*, 111(15):5586–5591, 2014.

[200] K. Suh, Y. K. Cho, I. B. Breinyn, and D. J. Cohen. E-cadherin biointerfaces reprogram collective cell migration and cell cycling by forcing homeostatic conditions. *bioRxiv 2023.07.25.550505*, 2023.

[201] K. Suh, R. H. Thornton, L. Nguyen, P. E. Farahani, D. J. Cohen, and J. E. Toettcher. Large-scale control over collective cell migration using light-activated epidermal growth factor receptors. *Cell Systems*, 16(3), 2025.

[202] K. Tanner, H. Mori, R. Mroue, A. Bruni-Cardoso, and M. Bissell. Coherent angular motion in the establishment of multicellular architecture of glandular tissues. *Proceedings of the National Academy of Sciences*, 109(6):1973–1978, 2012.

[203] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 9.4)*, 2021. `https://www.sagemath.org`.

[204] H. Togashi, K. Kominami, M. Waseda, H. Komura, J. Miyoshi, M. Takeichi, and Y. Takai. Nectins establish a checkerboard-like cellular pattern in the auditory epithelium. *Science*, 333(6046):1144–1147, 2011.

[205] C. Topaz, A. Bertozzi, and M. Lewis. A nonlocal continuum model for biological aggregation. *Bulletin of Mathematical Biology*, 68:1601–1623, 2006.

[206] P. L. Townes and J. Holtfreter. Directed movements and selective adhesion of embryonic amphibian cells. *Journal of Experimental Zoology*, 128(1):53–120, 1955.

[207] O. Trush, C. Liu, X. Han, Y. Nakai, R. Takayama, H. Murakawa, J. A. Carrillo, H. Takechi, S. Hakeda-Suzuki, T. Suzuki, et al. N-cadherin orchestrates self-organization of neurons within a columnar unit in the Drosophila medulla. *Journal of Neuroscience*, 39(30):5861–5880, 2019.

[208] T. Y.-C. Tsai, R. M. Garner, and S. G. Megason. Adhesion-based self-organization in tissue patterning. *Annual Review of Cell and Developmental Biology*, 38(1):349–374, 2022.

[209] F. Turci, R. L. Jack, and N. B. Wilding. Partial and complete wetting of droplets of active brownian particles. *Soft Matter*, 20:2060–2074, 2024.

[210] M. Uroz, S. Wistorf, X. Serra-Picamal, V. Conte, M. Sales-Pardo, P. Roca-Cusachs, R. Guimerà, and X. Trepat. Regulation of cell cycle progression by cell-cell and cell-matrix forces. *Nature Cell Biology*, 20(6):646–654, 2018.

[211] P. O. Valls and A. Esposito. Signalling dynamics, cell decisions, and homeostatic control in health and disease. *Current Opinion in Cell Biology*, 75:102066, 2022.

[212] J. L. Vazquez. *The Porous Medium Equation: Mathematical Theory*. Oxford University Press, 2007.

[213] S. T. Vittadello, S. W. McCue, G. Gunasingh, N. K. Haass, and M. J. Simpson. Mathematical models for cell migration with real-time cell cycle dynamics. *Biophysical Journal*, 114(5):1241–1253, 2018.

[214] 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.

[215] A. Volkening, M. R. Abbott, N. Chandra, B. Dubois, F. Lim, D. Sexton, and B. Sandstede. Modeling stripe formation on growing zebrafish tailfins. *Bulletin of Mathematical Biology*, 82:1–33, 2020.

[216] A. Volkening and B. Sandstede. Modelling stripe formation in zebrafish: an agent-based approach. *Journal of the Royal Society Interface*, 12(112):20150812, 2015.

[217] A. Volkening and B. Sandstede. Iridophores as a source of robustness in zebrafish stripes and variability in Danio patterns. *Nature Communications*, 9(1):3231, 2018.

[218] D. J. Warne, R. E. Baker, and M. J. Simpson. Optimal quantification of contact inhibition in cell populations. *Biophysical Journal*, 113(9):1920–1924, 2017.

[219] T. S. Weber, I. Jaehnert, C. Schichor, M. Or-Guil, and J. Carneiro. Quantifying the length and variance of the eukaryotic cell cycle phases by a stochastic model and dual nucleoside pulse labelling. *PLoS Computational Biology*, 10(7):e1003616, 2014.

[220] J. R. Worsfold, T. Rogers, and P. Milewski. Density fluctuations in stochastic kinematic flows. *SIAM Journal on Applied Mathematics*, 83(3):1000–1024, 2023.

[221] C. A. Yates, M. J. Ford, and R. L. Mort. A multi-stage representation of cell proliferation as a Markov process. *Bulletin of Mathematical Biology*, 79:2905–2928, 2017.

[222] J. Youssef, A. K. Nurse, L. Freund, and J. R. Morgan. Quantification of the forces driving self-assembly of three-dimensional microtissues. *Proceedings of the National Academy of Sciences*, 108(17):6993–6998, 2011.

[223] S. M. Zehnder, M. Suaris, M. M. Bellaire, and T. E. Angelini. Cell volume fluctuations in MDCK monolayers. *Biophysical Journal*, 108(2):247–250, 2015.