Modelling cell competition

Thomas Pak

Linacre College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2021

If the doors of perception were cleansed every thing would appear to man as it is: Infinite.

— William Blake

Acknowledgements

First and foremost, I would like to thank my supervisors, Professor Ruth Baker and Professor Joe Pitt-Francis, from the bottom of my heart for their truly incredible supervision and invaluable advice. They always went the extra mile, *especially* when I needed it the most.

I am also deeply grateful to my examiners, Professor Philip Maini and Doctor Rubén Pérez–Carrasco, for their thoughtful feedback and kind encouragement.

I feel especially blessed for the many incredible people I met during my time in Oxford. Conversing with so many interesting and diverse minds enriched my life in ways I could never have anticipated[1].

Thank you Bailey for being an incredible housemate and friend. I have learned so, so much from you, and I look forward to learning much more.

Thank you Alice for the endless conversation, laughter, and good times we shared through four years of being housemates. Words cannot express how grateful I am for our friendship.

I would also like to thank Eboni, Karel, and Katie for the many wholesome moments we enjoyed as housemates. You truly made Oxford feel like a second home to me.

Thank you David for being a kind-hearted friend and spiritual mentor. Our conversations about everything under the sun have given me a one-of-a-kind education that I will be forever grateful for.

Thank you John for obvious reasons. It is *always* a pleasure.

Thanks to my friends from home for keeping me grounded with their endless zever: King Woetie, Bever, Sander, Dylan, Pieter, and Purple Jacky.

Thanks as well to my parents for their tireless support and encouragement, without which none of this would have been possible.

Last but not least, I extend my deepest thanks to the ultimate Source of all Being. Whoever or whatever is responsible for whatever this is, I am a massive fan of your work. Please keep it going! 🙏🙏🙏

[1] Except Marios. No thanks given whatsoever to Marios.

Abstract

Competition in nature is traditionally conceptualised as a battle between individual organisms in the constant struggle for survival. The triumph of one organism results in the demise of the other. Cooperation on the other hand involves individuals acting in a mutually beneficial manner and is therefore often regarded as the opposite of competition. From the perspective of individual cells, multicellular life is perhaps the most extreme form of cooperation. However, even within multicellular organisms, cells compete for space and survival. Genetically damaged "loser" cells that grow at a slower rate than their neighbouring "winner" cells are eliminated from the body by apoptosis. Importantly, cell competition is context-dependent; loser cells are perfectly viable if the whole organism is composed of them. The demise of loser cells is triggered specifically by the presence of cells that are perceived to be more fit. Multiple triggers and pathways involved with cell competition have been discovered, but the mechanism by which cells measure and communicate their relative fitness remains elusive.

Current models of cell competition assert *a priori* winner or loser status to competing cell types. By their nature, such models cannot explain how the winner or loser identity is attained. In this thesis, we develop a modelling framework to study the emergence of winners and losers in cell competition. Specifically, we construct models of heterotypic populations where cell types can only vary in their model parameters. Using this approach, we show that i) variations in mechanical parameters are not sufficient for cell competition in a vertex-based model, and ii) winner or loser status in a competition system based on the exchange of death signals is determined by the emission rate of death signals *and* the tolerance to death signals. Finally, we make concrete suggestions for the experimental validation of our predictions.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Competition (*n.*)
The action of endeavouring to gain what another endeavours to gain at the same time.

Oxford English Dictionary [1]

Virtually all living creatures participate in some form of competition in their lives. Merely being viable is not enough to be successful in the ruthless struggle for existence. Confronted with scarce resources, one must compete with others to survive, or else be doomed to extinction. Indeed, as Malthus argued in 1798, populations have a tendency to grow exponentially, whereas the renewal of vital resources typically grows linearly at best[1] [2]. The inescapable conclusion is that there must be winners and losers in the game of life; losers fail to compete and are erased from existence, while winners live on to compete another day. This insight famously inspired Darwin to postulate the principle of natural selection [3]. Similarly, the competition for mates in sexually reproducing species is at the basis of sexual selection.

Since all life is subject to natural selection, it is no surprise that competition appears to pervade all of Nature. At the macroscopic scale, animals battle each other for food, territory, and mates [4]. Despite the symbolism of "flower power", plants are constantly competing for nutrients and sunlight, starving their neighbours in the process [5]. At the microscopic scale, bacteria employ a wide variety of tactics to engage in warfare [6, 7], including stabbing, poisoning, and even kamikaze attacks [8]. Competition also extends into

[1] The table of contents, the lists of figures and tables, citations, footnotes, URLs, and references to equations, chapters, sections, pages, figures, and tables in the digital version of this thesis are hyperlinked.

society; humans compete for status and wealth, and organise themselves into groups that compete for economic, political, and military dominion. Indeed, the struggle for existence has at times been described as "Nature's war"; a perpetual state of conflict where all living beings are at war with one another [9].

This, however, is not the full story. Cooperation, typically regarded as the opposite of competition, is also ubiquitous in Nature. Fungi associate symbiotically with plant roots in mycorrhiza, providing the plant with minerals and water in exchange for sugars [10]. Similarly, the human gut harbours bacteria that consume substrates that humans cannot digest directly, and convert them into nutrients that can be absorbed by the host [11]. In return, humans provide a stable habitat and a steady supply of nutrients. There is also evidence for cooperation among members of the same species under certain conditions. In particular, the Allee effect describes the cooperative phenomenon where, for low population densities, the survival rate of individuals increases with population density [12].

Perhaps the most extreme form of cooperation is multicellular life. When viewed as a collective of cellular individuals, the multicellular organism is a remarkable feat of collaboration. In order to avoid a Malthusian catastrophe among its members and maintain homeostasis, cell population dynamics are strongly regulated by cellular society. The multicellular collective does not only constrain proliferation, but also routinely induces programmed cell death, i.e. cell suicide [13, 14, 15]. In fact, this "social control" is so heavily ingrained in metazoan organisms that cells default to suicide unless it is suppressed by survival factors [14]. Indeed, the dysregulation of programmed cell death is one of the hallmarks of cancer [16], which can be regarded as the collapse of cellular society.

However, even within multicellular organisms, a struggle for survival can take place. **Cell competition** is a class of phenomena where cells in genetically heterotypic tissues determine their fitness in relation to each other to determine a winner or loser status [17, 18, 19, 20, 21, 22]. The winners take all and populate the entire tissue, while the losers are actively eliminated through programmed cell death. A defining characteristic of cell competition is that both winner and loser cell types are perfectly viable in homotypic environments, and only acquire a winner/loser status when confronting each other in the same tissue. It is somewhat counter-intuitive that cells belonging to the same organism should compete with each other, since competition and cooperation are often considered to be antithetical. In order to resolve this contrast, we first explore the link between competition and cooperation in more depth.

The field of ecology studies the relationships between organisms and their environment. In the ecological classification of pairwise interactions between species or individuals (see Table 1.1), competition is defined as a mutually harmful interaction, and mutualism

Table 1.1: Classification of ecological interactions based on the impact it has on the interacting species, which can be negative (−), neutral (0), or positive (+).

Species B \ Species A	−	0	+
−	Competition	Amensalism	Predation
0	Amensalism	Neutralism	Commensalism
+	Predation	Commensalism	Mutualism

or cooperation as a mutually beneficial interaction [23]. Instead of fighting over limited resources, cooperation has the potential to expand the resources available. Indeed, it is humans' remarkable capacity for cooperation that has enabled an escape from Malthusian catastrophe in modern times through a dramatic increase in the food supply [24, 25].

This poses a paradox, however: why compete at all if it harms everyone involved, and why are not all interactions cooperative if it is in everyone's interest? These questions reveal several important properties of competition. Firstly, the precondition for competition in the Malthusian model is scarcity. When the only alternative is starvation, competing with others is inevitable, even at the risk of being harmed. In contrast, there is no need for competition when resources are abundant, so there is more to be gained from cooperation than from competition. It is instructive in this context that the Allee effect, which models intraspecies cooperation, is observed in low population density regimes, when resources are plentiful relative to population size. Secondly, in the examples of gut bacteria and mycorrhiza, cooperation occurs between two species with different and complementary needs. As the dictionary definition of competition at the start of this chapter indicates [1], competition occurs when individuals are vying for similar resources.

The more subtle resolution to this paradox, however, is that competition and cooperation are not absolute opposites, and that they serve each other when viewed from multiple perspectives. After all, "survival of the fittest" within a species by definition selects for organisms that are most adapted to the environment, which is an advantage for the continuation of the species, even though it is a harsh reality from the perspective of the loser. At the level of ecosystems, the principle of **competitive exclusion** states that two species cannot occupy the same ecological niche in the long term [23]. A classic example involving barnacle species is illustrated in Figure 1.1. Thus, competing species are forced into divergent ecological niches, which can contribute to evolutionary changes over time [26]. Therefore, competition is a driving force of speciation and biodiversity, which undergird the resilience and robustness of the biosphere.

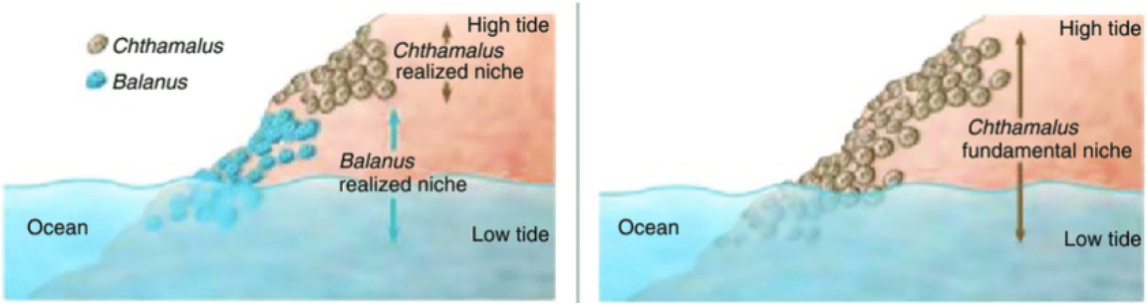

Figure 1.1: Connell [27] observed that *Chthamalus* and *Balanus* barnacles are always found on the upper and lower parts of coastal rocks, respectively. When he removed *Balanus* from the rock, he observed that *Chthamalus* readily colonised the lower reaches of the rock. Upon the reintroduction of *Balanus*, *Chthamalus* is displaced by the more competitive *Balanus*. Since *Chthamalus* is excluded by the presence of a competitor from a niche which it can populate in the absence of said competitor, it is an example of competitive exclusion. Figure reproduced from [28] by permission of Pearson Education, Inc.

In addition, competition can contribute to stability, as exemplified by the dominance hierarchies of social animals. Early studies of dominance hierarchies reported that birds living in groups adhere to a rigid social hierarchy where higher-ranking birds enjoy priority access to resources and mates [29, 30]. The bird hierarchy is known as the "pecking order" because, once a dominance–subordination relationship has been established, the dominant bird gains pecking rights over the subordinated bird, where the former routinely pecks the latter but not the other way around. Crucially, birds display aggressive behaviour while forming a pecking order, but stop doing so once a pecking order is established [31]. Here, competition is a mechanism for establishing a social order that reduces intragroup conflict and enables the group to function as a cohesive social unit [32].

To summarise, competition and cooperation are interdependent forces, despite appearing at opposite ends of the ecological spectrum of interactions. Competition creates winners and losers, to the detriment of the latter especially, but also at a significant cost to the former. However, the benefits of competition are not to be found at the level of the competing individuals, but rather at the level of the group to which they belong. This suggests that competitive processes can evolve within a cooperative structure, such as the multicellular organism, to benefit that structure.

An illustrative example is found in the adaptive immune system. During T cell maturation, a highly heterogeneous population of thymocytes (T cell precursors) is generated through a biological gene editing process that remixes gene fragments encoding parts of the T cell receptor in order to produce a diverse array of thymocytes with different T cell receptors [33, 34]. This diversity enables a wide range of antigens to be recognised by the

adaptive immune system. T cell receptors must navigate a delicate balance however [35]. On the one hand, they need to bind the major histocompatibility complex in order to perform their function. On the other hand, they cannot bind too strongly, or else the T cells may direct the immune system to fight healthy cells (this is known as autoimmunity). The maturation process selects for functional T cells by inducing apoptosis, a form of programmed cell death, in thymocytes that do not fulfil these criteria, killing over 95% of all thymocytes [36].

T cell maturation demonstrates how competitive processes can be leveraged by the multicellular organism for its benefit. There are, however, important differences between T cell maturation and cell competition. Firstly, in T cell maturation, the cells being eliminated would impede the function of the adaptive immune system if they were allowed to survive. In contrast, the losers in cell competition are perfectly viable on their own and capable of adequately performing the functions of the tissue. Secondly, the selection of thymocytes is based on the binding of T cell receptors to antigens that are presented to all thymocytes equally. Cell competition, on the other hand, is predicated on interactions *between* competing cells, as opposed to an interaction with a shared environment. Hence, cell competition is **context-dependent**; the same cell type can be both a loser or winner cell type, depending on the cell type that it is competing with.

We conclude this introduction with the proposition that cell competition is similar to conventional competition in the ecological sense with respect to the *outcome*, but different with respect to the *process*. As in conventional competition, only the winner cell type remains, which is analogous to the phenomenon of competitive exclusion. On the other hand, the manner by which cells compete is more akin to that of a sports competition, where both sides have agreed to compete under a predefined set of rules. A game, however, with the stakes of life and death. Instead of a brutal fight to the death, competing cells calmly exchange messages with their neighbours to determine their status through genetically encoded rules and gracefully initiate programmed cell death when they "lose". Indeed, the sportsmanship-like manner in which cell competition proceeds reflects the fundamentally cooperative nature of the multicellular structure in which it unfolds.

1.1 Thesis aims and structure

As we shall see in **Section 1.2**, where we review the literature on cell competition, experiments have uncovered various genes, messenger molecules, and pathways involved in cell competition. However, it is not known how relative fitness is determined or communicated among cells. Since cell competition evolved within the cooperative framework of the multicellular organism, it must involve certain rules to discriminate winners from losers.

However, the nature of these rules has remained elusive thus far. Therefore, one of the main research aims in the field of cell competition is to uncover the rules and mechanisms of cell competition. This is the underlying motivation for this thesis.

Cell competition is a process that unfolds at the level of the cell population, but is mediated by interactions at the cellular level. Verbal models are generally not rigorous enough to accurately predict complex behaviour of cell populations from mechanisms specified at the cellular level. Mathematical models, on the other hand, allow us to precisely define interactions and study their implications on system properties and dynamics. In particular, in cell-based models we can define the behaviours of individual cells and investigate the patterns that emerge at the level of the population. Hence, in this thesis the fundamental framework for investigating the mechanisms of cell competition is cell-based modelling.

Cell competition was first discovered in the *Drosophila* wing disc [37], and it has been the model system for cell competition ever since. The wing disc is an epithelial tissue that develops during the larval stage of *Drosophila* development, and is the precursor to the wing in the fully-formed adult. Although the modelling framework we develop in this thesis can be applied more widely, we use the *Drosophila* wing disc as a grounding example, particularly when studying the spatial dimensions of cell competition. In **Section 1.3**, we review different cell-based models and conclude that the most appropriate cell-based model for the *Drosophila* wing disc is the vertex model.

In **Section 1.4**, we review mathematical models of competition in general, and of cell competition in particular (Section 1.4.5). The main weakness of current models of cell competition is that they assume *a priori* winner and loser identities. Although they can simulate processes occurring downstream of winner/loser identification, they do not shed any light on *how* a cell becomes a winner or loser. Therefore, in this thesis we address the question: *how do cells acquire winner or loser status?* Specifically, the central aim of this thesis is to study the emergence of winners and losers in cell competition using mathematical modelling. We decompose this aim into three research objectives:

Research Objective 1: construct a model of cell competition that produces winners and losers as an emergent phenomenon.

Research Objective 2: use this model to identify and study the key factors driving winner/loser identification.

Research Objective 3: use predictions from the model to propose experiments for validating the model.

It has been observed, as Section 1.2.5 discusses in more depth, that mechanical interactions can play an important role in cell competition. Therefore, in our first attempt at

tackling Research Objective 1, we implement a vertex model for heterotypic populations that lets us independently vary the mechanical parameters of two distinct cell populations. We then run a systematic parameter sweep and record competitive outcomes to test whether differences in cell mechanical properties can result in cell competition. This is described in **Chapter 2**. Briefly summarised, we observe the emergence of winners and losers, but homotypic simulations of the loser cell type reveal that the losers are intrinsically nonviable. This does not meet the criteria for cell competition, so we conclude that mechanical interactions in this particular model are not sufficient to induce cell competition.

Motivated by these results, we propose a modelling framework that explicitly models competition-induced cell death in **Chapter 3**. In short, we assume that cells perceive **death signals** from the extracellular environment that are generated by other cells, directly or indirectly. Moreover, every cell is equipped with a **death clock**. This is a scalar variable that is incremented by death signals and represents the cell's memory of received death signals. When the death clock reaches a threshold, the cell initiates apoptosis. As such, the death clock abstracts the process of intercellular interactions leading to competition-induced apoptosis. We also discuss the implementation of the death clock framework in two concrete cell-based models (vertex model and well-mixed model). We use these in later chapters to run simulations that include permutations of the death clock mechanism and verify theoretical predictions.

One advantage of the death clock framework is that the death signal can take many forms and can therefore model virtually any type of intercellular interaction. In **Chapter 4**, we investigate the effects of a constant death signal, the simplest possible death signal. Although a constant death signal cannot represent cell competition because it is not based on intercellular interactions, it is a useful and simplified model for investigating basic properties of the death clock framework. Furthermore, in later chapters we approximate more complex death signals using constant values. Hence, Chapter 4 forms a theoretical foundation for later work.

For cell-based models used in this thesis, we assume a simplified cell cycle consisting of two phases: G1 phase and G2 phase. In G1 phase, the cell has not fully committed to division and is vulnerable to apoptosis. In G2 phase, the cell is fully committed to division and can no longer initiate apoptosis. Since the winners in cell competition are often cells with higher proliferation rates, and G2 phase is associated with a commitment to proliferation, we propose a model in which cells in G2 phase emit death signals to their neighbours. In **Chapter 5**, we examine the effects of emission of this G2 death signal in a homotypic population; the homotypic survival probability is significant because it is a measure of intrinsic viability. After making some approximations, we are able to predict the

survival probability of cells as a function of cell cycle and death clock parameters. Moreover, we use cell-based model simulations to verify these predictions.

In **Chapter 6**, we apply the G2 death signal model to heterotypic populations. Importantly, we can reproduce the emergence of winners and losers in simulations as a direct result of competitive interactions, without assigning *a priori* winner/loser identities, thus achieving Research Objective 1. A crucial distinction with Chapter 2 is that the emergent winner and loser cell types are intrinsically viable in homotypic conditions. We also develop a theoretical framework that enables us to classify competitive interactions and predict the outcome of competitive interactions. We also validate these results using cell-based simulations. In addition, we discuss how this framework lets us study different competition regimes and how they depend on model parameters. Hence, we achieve Research Objective 2 in this chapter as well.

The modelling framework developed in this thesis is able to reproduce the hallmarks of cell competition, but also makes a number of other predictions that can be experimentally verified. In **Chapter 7**, we summarise our work and discuss how the death clock framework can be experimentally validated, thus achieving Research Objective 3. In this thesis we only implement the death clock framework in two cell-based models and for two particular death signals, but the death clock framework is very flexible and can be implemented in other cell-based models and with other death signals. Hence, in Chapter 7, we suggest potential future avenues for research on cell competition using the death clock framework.

1.2 Overview of cell competition

1.2.1 *Minutes* and *Myc*

As mentioned earlier, cell competition inside a multicellular organism was first observed in the development of the wing disc in *Drosophila*, also known as the fruit fly [37]. Fruit flies are often used as model organisms for animal development because they are easily cultured, have a short reproductive cycle, and are very amenable to genetic manipulation. Moreover, the majority of studies in cell competition are performed on epithelial tissues, both *in vivo*, such as in the wing disc, and *in vitro*, using co-cultures, because it is easier to observe cell–cell interactions within a two-dimensional sheet of cells. In the study [37], the authors were investigating a set of mutants called *Minutes* (*M*)[2], which contain mutations in genes that encode ribosomal proteins and cause cells to divide at a slower rate [38]. Flies that are homozygous for a *Minutes* mutation are not viable, but heterozygous flies (*M*/+)

[2]In order to distinguish between genes and proteins, we italicise gene names.

can survive, although their growth is mildly retarded. When genetic mosaics containing both wild-type and *M*/+ cells were artificially induced, only wild-type cells remained in the adult. This seminal study suggested that proliferation rate is an important factor of cell fitness. Indeed, subsequent studies showed that the severity of cell competition is greater for variants of *Minutes* that grow slower [39]. Moreover, in conditions of starvation, when nutrients are the limiting factor for growth rather than protein synthesis, *Minutes* clones regain their competitive edge and survive to adulthood [40].

In the case of *Minutes* mutations, wild-type cells prevail and mutated cells, which proliferate at a lower rate, are removed. The fact that cell competition occurs in the wing disc during a period of rapid cell proliferation suggests that cell competition is a mechanism intended to keep tissues healthy by weeding out suboptimal cells, which in this case are slowly proliferating cells. However, later studies involving the growth regulator *Myc* in *Drosophila* have complicated the role of cell competition considerably [41, 42, 43]. Similarly to *Minutes*, cells that were heterozygous for *dmyc* (the *Drosophila* homologue of the *Myc* gene) grew slower and were outcompeted by wild-type cells. However, cells that were genetically transformed to have three copies of *dmyc* grew at a faster rate than the wild-type cells (who had just two copies), and were stronger competitors as a result. This was the first observation of **super-competition**; mutated cells that outcompete and displace wild-type cells. A schematic summary of cell competition and super-competition is given in Figure 1.2.

The discovery of super-competition through overexpression of *Myc* is particularly significant because *Myc* is a known proto-oncogene [44]. Super-competition may be a mechanism by which mutated cells establish themselves in a tissue at the expense of surrounding, healthy cells without producing detectable morphological abnormalities [45]. Additional oncogenic mutations within a population of super-competitors are then more likely to result in tumour formation. This mechanism could help explain the clinical phenomenon of "field cancerisation"; this occurs when a large field of cells in a tissue is affected by oncogenic mutations, from which multiple tumours develop [46].

1.2.2 Cell polarity

Interest in cell competition has exploded in the past decade, which has led to the discovery of many more competitive cell–cell interactions. Most notably, cell polarity in epithelial tissues has been identified as an additional parameter of cell fitness [47]. The apicobasal polarity of epithelial cells is established and maintained by an asymmetric distribution of protein complexes throughout the cell [48]. Some of the genes that encode cell polarity proteins, *Scribble*, *Discs large* and *Lethal giant larvae*, can be mutated and still produce viable cells.

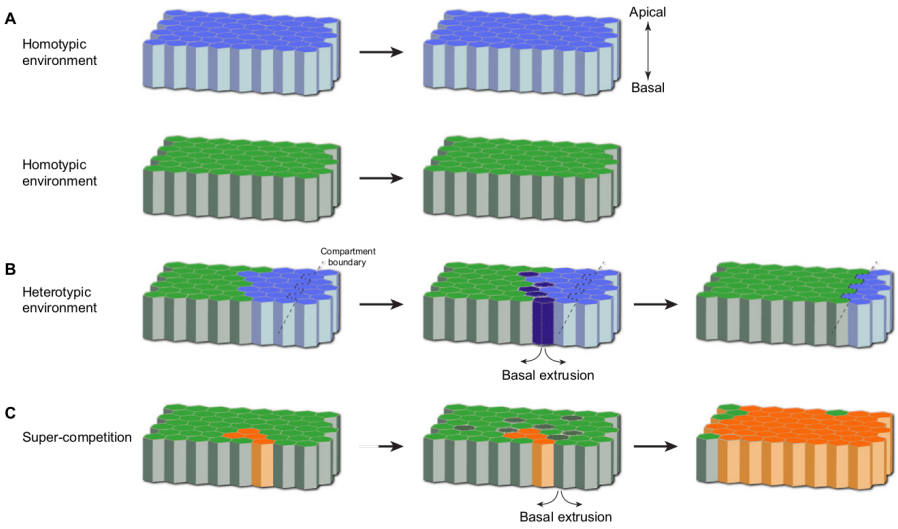

Figure 1.2: Schematic summary of cell competition and super-competition. (A) In a homotypic environment, both genotypes are viable and result in normal tissues. Blue cells (top) represent less fit cells and green cells (bottom) represent wild-type cells. (B) When wild-type and less fit cells are present in the same tissue, the cells compete and only the wild-type cells remain. (C) When wild-type cells and super-competitors (orange) are placed in the same environment, the super-competitors, which grow faster than wild-type cells, outcompete wild-type cells. Figure reproduced[3] from [21].

In fact, the mutated cells proliferate at a faster rate than wild-type cells. In addition, the loss of polarity compromises epithelial structure and results in tumour-like overgrowths. However, when polarity-deficient cells are competing with wild-type cells in a genetic mosaic, they are eliminated by cell competition [47, 49, 50, 51, 52]. Cell competition, it seems, is a double-edged sword with respect to cancer; it can promote tumour formation through super-competition but also prevent tumours by eliminating polarity-deficient cells.

1.2.3 Mediators of cell competition

At first sight, the prominent role of proliferation rate in determining cell fitness may suggest that cell competition is the result of simple arithmetic; loser cells divide less often and are thus less likely to populate the tissue. However, experiments show that suppression of the pro-apoptotic protein Hid inhibits cell competition, indicating that loser cells are actively eliminated by apoptosis [53]. Moreover, activating cell growth through alternative means, such as overexpressing the cell cycle regulators cyclin D and Cdk4 [42], does *not* trigger cell competition. We can therefore conclude that cell competition is an active and regulated

[3]With permission of The Company of Biologists Ltd.; permission conveyed through Copyright Clearance Center, Inc.

process. Several pathways whose activation correlate with competitive interactions have been identified, but a general mechanism for cell competition remains elusive.

Intriguingly, it was demonstrated that the signalling pathway leading to apoptosis in *Minutes*- and *Myc*-induced competition shares components with pathways involved in the innate immune system [54]. This raises the possibility that cell competition shares an evolutionary history with the immune system. Similarly to how immune cells target damaged or cancer cells in the body and eliminate them by inducing apoptosis, competing cells may be exerting a type of quality control among neighbouring cells.

A genetic study of cell competition in the wing disc revealed that the Flower (Fwe) cell membrane protein is responsible for "encoding" cell fitness [55]. The *Flower* (*Fwe*) gene encodes three splice isoforms; the Fwe^{ubi} isoform is expressed ubiquitously in the wing disc, while the isoforms $Fwe^{Lose\text{-}A}$ and $Fwe^{Lose\text{-}B}$ are found only in loser cells. Moreover, the presence of Fwe^{Lose} proteins is sufficient to mark cells as losers and trigger apoptosis. According to the "fitness fingerprint" hypothesis, cells compare their fitness locally with neighbouring cells based on their Fwe membrane proteins. However, the mechanism that causes differential expression of Fwe^{Lose} in the first place is not known.

1.2.4 Conservation in mammals

Cell competition is not restricted to *Drosophila*. The first evidence for mammalian cell competition came from research into a mutation in mice that is analogous to *Minutes*, specifically the *Belly spot and tail* (*Bst*) mutation [56]. Similar to *Drosophila*, *Bst* is a heterozygous mutation in a ribosomal protein gene that confers a growth disadvantage to afflicted cells. Later studies demonstrated that cell competition due to differential *Myc* expression occurs in the mouse epiblast, further confirming the conservation of cell competition in mammals [57, 58].

More recently, a study of mouse development suggests that cell competition is not only conserved in mammals, but that mammalian tissues have evolved additional modes of cell competition as they increased in complexity [59]. During mouse embryogenesis, the skin develops in two distinct stages. In the first stage, the epidermis, consisting of a single layer of epithelial cells, expands its surface area roughly 30-fold to accommodate a rapidly-growing embryo. In the second stage, the epithelium transitions to a stratified barrier with a layer of progenitor cells at its base. The progenitor cells can divide in roughly two ways: either it divides symmetrically to yield two progenitor cells, or it divides asymmetrically with one daughter cell remaining in the basal progenitor layer and the other daughter cell expelled outwards to contribute to the outer layers of the barrier. Importantly, the expelled cells are

programmed to terminally differentiate and are eventually shed as dead skin cells.

In the study [59], the authors proved that, during the first stage, cell competition is caused by differential expression of *Mycn*, an isoform of *Myc*. Moreover, the loser cells are cleared by apoptosis, which is similar to cell competition in the *Drosophila* wing disc. In the second stage, cell competition occurs as well, but loser cells are no longer removed by apoptosis. Instead, loser cells are induced to preferentially divide asymmetrically, such that a larger proportion of loser cells differentiate into cells that eventually die. These findings demonstrate that, even within the same tissue, multiple modes of cell competition are possible, and that tissues dynamically adapt their mechanism for cell competition to the developmental context.

1.2.5 Mechanical cell competition

When groups of cells in the same tissue proliferate at different rates, it inevitably leads to mechanical stresses. Therefore, it has been hypothesised that mechanical interactions could play a role in cell competition [60]. Experiments later demonstrated that cell crowding in epithelial tissues triggers an intrinsic form of cell competition where genetically identical cells compete for space [61, 62]. In these studies, the cell density was artificially increased, initiating a competitive process where some cells were squeezed out by their neighbouring cells to relieve excessive mechanical stresses in the tissue. The extruded cells then proceeded to die through apoptosis. Since the cells were genetically identical, the competition was presumably predicated on intrinsic, stochastic cell–cell variability of mechanical properties.

This discovery reignited interest in the contribution of mechanical stresses to competition between genetically distinct cell populations [20]. Perhaps phenotypic differences in mechanical sensitivity as a result of genetic divergence are at least partly responsible for cell competition. This hypothesis is supported by recent research; one study provided evidence of mechanical super-competition achieved by the overexpression of the growth regulator Ras to simultaneously stimulate cell growth and decrease mechanical sensitivity [63]. Another study demonstrated that cells silenced for *Scribble* are more likely to delaminate under mechanical stress [64]. In the presence of wild-type cells, *Scribble*-deficient cells are outcompeted due to their mechanical hypersensitivity.

Mechanical cell competition is a relatively new discovery. Hence, not much is known about the mechanoreceptors (receptors that respond to mechanical stress) and mechanotransduction pathways involved. The mechanotransducer Piezo1 plays a role in both cell crowding and cell stretching [62, 65]. However, it does not seem to be involved with *Scribble*-related cell competition [64]. On the other hand, the relationship between Ras-induced

mechanical super-competition and Piezo1 has not yet been investigated [66].

1.3 Overview of cell-based models

In this section, we provide a high-level review of cell-based computational models. This means that we restrict ourselves to relatively basic descriptions that nonetheless allow us to discuss potential strengths and weaknesses in modelling cell competition. The models are introduced in order of increasing complexity; we start with cellular automata, where cells are represented as shapeless entities inhabiting a fixed lattice, and end with vertex models, which explicitly model cell shape.

Broadly speaking, cell-based models can be categorised into two classes of models; on-lattice models and off-lattice models. In the former, space is discretised into a fixed lattice whose sites can be occupied by cells. In off-lattice models, on the other hand, cells are represented by objects that reside in a continuous space.

1.3.1 On-lattice models

On-lattice models assign each simulated cell to one or more sites of a two- or three-dimensional lattice, and iteratively update cell positions based on a set of "update rules". These update rules generally sample each cell's local neighbourhood to decide which action to take. For instance, a cell that is surrounded only by empty spaces may be eligible for cell division, which is performed by placing an additional cell in the empty space. Even with relatively simple update rules, it is possible to simulate phenomena such as cell migration, cell proliferation, cell–cell adhesion, and so on [67].

Cellular automaton model

In cellular automaton models (CAMs), each cell occupies exactly one site on the lattice (see Figure 1.3.a). This constraint makes the CAM the simplest type of cell-based model, as well as the most computationally efficient [68, 69, 70]. Hence, their main advantage is that a large number of cells, and their interactions, can be simulated at once. As a result, CAMs can be used to reproduce the dynamics of mesoscopic structures for which more sophisticated models are computationally intractable.

On the other hand, the simplified spatial representation of CAMs necessarily produces simulation artefacts, such as anisotropic biases in cell migration. Moreover, because CAMs evolve on the basis of discrete update rules instead of continuous laws of motion, it is not obvious how to translate between model parameters and experimentally derived parameters.

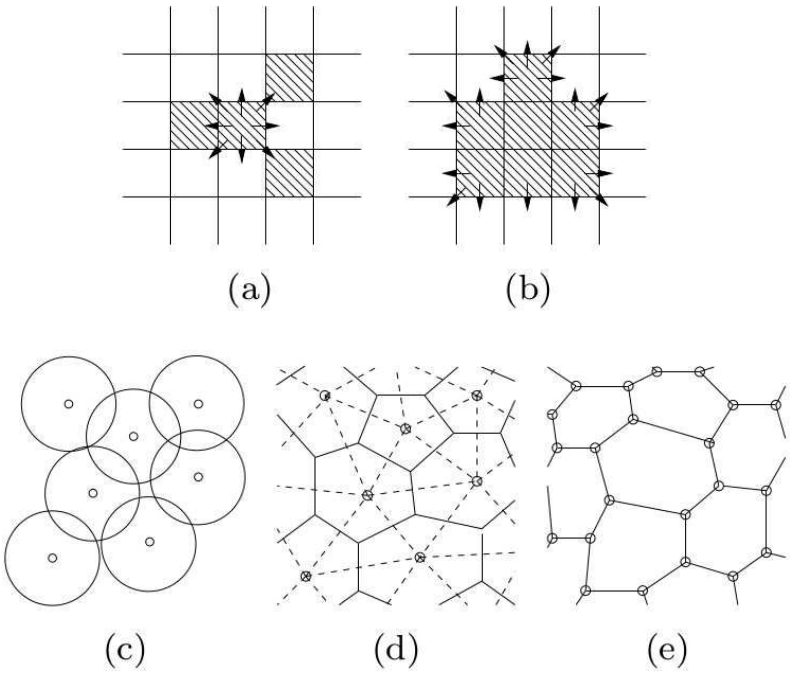

Figure 1.3: Overview of cell-based models. (a) Cellular automaton model. (b) Cellular Potts model. (c) Overlapping spheres model. (d) Voronoi tessellation model. (e) Vertex model. Figure reproduced[4] from [71].

Most importantly, CAMs do not model cell shape and are therefore unable to model processes where mechanical cell deformations play a role.

Cellular Potts model

The cellular Potts model (CPM) is also an on-lattice model, but, unlike the CAM, it explicitly accounts for cell shape by allowing cells to extend over multiple lattice sites [72]. The update rules for CPMs are derived by defining an energy function over the system and performing cell identity swaps to minimise the energy through a Monte Carlo process. Because cells are characterised in greater detail than in CAMs, it is possible to model mechanical properties such as cell membrane tension, cell–cell adhesion, and cell elasticity. In turn, this enables the application of CPMs to a wider range of biological processes, such as morphogenesis [73] and tumour growth [74].

Despite these advantages however, CPMs still suffer from many of the same drawbacks that CAMs have. For instance, it is difficult to interpret CPM parameters in an experimental context, and spatial anisotropy remains an issue. It is possible to reduce the spatial bias by using a very fine grid to resolve cell shapes, but this increases the computational cost. We

[4]Under CC BY license (`https://creativecommons.org/licenses/by/4.0/`).

will not use on-lattice models in this thesis, but this does not mean that on-lattice models have no value in the context of cell competition. On the contrary, there are examples of on-lattice models of competition in an ecological context [75, 76], in a microbial context [77], and in the context of cell competition in multicellular organisms [78].

1.3.2 Off-lattice models

In contrast to on-lattice models, off-lattice models treat the spatial domain as a continuous medium in which cells move and interact. Often, this medium is assumed to be a viscous environment, such that cell motion occurs in a low Reynolds number and inertial forces can be neglected [79]. Hence, the equation of motion for a "particle" i with position $\mathbf{r}_i$, experiencing the total force $\mathbf{F}_i$, has the form

$$\mu \frac{\mathrm{d}\mathbf{r}_i}{\mathrm{d}t} = \mathbf{F}_i \,, \tag{1.1}$$

where μ is the friction coefficient. The meaning of a particle and the calculation of the total force depends on the type of off-lattice model under consideration. For instance, in the overlapping spheres model (OSM) and Voronoi tessellation model (VTM), particles correspond to cell centres, while particles correspond to vertices in the vertex model.

Different types of off-lattice models are distinguished by their description of cell shape. For instance, the OSM and VTM both track cell centres, but they differ in how they represent cell shape and, as a result, how they calculate intercellular forces.

Overlapping spheres model

As the name suggests, the OSM assumes that cells have a spherical shape, which may deform upon contact with other cells or a substrate [80]. When cells come into contact (i.e. their spherical shapes "overlap"), they exert forces on each other. Typically, the intercellular force is attractive at an intermediate distance (cell–cell adhesion), but becomes repulsive when cells approach too closely.

In comparison with off-lattice models, the OSM provides a view of cellular interactions that is more physically relevant, and is therefore easier to relate to experiments. In particular, intercellular force laws are often adopted from experimentally derived cell–cell contact models [81]. However, the OSM lacks an explicit description of cell connectivity, which is an important feature of epithelial tissues and could contribute to mechanical cell competition.

Voronoi tessellation model

Similarly to the OSM, the VTM identifies cell positions with their cell centres. Unlike the OSM, cell shapes are derived by applying a Voronoi tessellation on the cell centres [82]. This means that a cell is defined as the set of all points that are closer to its centre than to the centres of any other cell. This operation generates a mesh of polygonal cells, where each cell has well-defined neighbours that they interact with (see Figure 1.3.d). Furthermore, the cell connectivity is given by the Delaunay triangulation of the cell centres (indicated by the dashed lines in Figure 1.3.d).

The VTM offers a number of advantages over the OSM in the context of epithelial tissues. Firstly, cell connectivity is explicitly expressed through a polygonal mesh, which mimics the spatial structure of epithelial tissues *in vivo*. Secondly, the Voronoi tessellation enables direct comparisons between computational results and the summary statistics that are commonly used to describe epithelial tissues in experiments. For example, the number of cell neighbours or the mean area per polygon class[5] can be computed from VTM simulations *and* can be derived from experimental results.

Unfortunately, in VTMs there is no explicit control of cell volume or shape. Although the cell connectivity is topologically similar to that of an epithelial tissue, cell–cell interactions are limited to pairwise forces between cell centres, as in the OSM.

Vertex model

The main issue with cell-centre models is that they localise cells to a single point in space. As a result, cell shape can only be inferred implicitly, by centring a sphere on the cell centre as in the OSM, or by applying a Voronoi tessellation as in the VTM. Of these two models, the VTM is more appropriate for epithelial tissues because it simulates the topology of epithelial tissues. However, the polygonal mesh structure is not an intrinsic feature of the VTM, but is instead an intermediate representation of the tissue that is derived from the positioning of the cell centres.

Vertex models, on the other hand, reverse this relationship by using an explicit polygonal mesh that represents the epithelial tissue and identifying cells directly with the polygons in the mesh [83, 84, 85, 86]. The notion of a cell centre does exist in vertex models, as the geometric centroid of the polygon corresponding to a cell, but it does not feature in the formulation of the vertex model. Rather, the primary components of the vertex model are the edges and vertices of the mesh, which correspond to cell–cell interfaces and locations

[5]The polygon class of a cell is defined as the number of edges it has. For instance, all triangular cells belong to the same polygon class.

where three or more cell–cell interfaces meet, respectively. Importantly, the dynamics of cells are determined by forces *acting on the vertices*, hence the name **vertex** model.

The benefits of the vertex model are similar to those of the VTM when modelling epithelial tissues since its polygonal mesh can be directly related to experimental data. In addition, however, we gain explicit control over cell shape, which enables the formation of more detailed hypotheses about the mechanical factors involved in cell competition (see Section 2.1 for more details on the vertex model). As a result, we chose vertex models as our initial modelling approach for modelling cell competition. No choice comes without a cost, however, and a potential downside of the vertex model is that it is computationally expensive. This can be partly mitigated by parallelisation, however.

As a final note, we should mention that there are other cell-based models which describe cells in even greater detail. For instance, a "many vertex" model has been proposed to model the gastrulation of sea anemones, where each cell is represented with a polygon composed of many vertices so that they can take on arbitrary cell shapes [87]. Another example is the subcellular element model, which takes a similar approach to the vertex model (or OSM), but also partitions the interior of the cell into polygons (or spheres) to simulate mechanics at a subcellular scale [88]. Although such models are more detailed than vertex models, they come at the price of further increased computational cost. In addition, there are currently no indications that such a level of detail is vital to understanding cell competition.

1.4 Overview of competition modelling

In this section, we review mathematical models relevant to cell competition. Population dynamics (Section 1.4.1) is a framework for modelling ecological interactions at the population level, and has provided theoretical insights such as competitive exclusion. Evolutionary game theory (Section 1.4.2) applies game theory to the evolution of strategies in populations. It is particularly successful at explaining frequency-dependent selection, where the fitness of a trait is dependent on the abundance of traits in the population. We also briefly discuss competition and models of competition in microbial ecology (Section 1.4.3) and cancer biology (Section 1.4.4), since these domains are concerned with similar spatiotemporal scales as cell competition. Finally, in Section 1.4.5, we review current models of cell competition.

1.4.1 Population dynamics

Population dynamics aims to describe the growth and decline of populations using mathematical modelling [89, 90]. Similarly to the general field of dynamical systems, time can be conceptualised as being discrete or continuous. In discrete-time models, populations are assumed to grow in non-overlapping generations, while generations overlap in continuous-time models. Such models are expressed in terms of difference and differential equations, respectively. In addition, the basic framework can be extended to include population structure, time delay, spatial distribution, stochasticity, and more [89, 90]. Nevertheless, the biological reality is usually more complicated than the relatively simple mathematical models used in population dynamics. These models are therefore most often used for conceptual insights, rather than quantitative predictions. A notable (and topical) exception, however, is infectious disease modelling, where the aim is to make concrete predictions and inform public health policy [91, 92].

Here, we will focus on summarising a select number of models relevant to competition. Denoting the population size by x, one of the simplest population models is the **Malthus model** [2],

$$\frac{\mathrm{d}x}{\mathrm{d}t} = rx\,, \quad x(0) = x_0\,, \tag{1.2}$$

which assumes a constant per capita growth rate, denoted r. When r is positive or negative, the model predicts exponential population growth or decline, respectively. This model is appropriate when resources are not a limiting factor.

A basic model for resource-constrained populations is the **Verhulst model** [93, 94],

$$\frac{\mathrm{d}x}{\mathrm{d}t} = rx\left(1 - \frac{x}{K}\right), \quad x(0) = x_0\,, \tag{1.3}$$

also known as the logistic model. In this model, the per capita growth rate declines with increasing population size as a result of intraspecific competition. When the population size reaches the carrying capacity, denoted K, the growth rate is zero. The carrying capacity is the maximum population size that the environment can sustain.

When two species, denoted x and y, compete for the same resources[6], we can represent this with a **Lotka–Volterra** model [89, 90, 95]:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = r_x x\left(1 - \frac{x + b_{xy}y}{K_x}\right), \quad x(0) = x_0\,, \tag{1.4a}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = r_y y\left(1 - \frac{y + b_{yx}x}{K_y}\right), \quad y(0) = y_0\,, \tag{1.4b}$$

[6]Or, more generally speaking, when two species inhibit each other's growth.

where the constant b_{xy} quantifies the competitive effect of species y on x, and vice versa for b_{yx}. Analysis of Equations (1.4a) and (1.4b) shows that stable solutions where $x > 0$ and $y > 0$ are only possible if $b_{xy} < K_x/K_y$ and $b_{yx} < K_y/K_x$. In other words, the competing species can only coexist if the effect of interspecific competition is smaller than intraspecific competition. In all other cases, the model predicts competitive exclusion.

Populations can also be modelled on spatial domains. When combining Equations (1.4a) and (1.4b) with a diffusive process representing dispersal, we obtain a **competition and diffusion model** [96]:

$$\frac{\partial x}{\partial t} = D_x \nabla^2 x + r_x x \left(1 - \frac{x + b_{xy} y}{K_x}\right), \quad x(\mathbf{r}, 0) = x_0(\mathbf{r}), \tag{1.5a}$$

$$\frac{\partial y}{\partial t} = D_y \nabla^2 y + r_y y \left(1 - \frac{y + b_{yx} x}{K_y}\right), \quad y(\mathbf{r}, 0) = y_0(\mathbf{r}), \tag{1.5b}$$

where D_x and D_y are the diffusion coefficients for species x and y, respectively, and where we used $\mathbf{r}$ to denote the spatial variables. This model has been used to show that the spread of an invading species outcompeting a resident species is not qualitatively different from the spread of a species in the absence of competition, except that the spread occurs more slowly [96].

1.4.2 Evolutionary game theory

Evolution can be conceived of as a process that optimises the fitness of organisms with respect to their environment. In fact, biological evolution is the inspiration for "evolutionary algorithms"; heuristic optimisation algorithms that use the principles of random mutation, reproduction, and selection to generate approximate solutions for complex optimisation problems [97]. The analogy between evolution and optimisation is therefore quite apt, but it has limitations. In particular, the mathematical notion of optimisation implies a static utility function. Therefore, evolution as optimisation portrays the struggle for existence as a struggle against a fixed environment. However, as we discussed at the start of this chapter, the struggle for existence is also a struggle against other organisms, i.e. a competition in a dynamic environment.

Therefore, when the success of an organism is not only determined by its environment, but also by the organisms that it encounters in its environment, the appropriate framework is game theory [98], not optimisation. Similarly to optimisation, the aim in game theory is to maximise utility. However, this aim is pursued from the perspective of an agent playing "games" with other agents. Despite the light-hearted term, a game simply means that each

agent can choose from a range of strategies, and that the utility of each strategy depends on the strategies chosen by the other agents. In classical game theory, there is therefore no "optimal" strategy, only Nash equilibria, in which each player's strategy is optimal given the strategies of the other players [99].

Evolutionary game theory (EGT) applies game theory to evolution by conceptualising interactions between organisms as games. In this framework, the phenotypes that organisms (players) display are the strategies that they employ in the game. Moreover, the utility of the game is determined by the fitness of the participating organisms that result from the interaction. The process of evolution is then represented by a population of organisms that compete through pairwise games and reproduce based on their fitness[7]. Crucially, the success of a phenotype is *context-dependent*, since its fitness depends on the distribution of phenotypes in the population. Again, there is no optimal phenotype, only an evolutionarily stable state[8], which is a distribution of phenotypes that cannot be invaded by a mutant phenotype. In many cases, the evolutionarily stable state is not a single phenotype, but a mixture of phenotypes, demonstrating that competition does not necessarily produce one single "fittest" phenotype, but rather a population with an evolutionarily stable composition of phenotypes.

EGT has been particularly successful in describing the evolution of traits that are under **frequency-dependent selection**, such as patterns of animal conflict [100], the sex ratio [101], and the evolution of cooperation [102], to name but a few. In addition, EGT has been applied to evolutionary processes at the cellular [103] and molecular [104] scale. The remarkable insight from EGT is that cooperative behaviours can emerge and thrive in a framework that describes, at its core, a competitive process among self-interested agents. EGT therefore demonstrates that cooperation is fundamentally compatible with the struggle for existence.

Standard assumptions in EGT often include random mixing of opponents. But evolutionary games can also be played on spatial domains, with players occupying sites on a lattice [105] or nodes in a network [106], and games being played locally between immediate neighbours. Spatial games permit the coexistence of strategies that are mutually exclusive

[7]For simplicity, EGT usually dispenses with the details of genetic inheritance, instead assuming that organisms reproduce asexually and that phenotypes are inherited without mutations.

[8]The evolutionarily stable state is closely related, but not identical, to the evolutionarily stable strategy. The former refers to a polymorphic population with organisms adopting "pure" strategies (or phenotypes), whereas the latter refers to a population where each organism has a mixed strategy, adopting each pure strategy with a certain probability. In most cases, the probability distribution of the mixed strategy is equivalent to the distribution of pure strategies in the polymorphic population, hence they are often used interchangeably. Even though the term "evolutionarily stable strategy" is more prominent, "evolutionarily stable state" is the more appropriate term here. For further discussion on this topic, see Appendix D in [98].

in random-mixing conditions [105], and can amplify or suppress the spread of strategies through network structures [106].

1.4.3 Microbial ecology

Microbes can be found virtually everywhere in the biosphere; they live in the water, the soil, and in the air. In addition, they inhabit multicellular organisms such as plants and animals in symbiotic relationships. According to endosymbiotic theory [107], mitochondria have a bacterial origin. Hence, in some sense bacteria may literally live *inside* our cells. Endosymbiosis aside, most microbes live in complex and diverse microbial communities. However, many microbial species cannot be cultured using current laboratory techniques. This makes it technically challenging to characterise all the species in a microbial community, let alone their interactions [108]. It is clear, however, that competitive interactions dominate in the microbial community [6, 109, 110, 111].

Similarly to the Malthusian model, the struggle for existence at the microscopic scale is centred on the competition for scarce resources [6]. Indeed, the resource ratio model of competition, proposed by Tilman [112] and based on the Monod equation for bacterial growth kinetics [113], suggests that the abundance of species can be predicted through patterns of nutrient consumption and availability. However, the resource ratio model can only account for **exploitative competition**, which is defined as competition that occurs indirectly through the consumption of a common limiting resource.

On the other hand, **interference competition** is competition where organisms actively inhibit their opponent's access to the contested resources. The latter form of competition is also common in bacteria [6], which have developed an impressive arsenal of weapons [7], including a diverse array of diffusing toxins [114, 115], molecular structures for perforating cell walls [116] and injecting toxins [117], and even viruses that specifically attack non-clonemates [118]. In contrast to animal conflict, which rarely escalates to physical fights, bacterial conflict is very aggressive and often lethal [7]. All of the modelling frameworks discussed so far have been applied in the study of microbial ecology, including cell-based models (Section 1.3), population dynamics (Section 1.4.1), and EGT (Section 1.4.2).

EGT models have been used to study the coexistence of cooperators and "cheaters" in microbial communities [103, 119, 120]. Many microbes produce extracellular enzymes (exoenzymes) to decompose complex nutrients into small molecules that can be absorbed through the cell wall. However, since the products of exoenzymes are released into the environment, they are "public goods" and the microbes producing exoenzymes can be considered cooperators. Cheaters, on the other hand, forgo the production of exoenzymes,

while still benefiting from the public goods. Game-theoretic models showed that cooperators and cheaters can coexist when the spatial scale in which public goods diffuse is smaller than the clustering size of cooperators [119, 120]. This result was verified *in vitro* [121], and it underlines the importance of spatial structure in microbial communities [110].

Population dynamics models play an important role in elucidating the large-scale structure of microbial communities from experimental data [108, 122], and provide a theoretical framework for describing interactions in microbial communities. For instance, it was demonstrated using Lotka–Volterra models that competitive interactions have a stabilising effect on microbial communities [123]. Even though cooperative interactions between different species of microbes can increase the total metabolic productivity, they introduce dependencies such that the decline of a single species may have a cascading effect on the whole community. Hence, there exists a trade-off in the composition of microbial communities between productivity and stability.

Cell-based models, on the other hand, are used to investigate fine-grained interactions between different microbial species and their environment [124]. For instance, cell-based modelling showed that spontaneous spatial segregation of genetic lineages is likely to occur in a wide range of conditions [125]. As discussed above, this emergent spatial structure promotes the evolution of cooperative phenotypes, since public goods can be preferentially shared among cooperators. Another application is in the production of polymers that form the backbones of biofilms. The predominant but naive view was that polymers are a public good, since non-polymer producing bacteria benefit from the protection that the biofilm offers. However, cell-based models revealed that polymer producing bacteria gain a competitive advantage by ensuring that their offspring are localised in the upper parts of the biofilm, and therefore have better access to oxygen [126].

1.4.4 Cancer biology

As discussed in Section 1.2, there are interesting parallels between cell competition and cancer, particularly the overlap in genes that are implicated in super-competition and cancer. Many of the modelling frameworks in cancer biology may therefore be relevant to cell competition. Mathematical oncology—the use of mathematical modelling to study cancer—is a very broad field for a number of reasons [127, 128, 129]. Firstly, applications span the entire spectrum from applied clinical research to fundamental biological research. Secondly, cancer is not a single disease, but a group of diseases with different properties depending on the tissue of origin and the mutations involved. Thirdly, tumour formation takes place at many different scales; genes, gene networks, signalling networks, cells, tissues, and organs.

Finally, tumour progression is a complex and multi-stage process. Hence, there is a plethora of models used in mathematical oncology tailored to different applications, types, scales, and developmental stages. Here, we can only provide a modest overview of models that may be relevant to cell competition.

The origins of cancer can be traced to mutations in genes that are responsible for controlling cell proliferation, cell death, and DNA repair, among others. Proto-oncogenes induce the cell to proliferate uncontrollably when overexpressed or improperly activated [130], while tumour-suppressing genes have a similar effect when their function is lost [131]. However, no gene lives in a vacuum, so these genes participate in complex intracellular signalling networks. The effects of perturbations on cellular responses can therefore not be derived by verbal reasoning alone. Hence, mathematical models are often used to describe their dynamics. Such models can take the form of deterministic differential equations [132], stochastic processes [133], or hybrids that combine both [134]. Examples include modelling of the MAPK pathway [135], the effects of cancerous mutations on the MAPK pathway [136], and p53 regulation [133].

As the tumour acquires more mutations and develops further, the interaction between the tumour and its microenvironment becomes increasingly important. This interaction is complex and multi-faceted; the tumour is spatially constrained by the surrounding tissue and its growth is limited by the diffusion of nutrients and oxygen, and, conversely, the microenvironment is modified by tumour secretions. Therefore, mathematical modelling is often necessary to describe and understand the interplay between the tumour and its microenvironment. For instance, many cancer cells upregulate the fermentation of glucose, even when sufficient oxygen is available for respiration [137, 138]. This creates lactic acid as a waste product, which is released into the environment. A reaction–diffusion model that represents cells by continuous cell densities was used to hypothesise that this creates an acidic gradient that degrades the surrounding host tissue, and contributes to malignancy [139]. Such a continuum model assumes that cell populations are large enough to be modelled by continuous densities. However, in the early stages of tumour formation cancer cells are few and this assumption does not hold. For these conditions, a hybrid model was developed where cells are instead simulated using cellular automata [140]. Similar hybrid models have been used to study the role of cell adhesion in tumour invasion [141], the evolutionary dynamics of glucose fermentation [142], breast cancer [143], and many more [144].

The evolutionary dynamics of cancer present major challenges for cancer treatment [145]. In particular, tumour cells often develop resistance to chemotherapeutic drugs, making it difficult or impossible to eradicate the tumour entirely. Even worse, aggressive

chemotherapeutic treatment can promote the proliferation of resistant cells by selectively eliminating drug-sensitive cells and thereby freeing up space and resources for resistant cells. The aim in "adaptive therapy", on the other hand, is to delay or avoid resistance by maintaining a population of sensitive cells in the tumour that compete with resistant cells [146]. Since resistance often comes at a fitness cost, sensitive cells can outcompete resistant cells in the absence of chemotherapeutic drugs. Therefore, in adaptive therapy the timing and dosage of drugs is dynamically adjusted to restrain overall tumour burden in a sustainable manner. The success of such an approach stands or falls with the ability to accurately simulate the dynamics of tumours for different dosing regimens. Mathematical modelling therefore plays a pivotal role in the development and application of adaptive therapy. For instance, a discrete-time model simulating the dynamics of sensitive and resistant subpopulations was used to investigate the impact of resistance costs [147]. The study demonstrates that adaptive therapy is most effective when resistance costs are high and resistant cells make up only a small fraction of the tumour. Similarly, simulations of cell-based models showed that competition for space inside the tumour is a key factor in suppressing the growth of resistant cells [148, 149]. Most recently, Lotka–Volterra models have been applied to examine the impact of tumour cell turnover in adaptive therapy [150].

1.4.5 Cell competition models

Nishikawa et al. present a Lotka–Volterra model for cell competition in epithelia based on short-range interactions [151]. The starting point is a competition model identical to Equations (1.4a) and (1.4b). In order to model the active elimination of loser cells by winner cells at the boundary of a loser clone in a background of winner cells, the authors introduce a predator–prey interaction term that models winner-induced loser cell death. Even though the model is not explicitly spatial, the clone geometry is reflected in the interaction term, which is defined to be proportional to the clone boundary length. For instance, if y is the loser cell count and the clone is perfectly circular, then the predation term is proportional to $y^{1/2}$. In addition, the authors construct an off-lattice, cell-based model implementing the same interactions. Using these models, the authors show that i) an active mechanism of loser elimination, such as a predator–prey interaction between winners and losers, is sufficient to achieve the outcome of cell competition, and ii) a difference in proliferation rates is by itself not sufficient to achieve cell competition. More recently, this work was extended to investigate the effect of cell death-induced proliferation on cell competition [152].

Shraiman derived a mechanical model that relates the pressure in an epithelial tissue to non-uniform growth [60]. Crucially, local buildup of excess pressure, which expresses the

amount of relative compression or stretching, results from a difference between the local growth rate and the average growth rate of the tissue. Additionally, Shraiman proposed a mechanical feedback mechanism where positive pressure (compression) causes cells to divide more slowly and vice versa for negative pressure (tension). If the absolute pressure becomes too extreme, however, the local growth rate turns negative and apoptosis is induced. Equipped with such a feedback mechanism, Shraiman showed that a mutant clone with a higher homeostatic pressure initially grows at a faster rate than the surrounding tissue, but adapts as local pressure accumulates, which dampens mutant growth and thus equalises its growth rate. However, wild-type cells surrounding the mutant clone may initiate apoptosis if the pressure exceeds the maximally tolerated pressure. In summary, the model shows that mechanical feedback can regulate tissue growth, and that differences in cellular responses to pressure can contribute to mechanical cell competition [60].

In addition to compression inside a fast-growing winner clone, a differential proliferation rate also results in the elongation of loser cells located along the boundary of the fast-growing clone [20]. Therefore, the shapes of loser cells at the boundary become anisotropic, with longer cell–cell junctions along the boundary and shorter ones perpendicular to the boundary. Tsuboi et al. show that this anisotropy is key to understanding how winner clones expand into the surrounding loser territory [153]. In particular, since junctional remodelling (see the T1 transition in Section 2.1.4) is more probable for shorter junctions, the topological rearrangements occurring in apoptotic loser cells at the boundary favour the expansion of neighbouring winner cells. Moreover, using a vertex model, the authors are able to demonstrate that i) biased junction remodelling in the absence of differential proliferation rates is sufficient for winner clone expansion, and ii) randomising junction remodelling inhibits winner clone expansion [153]. Taken together, model simulations demonstrate the importance of biased junction remodelling in epithelial cell competition. Another use of vertex models in cell competition is presented by Lee et al. [154]. The authors investigate mechanical cell competition in a tissue that is genetically homogeneous, but whose cells nonetheless display phenotypic variation in their mechanical properties. This phenotypic variation either is inherited, or randomly sampled at every cell division. In the former case, evolution takes place in the tissue, increasing tissue fitness relative to the latter case. Hence, the authors suggest a mechanism by which cell competition contributes to tissue fitness [154]. Both of these studies [153, 154] demonstrate the utility of vertex models in investigating mechanical aspects of cell competition in epithelia that involve cell-scale processes.

Generally speaking, it is becoming increasingly important to characterise interactions at the cellular level to elucidate the mechanisms of cell competition [155]. This presents

two challenges. On the one hand, it requires conducting cell competition experiments under various conditions while tracking many cells simultaneously and collecting data on their surface area, neighbourhood connectivity, cell cycle, and more. On the other hand, we need to construct cell-based models at a similar level of resolution to assimilate the data and reproduce the observed behaviour. Bove et al. tackle the first challenge by developing an experimental pipeline for mechanical cell competition experiments that combines "automated long-term microscopy and advanced image analysis" [156]. This enables them to track individual cell lineages over time and space, complete with information on their cell cycle and neighbourhood composition. In particular, the authors derive empirical relationships between apoptosis and division rates on the one hand, and cell density on the other hand. Gradeci et al. use these data to construct cellular Potts models of both mechanical and biochemical cell competition [78]. Through simulations, the authors conclude that tissue organisation strongly influences the outcome of biochemical cell competition, while differences in homeostatic density is sufficient for mechanical cell competition.

1.5 Concluding remarks

In this chapter, we introduced cell competition by first discussing competition more broadly and then addressing the seemingly paradoxical nature of cells belonging to the same multicellular organism competing with each other. In Section 1.1, we identified the need for a mechanistic model for winner/loser status acquisition, something which is currently missing from the literature, thus motivating the aim of this thesis to study the emergence of winner/loser status using mathematical modelling. We also summarised our strategy for achieving this aim with an outline of the thesis. In the remaining sections, we provided an overview of cell competition (Section 1.2), cell-based models (Section 1.3), and competition modelling (Section 1.4). The latter includes a discussion of current cell competition models. In the next chapter, we investigate whether varying the mechanical parameters in a heterotypic population is sufficient for inducing cell competition in a vertex model.

Chapter 2

Mechanical cell competition in epithelia

In Section 1.3, we argued that the vertex model is the most appropriate cell-based model for studying cell competition in the *Drosophila* wing disc because it has been shown to reproduce the dynamics of epithelial tissues in a variety of developmental processes [85]. In this chapter, we construct a vertex model for cell competition in Section 2.1. We then discuss its implementation using the open-source simulation package Chaste in Section 2.2. Finally, we describe a computational study of mechanical cell competition in Section 2.3.

2.1 Vertex model

Epithelial tissues are composed of cells that adhere tightly to form a continuous membrane. As a result, the apical surface of an epithelial tissue resembles a polygonal mesh where each polygon represents the apical membrane of an epithelial cell, each edge corresponds to a cell–cell junctional interface, and vertices mark the locations where multiple cell–cell interfaces meet (see Figure 2.1).

As mentioned in Section 1.3, vertex models represent epithelial tissues explicitly as polygonal meshes. The cell dynamics are determined by the equations of motion of the vertices. In Section 1.3.2, we explained that inertial forces can be neglected such that the equations of motion take on the form given in Equation (1.1).

There are multiple ways to define vertex forces [157], but the approach we use here is to associate the mesh with a phenomenological energy function E that models mechanical cell properties such as cell elasticity and cell–cell adhesion. The total force acting on vertex i is then defined as

$$\mathbf{F}_i = \nabla_i E \,, \tag{2.1}$$

where ∇_i is the gradient of E with respect to the spatial coordinates of vertex i. In our model,

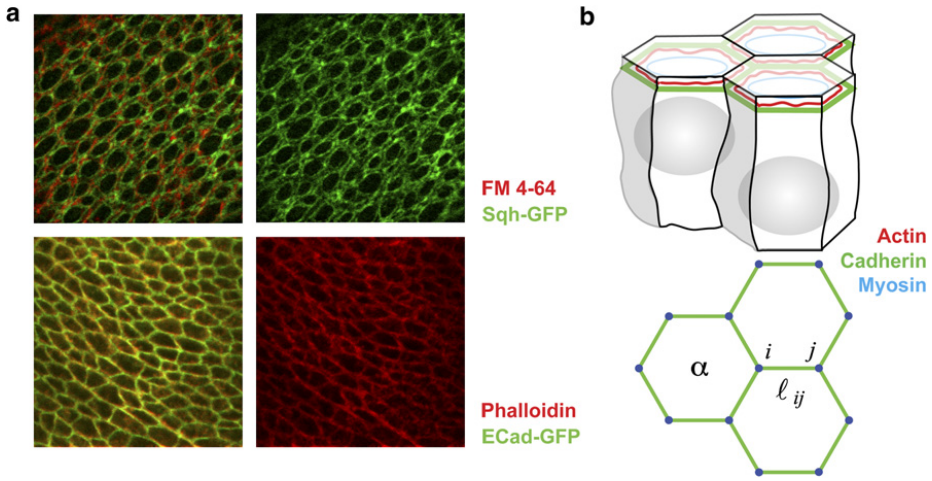

Figure 2.1: Epithelial tissues can be described by a two-dimensional network of cell–cell junctions. (a) Fluorescent microscopy image of an epithelial tissue showing the localisation of adhesive and contractile proteins. (b) The structure and dynamics of epithelial sheets can be approximated by a network of cell–cell junctions, which is modelled as a two-dimensional polygonal mesh in the vertex model. We use the symbol α to denote cells, and i or j to denote vertices. Furthermore, the length of the edge connecting i and j is written as ℓ_{ij}. Figure reproduced[1] from [84].

we assume that there are three important mechanical properties at work; cell elasticity, cell contractility, and cell–cell adhesion. Hence, the energy function can be written as

$$E = E_{\text{elasticity}} + E_{\text{contractility}} + E_{\text{adhesion}} \,, \tag{2.2}$$

where $E_{\text{elasticity}}$, $E_{\text{contractility}}$, and E_{adhesion} represent the energetic contributions of cell elasticity, cell contractility, and cell–cell adhesion, respectively. We describe each of these terms in detail below.

In addition to the vertex dynamics defined by Equations (1.1), (2.1), and (2.2), vertex models include discrete operations that modify the topology of the mesh. For instance, a "T2 transition" extrudes a cell whose cell area has fallen below a critical threshold by collapsing the cell into a vertex. Mesh rearrangements are discussed in Section 2.1.4. We also include cell proliferation in our model. Section 2.1.4 briefly discusses the implementation of cell divisions in the vertex model, and Section 2.1.5 describes the cell cycle model. Finally, in order to model cell competition, we assume that there are two cell types, denoted A and B. Cells belonging to the same cell type have identical parameters, but parameters are allowed to vary between cell types.

[1]With permission of Elsevier Science & Technology Journals; permission conveyed through Copyright Clearance Center, Inc.

2.1.1 Elasticity

The elasticity of a cell expresses its resistance against deformation; a cell that is being compressed or stretched resists this deformation by producing an opposing force. A common assumption in vertex models is that the cell α attempts to maintain a constant target cell area, denoted by S_α^0, and that the energy term for each individual cell is proportional to the squared deviation of the cell area from the target cell area. Summing over all cells, we obtain the following expression:

$$E_{\text{elasticity}} = \sum_\alpha \frac{K_\alpha}{2} \left(S_\alpha - S_\alpha^0\right)^2 , \tag{2.3}$$

where K_α and S_α are the elasticity parameter and cell area of cell α, respectively.

There are thus two parameters involved in cell elasticity, K_α and S_α^0. As mentioned before, these parameters vary between cell types. Thus, for cell α, we write

$$K_\alpha = \begin{cases} K_A & \text{for } \alpha \in A \\ K_B & \text{for } \alpha \in B \end{cases} , \tag{2.4}$$

and

$$S_\alpha^0 = \begin{cases} S_A^0 & \text{for } \alpha \in A \\ S_B^0 & \text{for } \alpha \in B \end{cases} . \tag{2.5}$$

2.1.2 Contractility

Cells in epithelial tissues tend to minimise the length of their perimeters. This is captured by the cell contractility energy term. Cell contractility penalises elongated cell shapes in favour of circular cell shapes. It is generally assumed that the contractility energy term is proportional to the square of the cell perimeter [84]:

$$E_{\text{contractility}} = \sum_\alpha \frac{\Gamma_\alpha}{2} L_\alpha^2 , \tag{2.6}$$

where Γ_α and L_α are the contractility parameter and the cell perimeter of cell α, respectively. The constant parameter Γ_α depends on the cell type as

$$\Gamma_\alpha = \begin{cases} \Gamma_A & \text{for } \alpha \in A \\ \Gamma_B & \text{for } \alpha \in B \end{cases} . \tag{2.7}$$

2.1.3 Adhesion

The final term in Equation (2.2) represents the line tension acting on cell–cell interfaces. This line tension models adhesive forces at cell–cell junctions and can be positive or negative. When the interaction between two cells is favourable, meaning that they adhere strongly to each other and prefer a large cell–cell interface, the line tension is low and vice versa. We denote the edge between vertices i and j as $\langle i, j\rangle$ and its edge length as ℓ_{ij} (see Figure 2.1). Furthermore, we assume that the line tension is proportional to edge length. We sum over all edges to obtain

$$E_{\text{adhesion}} = \sum_{\langle i,j\rangle} \Lambda_{ij}\ell_{ij}\,, \tag{2.8}$$

where Λ_{ij} is the line tension parameter. Unlike the other parameters discussed here, the line tension parameter can be negative.

Since the line tension parameter is dependent on the edge type, rather than the cell type, we need to specify values for every pairing of cell types. In addition, we need to account for edges at the boundary of the tissue, which border a cell on one side and empty space on the other. For two cell types, we thus specify a total of five line tension parameters. Denoting the two cells sharing the edge $\langle i, j\rangle$ as α and β, we write

$$\Lambda_{ij} = \begin{cases} \Lambda_{AA} & \text{for } \alpha,\beta \in A \\ \Lambda_{BB} & \text{for } \alpha,\beta \in B \\ \Lambda_{AB} & \text{for } \alpha \in A,\ \beta \in B \\ \Lambda_{A} & \text{for } \alpha \in A,\ \beta \in \emptyset \\ \Lambda_{B} & \text{for } \alpha \in B,\ \beta \in \emptyset \end{cases}, \tag{2.9}$$

where $\beta \in \emptyset$ signifies that $\langle i, j\rangle$ is a boundary edge.

2.1.4 Mesh rearrangements

The purpose of mesh rearrangements in the vertex model is twofold. Firstly, the energy formulation of the vertex dynamics does not prevent vertices from crossing cell boundaries. This has the potential to create meshes with overlapping cells, therefore invalidating the mesh. Mesh rearrangements handle these situations in a manner that restores the validity of the mesh. From a modelling perspective on the other hand, mesh rearrangements enable cells to break and form bonds so that they are able to swap neighbours and are not confined to a static topology. Moreover, it is not possible to create new cells or remove existing cells without restructuring the mesh. There are many technical details involved with mesh

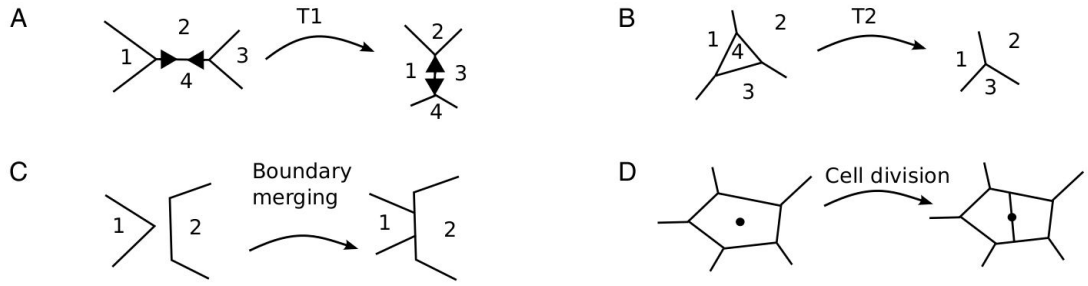

Figure 2.2: Overview of mesh rearrangements. (A) T1 transition. (B) T2 transition. (C) T3 transition (also known as a boundary merge). (D) Cell division. Figure reproduced[2] from [158].

rearrangements, so here we discuss only the most important mesh rearrangements and refer the reader to [157] for technical details.

For the first type of mesh rearrangement, consider the edge shared between cells 2 and 4 in Figure 2.2.A. All the energy contributions in Equation (2.2) are finite, hence it is possible that the net forces on the edge vertices cause them to converge until they meet each other, creating overlapping cells. To prevent cells from overlapping in this scenario, the mesh undergoes a T1 transition when the edge length falls below the threshold distance ℓ_{min} [83]. In order to prevent a second T1 transition from occurring again immediately afterwards, the newly formed edge has a slightly longer length of $\rho\ell_{\text{min}}$, where $\rho > 1$ is the separation ratio. As Figure 2.2.A shows, cells 2 and 4 are no longer neighbours after the transition. This scenario is more likely when cells 2 and 4 have a large line tension parameter, corresponding to weak cell–cell adhesion.

A second type of mesh rearrangement, called the T2 transition, occurs when the area of a triangular cell becomes smaller than the threshold area S_{min}. In the T2 transition, the cell is removed from the mesh and replaced by a single vertex (see Figure 2.2.C). Cells lose neighbours through T1 transitions as they decrease in size, so it is sufficient to only consider the removal of triangular cells. A T3 transition, or boundary merge, occurs when two boundary cells intersect (see Figure 2.2.C).

The final type of mesh rearrangement discussed here is cell division. Whereas T1, T2, and T3 transitions are necessary to avoid mesh invalidation, this operation is implemented to model cell proliferation. During cell division, the cell is split into two daughter cells of equal size by a new edge that intersects the cell's centroid, creating two new vertices on its perimeter. There are multiple choices for determining the orientation of this edge [157]. We chose the shortest axis as the dividing line[3].

[2]Under CC BY license (`https://creativecommons.org/licenses/by/4.0/`).

[3]If the cell is shaped as a regular polygon, we choose a random axis of division.

2.1.5 Cell proliferation

We implement a cell cycle with two phases: a stochastic G1 phase and a deterministic G2 phase. For a cell α, the duration of the G1 phase is exponentially distributed with the mean given by $t_{\mathrm{G1},\alpha}$, and the G2 phase lasts for the fixed duration $t_{\mathrm{G2},\alpha}$. This two-phase cell cycle model is motivated by experiments with *in vitro* cell cultures [159].

Furthermore, we assume that cell cycles are uncorrelated at the start of the simulation. We achieve this by sampling the birth times of the initial cell population uniformly from the interval $[-t_{\mathrm{G1},\alpha} - t_{\mathrm{G2},\alpha}, 0]$ where $t = 0$ is the start of the simulation. The purpose of this randomisation is to avoid the artificial synchronisation of cell divisions that would occur if all cells were initialised to the same stage of the cell cycle at the start of the simulation.

We assume two distinct cell types, so we write

$$t_{\mathrm{G1},\alpha} = \begin{cases} t_{\mathrm{G1},A} & \text{for } \alpha \in A \\ t_{\mathrm{G1},B} & \text{for } \alpha \in B \end{cases}, \tag{2.10}$$

and

$$t_{\mathrm{G2},\alpha} = \begin{cases} t_{\mathrm{G2},A} & \text{for } \alpha \in A \\ t_{\mathrm{G2},B} & \text{for } \alpha \in B \end{cases}. \tag{2.11}$$

Moreover, each division results in two daughter cells that are of the same type as the mother cell, i.e. a cell of type A divides into two daughter cells of type A.

2.2 Chaste implementation

Now that we have discussed the heterotypic vertex model, we discuss its computational implementation. The main challenge in implementing a vertex model is its complexity. Programming a mesh that is continuously being modified, not only by vertex dynamics, but also by discrete mesh rearrangements, is a task that takes considerable effort and is prone to errors. Moreover, because the implementation of a vertex model involves many technical details, even two implementations with the same underlying mathematical formulation are unlikely to be identical if they are developed independently.

We therefore implemented the heterotypic vertex model within Chaste, an open-source simulation package for computational physiology and biology [160] that includes a range of cell-based models [71]. Chaste offers a number of advantages that address the issues mentioned above. Firstly, because Chaste is an open-source project, every implementation detail is publicly available. As a result, the scientific community can review every line of

code that went into a simulation, ensuring that simulation results are fully reproducible.

Secondly, Chaste exposes a large library of C++ classes for cell-based models that researchers can build on or adapt to their specific application. This eliminates the need for "reinventing the wheel", and saves time and effort on behalf of computational scientists. Moreover, the development of Chaste is centred on a test-driven framework, which means that every piece of functionality coded in Chaste must be accompanied by a suite of unit tests verifying that it behaves as expected [161], minimising the chance of software bugs going undetected.

The Chaste project is thoroughly documented on its website[4], so we will not discuss its software architecture here. Instead, we focus on the specific contributions that we made to implement the heterotypic vertex model. The code implementing the heterotypic vertex model can be found in the following GitHub repository: `https://github.com/ThomasPak/cell-competition`.

2.2.1 `FarhadifarDifferentialForce` class

The Chaste project is developed using an object-oriented programming (OOP) framework. The calculation of vertex forces, Equation (2.1), is encapsulated by the `AbstractForce` class. This is an abstract class that declares the function `AddForceContribution()`, which takes a cell population, iterates over the vertices of the associated mesh, and computes the forces applied to each vertex.

Chaste comes with a number of vertex force models included in its library, including the `NagaiHondaForce` and `FarhadifarForce`, which implement the vertex models introduced in [83] and [84], respectively. The `FarhadifarForce` energy formulation is identical in form to the one we defined in Section 2.1. However, the class assumes that every cell has the same mechanical properties. Hence, the mechanical parameters are assumed equal for all cells in the population.

To implement force calculations in the heterotypic vertex model, we created the class `FarhadifarDifferentialForce`, which takes into account cell types when computing vertex forces. Furthermore, in keeping with Chaste's test-driven development framework, we wrote unit tests verifying that the class behaves as expected.

2.2.2 Cell labelling

In Chaste, cells are represented by objects that own a collection of "cell properties". These cell properties track the states of each individual cell, such as its mutation state and pro-

[4] `https://chaste.cs.ox.ac.uk/`

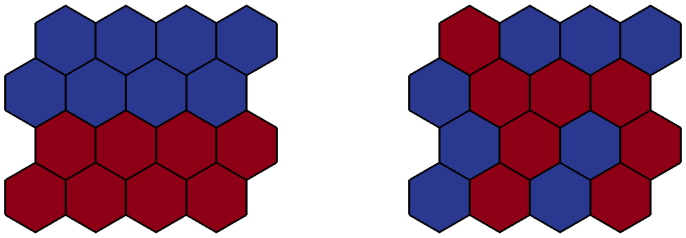

Figure 2.3: Initial spatial configuration of the vertex model. Cells are initially arranged in a honeycomb pattern. A-type cells are coloured blue and B-type cells are coloured red. **Left**: segregated initial pattern. **Right**: random initial pattern.

liferation type (e.g. whether it is a stem cell or a differentiated cell). The components of a cell-based simulation can access and modify these cell properties. In addition, Chaste supports writing out cell properties so that they can be plotted when visualising the results (see Figure 2.3 for an example).

We distinguish between cell types A and B by adding the `CellLabel` cell property, which was already implemented in Chaste, only to B-type cells at the start of the simulation. As the `FarhadifarDifferentialForce` class iterates over vertices, it uses this cell property to identify the cell types involved in each force calculation and chooses mechanical parameters accordingly.

At the start of the simulation, the cells are arranged in a honeycomb pattern (see Figure 2.3). We denote the initial number of cells as $N^0 = N_A^0 + N_B^0$. Since we want to explore the effects of spatial segregation and mixing, we implemented two different methods for assigning cell type identities. For the "segregated" pattern, we order the cells from left to right, bottom to top, and label the first N_B^0 cells as B-type cells (see Figure 2.3). For the "random" pattern, we generate a random permutation of cell types.

2.2.3 Simulator executable

To simplify running large numbers of simulations, we packaged the simulation code as a "simulator" executable with a text-based interface. Concretely, this means that parameters are specified using command-line arguments or through the standard input. Moreover, all the parameters used by the simulator are written to standard output before running the simulation, so that the user can verify that the parameters were specified correctly.

Table 2.1: Default, low and high parameter values for parameter sweep.

Parameter	Default value	Low value	High value
S_A^0	1	0.9	1.1
K_A	1	0.9	1.1
Γ_A	0.04	0.03	0.05
Λ_{AA}	0.12	0.10	0.14
Λ_{AB}	0.12	0.10	0.14
$t_{\mathrm{G1},A}$	30	20	40
$t_{\mathrm{G2},A}$	70	60	80

2.2.4 Post-processing

We implemented a number of C++ classes for writing data to text files during simulation. In addition, we wrote Python scripts to process the data files for all the individual simulations in a simulation suite and collect summary statistics in comma-separated value files.

2.3 Computational study of mechanical cell competition

In the previous sections, we introduced the heterotypic vertex model and discussed its numerical implementation. In this section, we perform a systematic parameter sweep to investigate the behaviour of the model for different parameters to see if we find evidence for cell competition. In particular, we vary the parameters of cell type A and run heterotypic simulations with 18 initial A-type cells and 18 initial B-type cells. In addition, for every heterotypic simulation, we also run a homotypic simulation of cell type A with the same parameters to verify the intrinsic viability of cell type A.

2.3.1 Methodology

Parameter choice We vary the parameters of cell type A as listed in Table 2.1. For every model parameter, we pick three values; one equal to the default value[5], one less than the default value, and one greater than the default value. These values are given in Table 2.1. We then take the Cartesian product over all of the parameters to generate the parameter sweep. In addition, we run every parameter set with segregated and random initial conditions. The parameters for cell type B are constant and set to the default values given in Table 2.1.

[5]Chaste specifies a default value for all model and simulation parameters.

Any remaining simulation parameters are set to the default Chaste values, with the exception of the simulation timestep (set to 0.05 hours), the simulation time (set to 250 hours), the threshold distance d_{T1} (set to[6] 0.1), and the initial number of cells (set to 36 cells). In addition, each simulation was given a distinct seed for generating random numbers.

After taking the Cartesian product of the parameter values in Table 2.1, we obtain 3^7 unique parameter sets. We also run every parameter set with segregated and random initial conditions, and perform a homotypic simulation for every heterotypic simulation. Therefore, the total number of simulations is $2 \times 2 \times 3^7 = 8\,748$.

Computational materials and methods Since a typical simulation of the vertex model takes on the order of 2.5 minutes, running all 8 748 simulations sequentially would take roughly two weeks. This is impractical, so we parallelised our computational workload using the Advanced Research Computing (ARC)[7] service, which provides high-performance computing resources and training at the University of Oxford. ARC uses the open-source job scheduler Slurm Workload Manager[8]. We used Slurm's job array functionality to submit the entire parameter sweep using a single shell script. We ran our computations on ARC's throughput cluster, specialised for high throughput of single-core jobs such as our vertex simulations. The computational nodes in this cluster contain a mix of Broadwell, Haswell, and Cascade Lake CPUs. All vertex simulation suites in this thesis were executed in this manner.

Data processing We summarise the proliferative behaviour of a cell type in a simulation using the **survival frequency**. This is the number of division events divided by the sum of division and death events. For a growing population, there are more division than death events, hence the survival frequency is greater than $1/2$ and vice versa for a declining population. We could have chosen other summary statistics to discriminate between population growth and decline, but we chose the survival frequency because it will become relevant in later chapters.

For each homotypic simulation, we computed the homotypic survival frequency, denoted by $\hat{\lambda}$, as

$$\hat{\lambda} = \frac{\#\text{ division events}}{\#\text{ division events} + \#\text{ death events}}\,. \tag{2.12}$$

[6] We do not specify units here because length is dimensionless in cell-based Chaste (see `https://chaste.cs.ox.ac.uk/trac/wiki/ChasteGuides/ChasteUnits`).

[7] `https://www.arc.ox.ac.uk/`

[8] `https://slurm.schedmd.com/`

For each heterotypic simulation, we computed the heterotypic survival frequencies as

$$\hat{\xi}_{A|B} = \frac{\text{\# A division events}}{\text{\# A division events + \# A death events}}, \tag{2.13}$$

$$\hat{\xi}_{B|A} = \frac{\text{\# B division events}}{\text{\# B division events + \# B death events}}, \tag{2.14}$$

for cell types A and B, respectively.

In addition, we define the heterotypic survival difference as

$$\hat{\Delta}^{\neq}_{A|B} = \hat{\xi}_{A|B} - \hat{\xi}_{B|A}\,. \tag{2.15}$$

When the average cell cycle length is the same for both cell types, this quantity indicates which cell type is at a proliferative advantage; for positive heterotypic survival differences, cell type A is fitter, and vice versa for negative heterotypic survival differences. When average cell cycle lengths are different, it is not obvious which cell type proliferates faster; for example, a cell type may compensate for a lower survival frequency with a shorter cell cycle length. However, in the context of cell competition, we are more interested in discriminating between population growth and decline than relative proliferation rates. As discussed above, survival frequencies are more suitable for this purpose. Hence, we use the heterotypic survival difference as an indicator of relative proliferative advantage, even in the case of unequal cell cycle lengths.

In addition, we want to compare the heterotypic survival frequency to the homotypic survival frequency. This quantifies the effect that the heterotypic interaction has on a cell type relative to homotypic conditions. For each pair of homotypic and heterotypic simulations with the same parameters, we define the homotypic survival difference for cell type A as

$$\hat{\Delta}^{=}_{A|B} = \hat{\xi}_{A|B} - \hat{\lambda}_{A}\,. \tag{2.16}$$

The homotypic simulation suite for cell type A contains two simulations where the homotypic population has the same parameters as cell type B (the default values), one with segregated initial conditions and one with random initial conditions. Hence, we can consider these simulations as homotypic simulations of cell type B, and we use them to estimate the homotypic survival frequency of cell type B. For a homotypic simulation, the random and segregated initial conditions are equivalent, so we average the homotypic survival frequency over these two simulations to estimate the homotypic survival frequency of cell type B, which we denote as $\overline{\lambda}_B$. Hence, for each heterotypic simulation, we compute the homotypic

survival difference for cell type B as

$$\hat{\Delta}^{=}_{B|A} = \hat{\xi}_{B|A} - \overline{\lambda}_B \,. \tag{2.17}$$

We used the open-source Python package pandas[9] to process and analyse the data.

Data visualisation We used the open-source Python packages Matplotlib[10] [162] for plotting histograms and Seaborn[11] [163] for plotting heat maps.

2.3.2 Results

Since the survival frequencies indicate whether a population has grown or declined, we can use them to check whether cell competition has taken place. We define a cell type as viable in a simulation if its survival frequency is greater than or equal to a half, since that signifies that the population size has grown or stayed the same, respectively. If the survival frequency is less than a half, then the population has declined and we count the cell type as nonviable.

We found that cell type B is viable for all heterotypic simulations and homotypic simulations. We therefore did not observe any cases where cell type B is outcompeted by cell type A. For cell type A, we applied the viability criteria for each parameter set to the homotypic and heterotypic survival frequencies to distinguish between four cases:

- $\hat{\lambda}_A < 1/2$, $\hat{\xi}_{A|B} < 1/2$: cell type A is nonviable in homotypic and heterotypic conditions.
- $\hat{\lambda}_A < 1/2$, $\hat{\xi}_{A|B} \geq 1/2$: cell type A is nonviable in homotypic conditions, but is viable in heterotypic conditions. This implies that cell type A fares better in heterotypic conditions than in homotypic conditions.
- $\hat{\lambda}_A \geq 1/2$, $\hat{\xi}_{A|B} < 1/2$: cell type A is viable in homotypic conditions, but nonviable in heterotypic conditions. These are the conditions for cell competition.
- $\hat{\lambda}_A \geq 1/2$, $\hat{\xi}_{A|B} \geq 1/2$: cell type A is viable in homotypic and heterotypic conditions.

For each unique parameter set, we checked whether the corresponding homotypic and heterotypic simulations satisfied these criteria, and listed the number of parameter sets satisfying each case in Table 2.2. For the overwhelming majority of parameter sets, cell type A was viable in homotypic and heterotypic conditions. Only seven out of 4 374 parameter

[9] `https://pandas.pydata.org/`
[10] `https://matplotlib.org/`
[11] `https://seaborn.pydata.org/`

Table 2.2: Count of simulations satisfying the viability criteria.

	$\hat{\lambda}_A < 1/2$	$\hat{\lambda}_A \geq 1/2$
$\hat{\xi}_{A\|B} < 1/2$	84	7
$\hat{\xi}_{A\|B} \geq 1/2$	58	4 225

sets, approximately 0.16%, satisfy the criteria for cell competition. This indicates that cell competition is very rare among the parameter sets that we simulated. In fact, there are many more cases of the opposite scenario, where cell type A is nonviable in homotypic conditions and becomes viable in heterotypic conditions.

In order to check whether those seven parameter sets truly represent cell competition, we ran an additional simulation suite targeting those seven parameter sets in particular. Specifically, we ran 20 pairs of simulations (one homotypic simulation and one heterotypic simulation, with random initial conditions in each pair) for each parameter set with different random seeds for the random number generator. The total number of simulations in the targeted simulation suite is $2 \times 20 \times 7 = 280$. We then applied the viability criteria and counted the number of pairs of simulations that match the cell competition criteria.

Among the seven parameter sets, there was one parameter set with three simulation pairs out of 20 that satisfied the cell competition criteria. For the remaining parameter sets, we only found one or two matching simulation pairs. Hence, the targeted simulation suite yielded only a small number of positive results per parameter set. This strongly suggests that the matches we encountered in the parameter sweep were statistical flukes rather than genuine indicators of competitive behaviour.

We bolster this claim statistically by treating each parameter set independently as a Bernoulli process. In this view, a pair of simulations for a given parameter set is a Bernoulli trial with probability q of matching the cell competition criteria. If a parameter set is more likely to match the cell competition criteria than not, i.e. $q > 1/2$, then we say that the parameter set gives rise to cell competition. In order to test a given parameter set for competitive behaviour, we discriminate between the following hypotheses:

$$\mathbf{H_0} : q \leq \frac{1}{2}, \tag{2.18}$$

$$\mathbf{H_1} : q > \frac{1}{2}. \tag{2.19}$$

In other words, the null hypothesis is that the parameter set does not give rise to cell competition, i.e. $q \leq 1/2$, and the alternative hypothesis is that it does, i.e. $q > 1/2$.

We decide whether or not to reject the null hypothesis by calculating how likely it is to obtain results that are at least as extreme as the observed results under the null hypothesis. This probability is also known as the p-value. By "extreme" we mean results that support the alternative hypothesis rather than the null hypothesis. In this case, the more simulation pairs that match the cell competition criteria, the more extreme the observed data are. For example, observing ten matches provides stronger evidence in favour of cell competition than three matches. A small p-value indicates that the observed results are unlikely to be obtained under the null hypothesis, supporting the alternative hypothesis, and vice versa for a large p-value.

The largest number of matches among the seven selected parameter sets was 3 out of 20 simulation pairs. Under the null hypothesis, the number of matches obeys a binomial distribution with 20 trials and $q \leq 1/2$. However, we only need to compute the p-value for the largest value of q that the null hypothesis permits, i.e. $q = 1/2$, because this corresponds to the null distribution that maximises the p-value. Hence, rejecting the null hypothesis for $q = 1/2$ would allow us to reject the null hypothesis for all $q \leq 1/2$. Concretely, we find that

$$p\text{-value} = \sum_{k=3}^{20} \binom{20}{k} \left(\frac{1}{2}\right)^k \left(1 - \frac{1}{2}\right)^{20-k} \approx 0.9998\,, \tag{2.20}$$

where we summed over $k = 3, \ldots, 20$ to compute the probability that we encounter *at least* three matches. The remaining parameter sets had fewer matches, and therefore Equation (2.20) is a lower bound on the p-value across all seven parameter sets. As Equation (2.20) indicates, the computed p-value is close to unity, so we do **not** reject the null hypothesis. In other words, there is not enough evidence in the data to support competitive behaviour.

These results suggest that cell competition is not taking place in the heterotypic populations, and that interactions occurring between cell types have little impact on their viability. Figure 2.4 shows the histograms of the homotypic and heterotypic survival frequencies and differences. The histograms for the homotypic survival differences show that the homotypic survival difference is very narrowly centred on zero, indicating that the cell populations do not experience significant changes in a heterotypic population as compared to homotypic conditions. This is further supported by the histograms for the homotypic and heterotypic survival frequencies of cell type A, which have roughly the same shape.

Finally, we examine the correlation between our summary statistics and the parameters that we varied in the parameter sweep. In Figure 2.5, we plotted the Pearson correlation coefficients using a heat map. We see that the homotypic survival frequency of cell type A correlates positively with the cell target area and the elasticity constant, and negatively

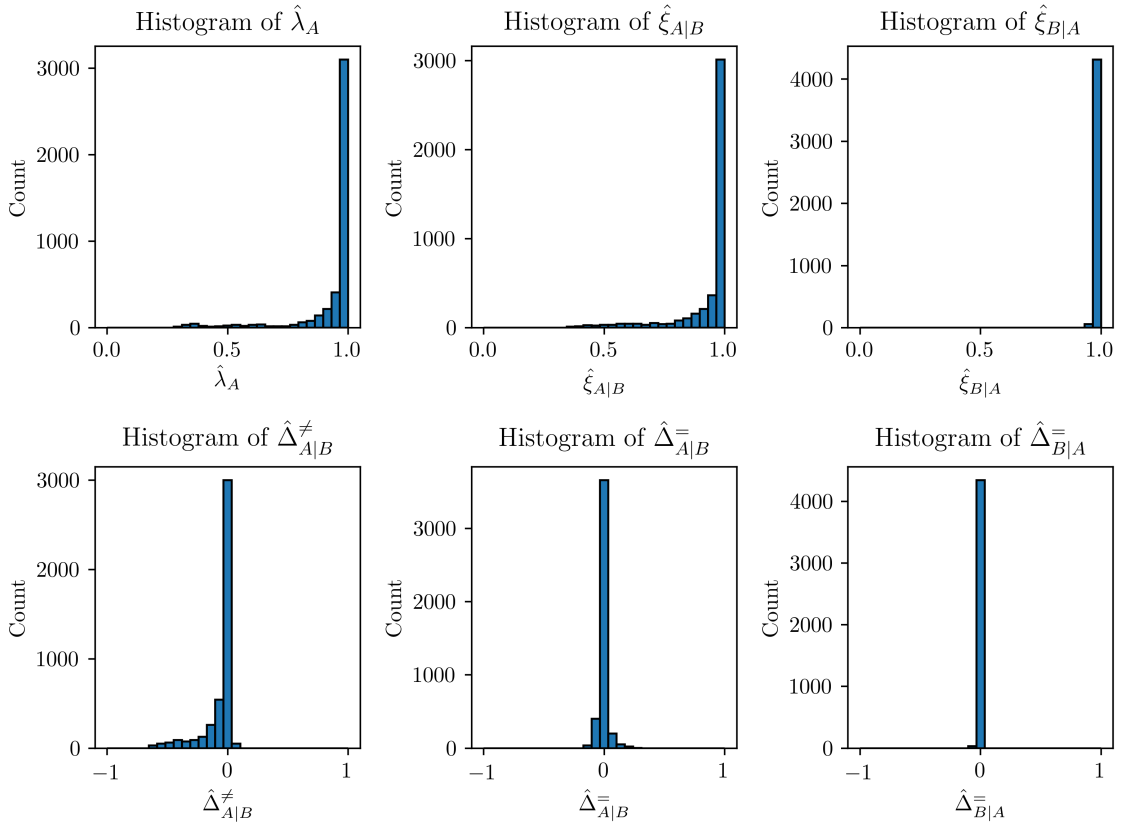

Figure 2.4: Histograms of the homotypic survival frequency $\hat{\lambda}_A$, defined in Equation (2.12), the heterotypic survival frequencies $\hat{\xi}_{A|B}$ and $\hat{\xi}_{B|A}$, defined in Equations (2.13) and (2.14), respectively, the homotypic survival difference $\hat{\Delta}^{\neq}_{A|B}$, defined in Equation (2.15), and the homotypic survival differences $\hat{\Delta}^{=}_{A|B}$ and $\hat{\Delta}^{=}_{B|A}$, defined in Equations (2.16) and (2.17), respectively.

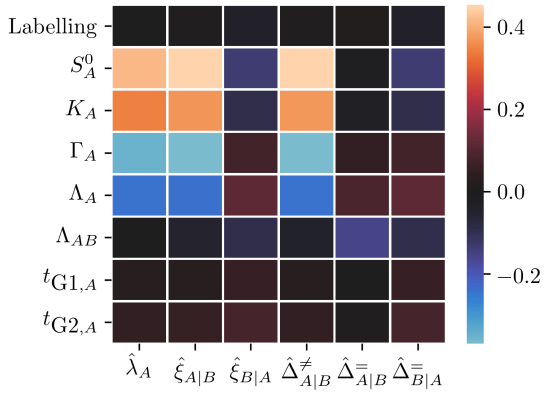

Figure 2.5: Heat map of Pearson correlation coefficients of summary statistics with respect to parameters for cell type A. The "labelling" parameter is a Boolean variable that indicates whether the initial conditions were random or segregated, for which the parameter takes on the values true and false, respectively.

with the contractility parameter and the line tension parameter. Moreover, we see the same correlation for the heterotypic survival frequency of cell type A and the heterotypic survival difference. This indicates that the model parameters influence the homotypic and heterotypic survival frequency in a similar way, suggesting that the viability of A-type cells is mostly determined by intrinsic factors, rather than by competitive interactions.

2.4 Discussion

In this chapter, we formulated the heterotypic vertex model, discussed its implementation, and performed a systematic parameter sweep to explore whether competitive behaviour can arise from differences in mechanical parameters. Out of 4 374 pairs of homotypic and heterotypic simulations, only seven parameter sets satisfied the criteria for cell competition. We ran more simulations of those seven parameter sets in particular, and found that we could not reliably reproduce competitive behaviour. Moreover, the empirical distributions of summary statistics, and the correlations between model parameters and summary statistics, also did not provide any evidence for cell competition.

Hence, we conclude that, for the parameter space investigated, simply varying the mechanical parameters of cell types in our heterotypic vertex model is not sufficient for cell competition. Indeed, experiments suggest that cell competition generally depends on an active mechanism of cell death, such as apoptosis [53], including in the case of mechanical cell competition [164]. Hence, mechanical interactions may lead to cell competition if paired with such a mechanism. We will briefly touch on this in the final chapter. In the next

chapter, we present a framework for implementing active cell death in cell competition.

Chapter 3

The death clock modelling framework

In Chapter 2, we learned that mechanical interactions by themselves are insufficient to drive cell competition. Even though we observed loser cell types being eliminated by winner cell types, homotypic simulations revealed that those loser cells were eliminated because they were intrinsically nonviable, and not because of competitive interactions. The definition of cell competition requires that both populations are viable in homotypic conditions, so we conclude that variation in mechanical properties is insufficient for cell competition.

This suggests that cell competition requires an active and non-autonomous mechanism of cell death. "Active" meaning that cells die by triggering apoptosis, a process which, once initiated, inexorably leads to cell death, and "non-autonomous" meaning that the decision to trigger apoptosis is influenced by signals produced by other cells. Indeed, experiments have established that apoptosis is a common feature of cell competition [53, 42, 164]. Hence, published models of cell competition feature active and non-autonomous mechanisms for apoptosis [78, 151, 152].

However, these mechanisms are based on predetermined winner/loser identities, with the rate of cell death in losers driven by heterotypic contacts with winner cells. Because the winner/loser identity is assumed, these models are fundamentally unable to describe *how* cells acquire a winner or loser status. In addition, experiments indicate that cell competition also occurs endogenously in homotypic populations [42, 57]. This suggests that the signals responsible for competition-induced cell death are not only being exchanged between winners on the one hand and losers on the other hand, but among all cells.

Therefore, because the aim of this thesis is to understand how winner/loser identities emerge, in this chapter we propose a framework that treats all cells as equal actors with the same ability to compete with other cells. Specifically, all cells share a common mechanism for initiating apoptosis and have the ability to receive and emit signals that may trigger apoptosis. Differences in cell fitness should only emerge from variations in parameters

between cell types and the interactions resulting from those variations, since the fundamental machinery is the same for both winners and losers. This is the core principle underlying the *death clock framework* developed as part of this thesis.

The rest of this chapter is structured as follows: in Section 3.1, we state the biological assumptions that motivate our modelling choices, and introduce the death clock framework. In brief, the cell senses "death signals" from the environment and accumulates them in an abstract quantity called the "death clock". Once the death clock reaches a threshold value, apoptosis is triggered. In Section 3.2, we derive an expression for the survival probability of a cell experiencing a known death signal. In Section 3.3, we explore the integration of the death clock framework in two cell-based models: the vertex model, first described in Chapter 2, and the well-mixed model, where every cell interacts equally with every other cell. Finally, we end the chapter with a summary in Section 3.4.

3.1 Death clock

3.1.1 Biological assumptions

Apoptosis is triggered by a threshold mechanism Apoptosis is one of the most dramatic events that a cell can undergo. It is therefore not a surprise that the decision to initiate apoptosis is controlled by a complex web of regulatory components [165, 166, 167]. In mammals, the two major apoptotic pathways are the intrinsic pathway and the extrinsic pathway. The intrinsic pathway is triggered by diverse stimuli associated with cell damage, such as hypoxia, irradiation, and growth factor deprivation. The extrinsic pathway, on the other hand, is mediated by cell surface death receptors that bind death ligands [167].

Complicating matters even further, the conditions and stimuli leading to apoptosis are dependent on the cell type [165]. For instance, some hormones stimulate apoptosis in some cell types, but have a neutral or inhibitory effect on other cell types. The many roads leading to apoptosis reflect the widespread use of cell death by the body in development, tissue homeostasis, and immune responses [168].

The complexity of the apoptotic pathway makes it challenging to come up with a model for apoptosis that is sufficiently representative of cell behaviour, yet simple enough to be useful. Moreover, the fact that the cell competition pathways upstream of apoptosis have not yet been identified means that we necessarily need to make some assumptions. Fortunately, because of previous work done on mathematical models of apoptosis [169, 170], we can make an educated guess.

In particular, a series of studies involving modelling and experiments have revealed

the importance of threshold mechanisms in the intrinsic pathway [171, 172] as well as the extrinsic pathway [173, 174]. For instance, it was shown that death ligand-induced apoptosis requires a threshold proportion of ligand to receptor numbers [173, 174], below which apoptosis does not occur. Given this precedent, we propose a model in which competition-induced apoptosis is triggered by the accumulation of some death signal, for example a ligand, above a threshold value.

Apoptosis is coupled to the cell cycle It has been well established in the literature that apoptosis and the cell cycle are closely coupled [175, 176, 177, 178]. This manifests in morphological similarities between apoptotic and mitotic cells [175], and common regulatory components [176, 178, 179]. Cell cycle checkpoints occur at key transitions in the cell cycle and allow the cell to review its internal state before committing to the next cell cycle phase. If the conditions are right, the cell cycle transitions to the next phase. However, if the conditions are not right, for example when the cell has accrued excessive DNA damage, cell cycle arrest or apoptosis may occur at the checkpoint [180].

Notably, the regulatory protein *Myc* is known to affect both cell cycle progression and apoptosis [181, 182, 183]. Specifically, *Myc* is necessary for the transition of G1 to S phase, and induces cell cycle progression in quiescent cells [181, 183]. On the other hand, *Myc* has been associated with increased rates of cell death [182]. Coupled with the fact that differential *Myc* expression results in cell competition (as discussed in Section 1.2.1), we will use our model to test the hypothesis that apoptosis, competition, and the cell cycle are interrelated.

Concretely, we assume that the cell is only susceptible to competition-induced apoptosis in G1 phase, and that from S phase onwards, the cell is committed to division. In reality, apoptosis may still occur in phases other than G1 for reasons unrelated to competition, for example DNA damage. However, we assume that those apoptosis events are negligible compared to competition-induced apoptosis in the context of cell competition and therefore omit them from our model.

Death signals can be coupled to the cell cycle So far, we have not made any specific assumptions regarding the form of death signals. The reason for this is that there are multiple competing hypotheses regarding the mode of intercellular communication in cell competition. These include short-range signalling at cell–cell junctions, long-range signalling through diffusible ligands, and mechanical stresses [184]. The reason for this diversity is that the nature of the signals exchanged to establish loser and winner status remains an open question, and may depend on the type of cell competition under consideration. Here, we highlight

some studies suggesting that death signals could be correlated with cell cycle progression.

One of the main features of cell competition is non-autonomous induction of apoptosis in loser cells by proliferating winner cells. Clues as to how the cell cycle is potentially involved in this process come from studies investigating the replacement of larval abdominal epidermis by adult epithelium in *Drosophila* [185]. During a process that resembles cell competition, larval epidermal cells (LECs) undergo apoptosis to make space for proliferating and migrating histoblasts (adult precursor cells). Similarly to cell competition, apoptotic LECs are mostly found at the boundary between LECs and histoblasts. Intriguingly, inhibiting cell cycle progression of histoblasts locally suppresses apoptosis in LECs [186, 187], which leads to three conclusions. Firstly, apoptosis in LECs is non-autonomous, i.e. caused by other cells (specifically histoblasts in this case). Secondly, the induction of apoptosis is a result of local interactions with histoblasts[1]. Thirdly, these local interactions, which can be considered death signals because of their effect, are coupled to the histoblast cell cycle. Since this process happens at the boundary between two distinct tissues and not within the same tissue, strictly speaking, it is not cell competition. However, it shares many characteristics with cell competition, so we hypothesise that death signals in the context of cell competition may similarly be coupled to the cell cycle.

In addition, we note that there is evidence for apoptosis itself being a death signal for surrounding cells [188]. In fact, it has been demonstrated that apoptotic cells are capable of inducing a wide array of behaviours in neighbouring cells, including proliferation [189, 190, 191, 192] and apoptosis [188], depending on the context [193]. Therefore, it appears that the process of apoptosis itself is not "silent". Instead, apoptotic cells are actively communicating with nearby cells [194]. In this thesis, we did not incorporate apoptosis as a death or proliferation signal, but our modelling framework can easily be extended to include such effects.

3.1.2 Modelling choices

As established in the previous section, we assume that the cell is only susceptible to competition-induced apoptosis during G1 phase. As a result, only two parts of the cell cycle are relevant to cell competition: G1 phase, and the rest of the cell cycle (composed of S, G2, and M phases). Therefore, we decided to model the cell cycle with just two phases, where the first phase corresponds to G1 phase and the second phase lumps together S, G2, and M phases. For convenience, we refer to the first and second phases as the G1 and G2 phases, respectively.

[1]Note that the study did not uncover *how* the local interactions were mediated, hence they could have been caused by direct cell–cell contact, or a signal that diffuses over small distances.

Cell cycle progression is a complex process regulated by environmental signals and intracellular checkpoints. As a result, the length of each cell cycle phase varies depending on the cell type, and also among cells of the same kind [195]. However, it has been observed that the variability in the length of G1 phase is often greater than that of other phases [159, 196]. Therefore, we decided to treat G1 duration as a random variable, while assigning a fixed duration to G2 phase. Hence, once the cell has committed to cell division, it takes a fixed amount of time to divide.

In this thesis, we use two different distributions to sample the G1 duration: the exponential distribution and the uniform distribution. We chose these distributions because they are able to represent the stochasticity of the G1 phase, while being simple enough to allow for mathematical analysis. In addition, using multiple cell cycle models enables us to test the sensitivity of the death clock framework with respect to the choice of cell cycle model. The exponential distribution has only one parameter and is "memoryless", which means that the time left until the cell transitions to G2 phase is independent of how much time has already elapsed. In contrast, the uniform distribution has two parameters and therefore spans a wider range of behaviours. The benefits are that it models a cell cycle where the time until transition is dependent on how much time has already elapsed, and that the mean and variance can be controlled independently, giving us more freedom to experiment with varying the G1 duration.

We note that an exponentially distributed G1 phase means that there is a high probability of a cell transitioning to G2 phase immediately after division. In reality, however, multiple biochemical steps must occur before the G1 phase is completed, so an instant transition is not realistic. It is this observation that led Yates et al. to represent the cell cycle as a series of exponentially distributed stages, instead of a single exponentially distributed stage [197]. Similarly, the Erlang, or generalised Erlang, distribution may be a more biologically accurate choice to model the G1 phase than the exponential distribution. Nevertheless, we use the exponential distribution here for simplicity. Moreover, we also use the uniform distribution, which does not suffer from this issue. Finally, the death clock framework is fully modular with respect to the cell cycle model, so more biologically accurate distributions can easily be incorporated in future iterations.

We model the accumulation of death signals using an ordinary differential equation (ODE) model. Specifically, the death signal is represented as a time-dependent quantity, and the **death clock** is a scalar variable whose rate of increase is given by the death signal. We do not consider signals that inhibit apoptosis and assume that only positive death signals are possible. In the absence of a death signal, the death clock is static. Therefore, the death clock can be interpreted as the cell's memory of past competition-related signalling events. If the

death clock reaches a threshold value while the cell is in G1 phase, it triggers apoptosis.

3.1.3 Death clock framework

The death clock framework comprises two coupled cellular processes: the cell cycle and the death clock. The cell cycle governs the transition of the cell from G1 phase to G2 phase, as well as cell division. On the other hand, the death clock governs the initiation of apoptosis in response to death signals. We consider the cell cycle to be an **autonomous** process, meaning that it is not affected by other cells. On the other hand, the death clock is a **non-autonomous** process because it relies on extracellular signals produced by other cells. Together, these processes determine whether and when the cell divides or initiates apoptosis.

At birth, we sample a stochastic G1 duration, denoted as t^*, from the **G1 duration distribution** C, i.e.

$$t^* \sim C\,, \tag{3.1}$$

where C is subject to the following constraints:

$$t^* \in [0, \infty)\,, \tag{3.2}$$

i.e. the stochastic G1 duration is non-negative, and

$$E(t^*) = t_{\mathrm{G1}}\,, \tag{3.3}$$

i.e. the mean G1 duration is equal to the **autonomous G1 duration** t_{G1}, a non-negative parameter. Note that t^* is a stochastic variable and t_{G1} is a fixed parameter. We discuss the concrete G1 duration distributions used in this thesis in Section 3.1.4.

If apoptosis is not triggered by the death clock, the cell spends a duration t^* in G1 phase and then transitions into G2 phase. After spending a fixed duration in G2 phase, determined by the non-negative parameter t_{G2}, the cell divides and the process repeats for each of the daughter cells (see Timeline A in Figure 3.1).

The **death clock**, denoted by the symbol $\tau(t)$, evolves according to the ODE

$$\frac{\mathrm{d}\tau}{\mathrm{d}t} = f(t)\,, \tag{3.4}$$

where $f(t)$ is the **death signal** experienced by the cell. We assume that $f(t)$ is non-negative, i.e.

$$f(t) \geq 0\,. \tag{3.5}$$

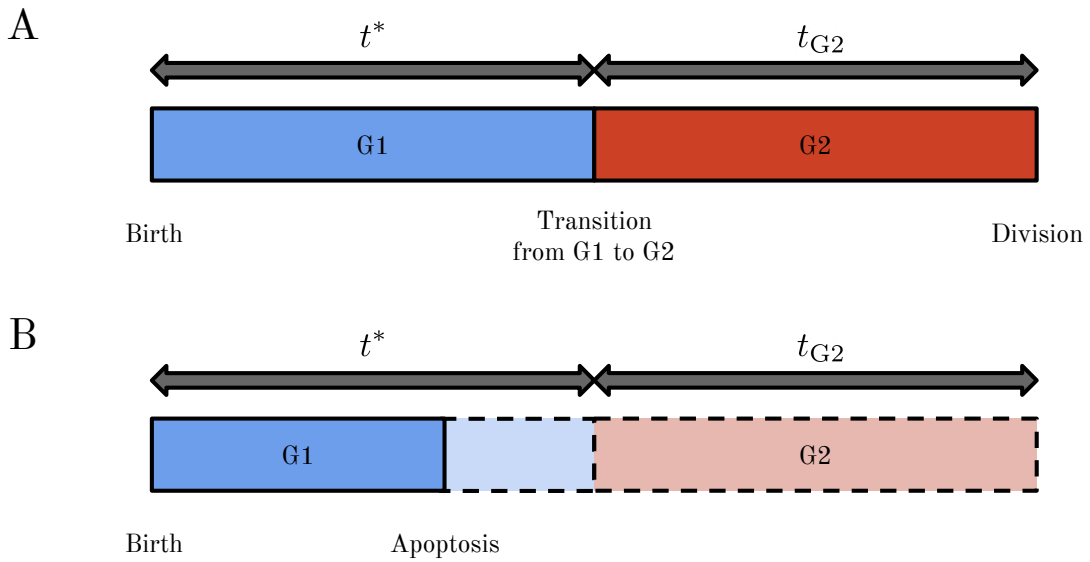

Figure 3.1: Cell timelines. (A) The timeline for a cell whose death clock does not initiate apoptosis. (B) The timeline for a cell whose death clock initiates apoptosis. Note that apoptosis shortens the amount of time the cell spends in G1 phase.

At birth, the death clock is initialised to zero, i.e.

$$\tau(t = 0) = 0\,, \tag{3.6}$$

where we have defined $t = 0$ as the time of birth. The apoptosis rule is as follows:

$$\text{Cell is in G1 phase} \quad \textbf{and} \quad \tau(t) \text{ reaches } T_\dagger \quad \Rightarrow \quad \text{initiate apoptosis}\,, \tag{3.7}$$

where $T_\dagger$ is the **death threshold**. See also Timeline B in Figure 3.1.

For mathematical convenience, it is often easier to formulate the apoptosis mechanism as a condition, rather than a rule. Hence, although the death clock is meaningless once apoptosis has been initiated, we let a fictitious death clock run until $t = t^*$ and define the **death condition** as

$$\exists\, t_\dagger \in [0, t^*] \quad \text{such that} \quad \tau(t_\dagger) = T_\dagger\,. \tag{3.8}$$

If the death condition is satisfied, then we conclude in hindsight that the cell has died at some time $t_\dagger \in [0, t^*]$. In reality, the cell initiates apoptosis at $t = t_\dagger$, so the death clock does not actually exist for $t \in (t_\dagger, t^*]$. By the same token, the value of $\tau(t)$ in $(t_\dagger, t^*]$ does not affect the outcome of the death clock model, justifying the use of a fictitious death clock.

In summary, there are two sources of uncertainty in the death clock framework: the variability in G1 duration and the death signal. The former originates from the cell cycle,

and the latter from intercellular interactions. Both contribute to the decision of the cell to initiate apoptosis. Our framework can thus be seen as a minimalist model of autonomous and non-autonomous processes interacting to govern competition-induced apoptosis. The model is summarised by the flowchart in Figure 3.2.

The parameters in the death clock model are as follows:

Cell cycle parameters The constants used to parameterise C, including the autonomous G1 duration t_{G1} and the fixed G2 duration t_{G2}.

Death signal parameters The constants used to parameterise $f(t)$.

Death threshold The threshold value for the death clock that triggers apoptosis, represented by the symbol $T_\dagger$.

In this section, we have discussed the death clock framework within the context of a single cell. There are two components missing from this description before we can implement a complete model for cell competition. The first is a cell-based model, where each cell has an internal state, which includes the cell cycle and death clock state, and an external representation that is perceived by other cells. In Section 3.3, we discuss the implementation of the death clock framework in two types of cell-based models. The second component is a concrete choice of death signal, such as contact-based signalling or mechanical stresses. The form of the death signal is constrained by the choice of cell-based model, since the cell-based model determines which aspects of the extracellular environment are observable to cells. For example, we can only use mechanical properties in the death signal if the cell-based model under consideration has a mechanical basis. We will introduce concrete death signals when discussing specific death clock models in later chapters.

3.1.4 Cell cycle models

Exponential cell cycle model

For the exponential cell cycle model, the probability density function for G1 phase is

$$\psi(t) = \frac{1}{t_{\mathrm{G1}}} \exp\left(-\frac{t}{t_{\mathrm{G1}}}\right), \tag{3.9}$$

where we set the rate parameter of the exponential distribution equal to $1/t_{\mathrm{G1}}$ in order to satisfy Equation (3.3). The cumulative distribution function is then

$$\Psi(t) = 1 - \exp\left(-\frac{t}{t_{\mathrm{G1}}}\right). \tag{3.10}$$

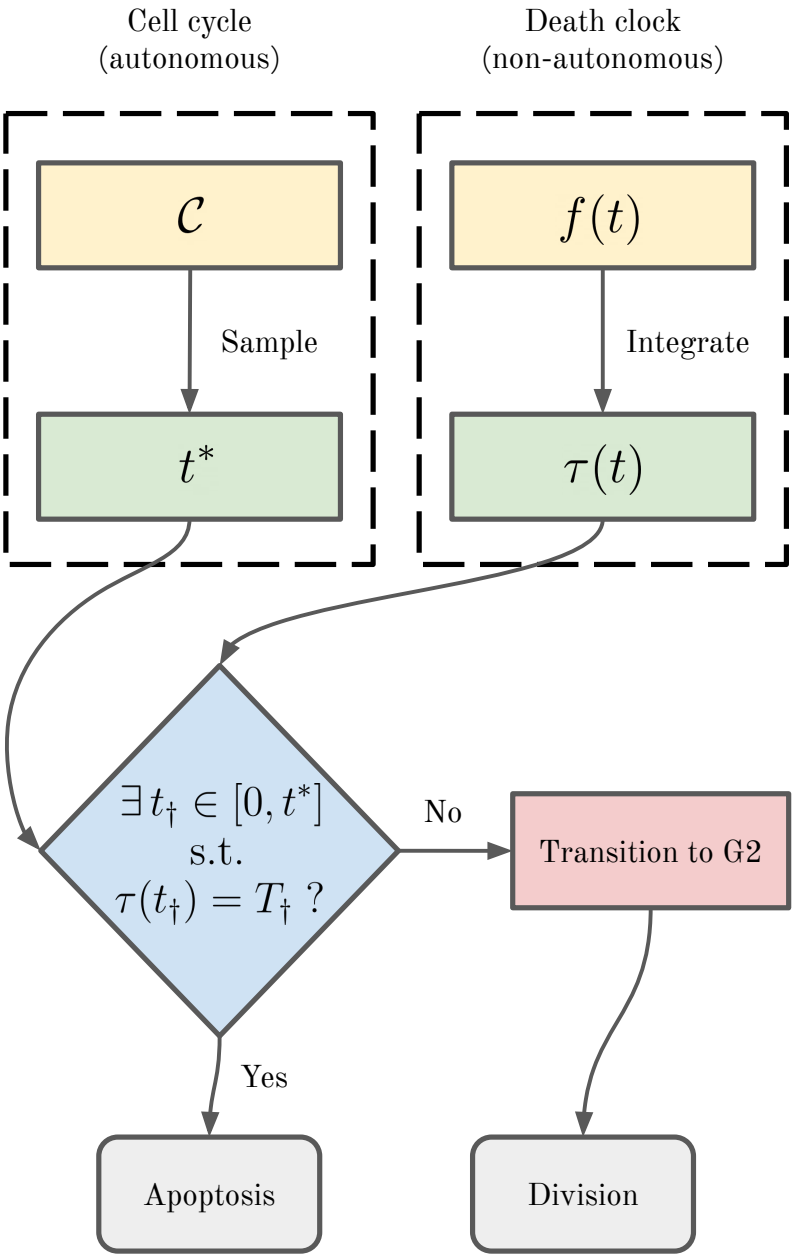

Figure 3.2: Death clock flowchart. The "Sample" step corresponds to Equation (3.1). The "Integrate" step corresponds to Equation (3.4) with initial conditions given by Equation (3.6). The condition in the decision block is the death condition, corresponding to Equation (3.8).

Uniform cell cycle model

For the uniform cell cycle model, the probability density function for G1 phase is

$$\psi(t) = \begin{cases} \frac{1}{r} & \text{for} \quad t \in \left[t_{\mathrm{G1}} - \frac{1}{2}r, t_{\mathrm{G1}} + \frac{1}{2}r\right] \\ 0 & \text{otherwise} \end{cases} , \tag{3.11}$$

where r is the width of the uniform interval. We parameterise the uniform distribution by the mean and the width of the interval to enforce Equation (3.3). We note that r must be positive and that the lower bound of the uniform interval, $t_{\mathrm{G1}} - r/2$, must be non-negative to satisfy Equation (3.2). Therefore, r is bounded as

$$0 < r \leq 2t_{\mathrm{G1}} \, . \tag{3.12}$$

The cumulative distribution function is given by

$$\Psi(t) = \begin{cases} 0 & \text{for} \quad t < t_{\mathrm{G1}} - \frac{1}{2}r \\ \frac{t-\left(t_{\mathrm{G1}} - \frac{1}{2}r\right)}{r} & \text{for} \quad t \in \left[t_{\mathrm{G1}} - \frac{1}{2}r, t_{\mathrm{G1}} + \frac{1}{2}r\right] \\ 1 & \text{for} \quad t > t_{\mathrm{G1}} + \frac{1}{2}r \end{cases} . \tag{3.13}$$

In practice, we permit $r = 0$ as a valid parameter in our simulations; this corresponds to a fixed G1 duration exactly equal to t_{G1}. However, the distribution is technically not a uniform distribution in that case, but instead a Dirac delta function centred on t_{G1}.

3.2 Survival probability

The outcome of the death clock mechanism is dependent on the interaction of death signals, representing non-autonomous processes, and the cell cycle, representing autonomous cellular processes. Since both of these components are uncertain, it is hard to make predictions on the cell fate when considering them simultaneously. However, if we control for the variability in death signals by considering a fixed death signal, we are left only with the cell cycle variance.

In this section, we take this approach to characterise the probability of different outcomes, given a fixed death signal. Specifically, we will derive an expression for the "survival probability", defined as the probability of a cell transitioning to G2 phase and dividing. Since the death clock framework only allows for two outcomes, the "death probability" is

simply the complement of the survival probability.

We begin by defining the survival condition as the negation of the death condition, given in Equation (3.8). Since the death clock is initialised to zero at birth, we can write the survival condition as

$$\tau(t) < T_\dagger \quad \text{for all} \quad t \in [0, t^*] \,. \tag{3.14}$$

Moreover, we assume that $f(t) \geq 0$ (Equation (3.5)), so we can reduce the condition in Equation (3.14) to

$$\tau(t^*) < T_\dagger \,. \tag{3.15}$$

3.2.1 Special case: constant death signal

We first derive the survival probability for a constant death signal, i.e. $f(t) = c$, before moving on to the general case. We solve Equation (3.4) with initial conditions given by Equation (3.6) to obtain

$$\tau(t) = ct \,. \tag{3.16}$$

We thus have a closed-form solution for the death clock. The remaining uncertainty lies solely with the stochastic G1 duration t^*. We define the **survival probability** θ as the probability that Equation (3.15) is satisfied, i.e.

$$\theta \equiv P(\tau(t^*) < T_\dagger) \,. \tag{3.17}$$

Substituting Equation (3.16), we obtain

$$\theta = P(ct^* < T_\dagger) = P\left(t^* < \frac{T_\dagger}{c}\right) . \tag{3.18}$$

Importantly, we have written the survival condition so that t^* appears as a single term in the inequality. This allows us to substitute the cumulative distribution function of t^*, denoted as $\Psi(t)$, hence

$$\theta = P\left(t^* < \frac{T_\dagger}{c}\right) = \Psi\left(\frac{T_\dagger}{c}\right) . \tag{3.19}$$

For example, in the case of an exponential cell cycle model, the survival probability is

$$\theta = 1 - \exp\left(-\frac{T_\dagger}{ct_{\mathrm{G1}}}\right) . \tag{3.20}$$

3.2.2 General case

We now derive the survival probability for an arbitrary death signal $f(t)$, subject to Equation (3.5) and assuming it is an integrable function. We define

$$F(t) \equiv \int_0^t f(t')\,\mathrm{d}t' \,. \tag{3.21}$$

Therefore, the death clock evolves as

$$\tau(t) = F(t) \,. \tag{3.22}$$

In particular, this gives the value of the death clock at time t^* as

$$\tau(t^*) = F(t^*) \,. \tag{3.23}$$

Substituting Equation (3.23) in the survival condition (Equation (3.15)), we have

$$F(t^*) < T_\dagger \,. \tag{3.24}$$

In order to obtain an expression for θ using the cumulative distribution function for t^*, we need to separate out t^* in Equation (3.24). Therefore, we need to "invert" $F(t)$. Because of Equation (3.5), we know that $F(t)$ is a monotonically increasing function. However, it is not *strictly* monotonically increasing. Intuitively, in the absence of a death signal (i.e. $f(t) = 0$) the death clock is not ticking. Thus, the death clock may have the same value for multiple time points.

As a result, it is not guaranteed that an inverse function of $F(t)$ exists. However, we are only interested in the earliest time point at which the death clock threshold is reached, since that uniquely defines the cut-off between survival and death. Therefore, we can make progress by defining the **pseudoinverse function** of $F(t)$ as

$$F^{-1}(\tau) \equiv \min\{t \in [0, \infty) : F(t) = \tau\} \,. \tag{3.25}$$

Because $F(t)$ is a monotonically increasing function, the pseudoinverse $F^{-1}(\tau)$ must be strictly monotonic increasing. Hence, we can apply it to Equation (3.24) to reformulate the survival condition as

$$t^* < F^{-1}\left(T_\dagger\right) \,. \tag{3.26}$$

Therefore, the general survival probability is given by

$$\theta = P\left(t^* < F^{-1}\left(T_\dagger\right)\right) = \Psi\left(F^{-1}\left(T_\dagger\right)\right) . \tag{3.27}$$

Since the right-hand side of Equation (3.26) is the amount of time that needs to pass after birth for a cell to reach the death threshold and initiate apoptosis, we name this quantity the **death time**, and denote it as

$$t_\dagger = F^{-1}\left(T_\dagger\right) . \tag{3.28}$$

3.2.3 Applications

The main use case for the survival probability is making theoretical predictions of cell fate outcome for a known death signal. This is particularly useful when the death signal is predictable, such as in the case of a constant death signal, which we discuss further in Chapter 4. When the death signal is variable, it is nonetheless possible in some cases, as we will see in Chapters 5 and 6, to approximate the death signal by a constant and compute an approximate survival probability.

Finally, the survival probability can be used as a measure of cell fitness. There are many definitions of fitness, but one of the most common definitions is *the expected number of offspring in one generation* [198]. In order to apply this definition to cells, we consider the cell cycle as one generation and any daughter cells as the offspring of the cell. We then compute the expected number of offspring in one generation by enumerating the outcomes of a single cell completing a cell cycle. In the death clock framework, there are two possible outcomes: survival or death. In the former case, the mother cell produces two daughter cells. In the latter case, there are no offspring. Furthermore, survival occurs with probability θ, and death occurs with probability $1 - \theta$. Thus, for a given survival probability θ, we define the fitness W as

$$W = E(\text{offspring}) = 2 \times \theta + 0 \times (1 - \theta) = 2\theta . \tag{3.29}$$

This equation provides a direct conversion between survival probability and fitness.

3.3 Cell-based death clock models

So far, we have discussed the death clock framework from the perspective of a single cell. However, the aim of the death clock framework is to model cell competition, which necessarily involves multiple cells. We therefore need to embed the death clock framework within a cell-based model so that we can simulate the interaction of multiple cells, each with

an internal death clock mechanism. The cell-based model provides the broader context for the death clock framework and simulates the extracellular environment that the death signal is derived from.

In this section, we discuss the implementation of the death clock framework in two distinct cell-based models: the well-mixed model and the vertex model (see Chapter 2 for a detailed discussion of the vertex model). The main difference between these models is the spatial representation of the tissue. In the well-mixed model, we assume that the cells are in a well-mixed culture where every cell interacts equally with every other cell. On the other hand, cells in the vertex model interact only with cells in their local environment. Therefore, we expect the vertex model to display more complex behaviours than the well-mixed model.

We note that the assumption of a well-mixed culture is a poor fit for the model system of cell competition, the *Drosophila* wing disc. Despite this, the well-mixed model can be considered a minimal cell-based model for the death clock. Therefore, it is useful as a tool to study the properties of the death clock model when spatial organisation can be ignored or neglected. In addition, it is much more computationally efficient to run simulations of the well-mixed model because it lacks an explicit spatial representation.

In summary, the well-mixed model is a "coarse-grained" cell-based model that treats the tissue as a homogeneous cell culture. As a result, we lose the ability to model the spatial aspects of competition, but we gain the benefit of cheap simulations and increased analytical tractability. On the other hand, the vertex model is a "fine-grained" cell-based model that simulates the spatial structure of the tissue, but is much more costly to run. As we study the properties of the death clock model, we will use both models in a complementary manner. The code implementing the cell-based death clock models can be found in the following GitHub repository: `https://github.com/ThomasPak/cell-competition`.

3.3.1 Well-mixed model

Because there is no explicit spatial structure in the well-mixed model, the state of the system is fully determined by simply aggregating all individual cell states. We represent the cell state with the **cell vector** $\mathbf{y}_\alpha(t)$, where $\alpha = 1, \ldots, N(t)$, and $N(t)$ is the number of cells at time t. The number of cells $N(t)$ is updated whenever a division or death event occurs, accompanied by a reindexing operation in the case of a death event. We equip every cell with a death clock model, so the cell vector must contain, at a minimum, all the information required to compute the individual cell cycle and death clock.

We distinguish three distinct parts in the cell vector, ordered by the degree of temporal transience. The first part contains one element: the death clock $\tau_\alpha(t)$, which changes

Table 3.1: Summary of cell vector elements in the well-mixed model. The type refers to the data type used to store the element in the Python implementation.

Symbol	Description	Type
$\tau_\alpha(t)$	Death clock	float
t^*_α	Sampled G1 duration	float
t^0_α	Birth time	float
C_α	G1 duration distribution	function object
$t_{\mathrm{G2},\alpha}$	G2 duration	float
$f_\alpha(\cdot)$	Death signal function	function object
$T_{\dagger,\alpha}$	Death threshold	float

continuously over time. The second part consists of the sampled G1 duration t^*_α and the birth time t^0_α, which are assigned at birth and remain fixed over the cell's lifetime. Finally, the third part comprises the cell cycle and death clock components. Specifically, the G1 duration distribution C_α (which is implicitly parameterised by the autonomous G1 duration $t_{\mathrm{G1},\alpha}$ and other cell cycle parameters, if applicable), the G2 duration $t_{\mathrm{G2},\alpha}$, the **death signal function** $f_\alpha(\cdot)$ (discussed below), and the death threshold $T_{\dagger,\alpha}$. These components are inherited from mother to daughter cell and are therefore conserved throughout the whole simulation. We write the cell vector as

$$\mathbf{y}_\alpha(t) \equiv \left[\; \tau_\alpha(t) \;\middle|\; t^*_\alpha \quad t^0_\alpha \;\middle|\; C_\alpha \quad t_{\mathrm{G2},\alpha} \quad f_\alpha(\cdot) \quad T_{\dagger,\alpha} \; \right], \tag{3.30}$$

and summarise its contents in Table 3.1. The state of the system, denoted $S(t)$, is then given by

$$S(t) \equiv \{ \mathbf{y}_1(t), \mathbf{y}_2(t), \ldots, \mathbf{y}_{N(t)}(t) \}. \tag{3.31}$$

Accordingly, the initial conditions of the well-mixed model are given by an initial system state of the form

$$S(t=0) = \{ \mathbf{y}_1(t=0), \mathbf{y}_2(t=0), \ldots, \mathbf{y}_{N(t=0)}(t=0) \}. \tag{3.32}$$

These initial conditions must conform to the simulation invariants discussed below.

We distinguish between two distinct rule sets that determine the evolution of the well-mixed model. The first rule set controls the evolution of the death clocks, and the second rule set governs division and death events. The former rule set pertains to continuous quantities, specifically the death clock, while the latter rule set involves discrete operations.

For the first rule set, we apply the death clock ODE, Equation (3.4), to each cell. Since the well-mixed model assumes spatial homogeneity, it follows that all cells observe the same homogeneous environment. On the other hand, as a cell-based model, each cell responds to this environment as an individual. The conversion of extracellular signals to a concrete death signal is mediated by the death signal function $f_\alpha(\cdot)$. Because the extracellular signals are derived from the system state $S(t)$, we write

$$\frac{\mathrm{d}\tau_\alpha}{\mathrm{d}t} = f_\alpha(S(t)), \tag{3.33}$$

where $f_\alpha(S(t))$ is the concrete death signal that cell α experiences as a result of the system state $S(t)$.

Simulation invariants

The discrete division and death operations are triggered when the **division invariant** and the **death invariant**, respectively, are violated. The division invariant is as follows:

$$\forall \alpha = 1, \ldots, N(t) : t - t_\alpha^0 < t_\alpha^* + t_{\mathrm{G2},\alpha}\,. \tag{3.34}$$

The division invariant can be read as "every cell's age must be less than its total cell cycle duration". The death invariant is as follows:

$$\forall \alpha = 1, \ldots, N(t) : t - t_\alpha^0 < t_\alpha^* \Rightarrow \tau_\alpha(t) < T_{\dagger,\alpha}\,. \tag{3.35}$$

The death invariant can be read as "for every cell in G1 phase, its death clock must be below its death threshold".

These invariants must be respected at all times during the simulation, including the initial conditions. When an invariant is violated, the corresponding operation is instantly triggered. The operation is guaranteed to restore the invariant, therefore allowing the simulation to proceed while respecting the invariants.

Division operation

As soon as the division invariant is violated, i.e. if

$$\exists\, \alpha \in 1, \ldots, N(t) : t - t_\alpha^0 = t_\alpha^* + t_{\mathrm{G2},\alpha}\,, \tag{3.36}$$

or, in words, "as soon as any cell's age reaches its total cell cycle duration", division is triggered. Biologically, the mother cell splits into two daughter cells. Computationally, we

perform this operation by adding a daughter cell vector and reusing the mother cell vector to construct the second daughter cell in-place.

Concretely, we update the population as[2]

$$S(t) \leftarrow S(t) \cup \{\, \mathbf{y}_{N(t)+1}(t) \,\}, \tag{3.37}$$

and complete the contents of the cell vectors $\mathbf{y}_\alpha(t)$ and $\mathbf{y}_{N(t)+1}(t)$ as below. Finally, we increment $N(t)$ to reflect the increased population count.

Inherit cell cycle and death clock components Both daughter cells inherit the mother cell's cell cycle and death clock components. Since we repurpose the mother cell vector $\mathbf{y}_\alpha(t)$ for the first daughter cell, we need only copy the relevant fields to $\mathbf{y}_{N(t)+1}(t)$:

$$C_{N+1} \leftarrow C_\alpha\,, \tag{3.38a}$$
$$t_{\mathrm{G2},N+1} \leftarrow t_{\mathrm{G2},\alpha}\,, \tag{3.38b}$$
$$f_{N+1}(\cdot) \leftarrow f_\alpha(\cdot)\,, \tag{3.38c}$$
$$T_{\dagger,N+1} \leftarrow T_{\dagger,\alpha}\,. \tag{3.38d}$$

Sample G1 duration We sample a G1 duration for each daughter cell from the cell cycle distribution as[3]

$$t^*_\alpha \leftsquigarrow C_\alpha\,, \tag{3.39a}$$
$$t^*_{N+1} \leftsquigarrow C_{N+1}\,. \tag{3.39b}$$

Set birth time We set the birth times of the daughter cells as

$$t^0_\alpha, t^0_{N+1} \leftarrow t_{\text{division}}\,, \tag{3.40}$$

where t_{division} is the time at which the division invariant was violated.

Reset death clock We reset the death clocks of the daughter cells so that they satisfy the death clock initial conditions given by Equation (3.6):

$$\tau^0_\alpha, \tau^0_{N+1} \leftarrow 0\,. \tag{3.41}$$

[2]The symbol '$\leftarrow$' represents the update operator, meaning that the left operand(s) are updated with the value of the right operand.

[3]The symbol '$\leftsquigarrow$' represents the sample-and-update operator, meaning that the left operand is updated with a value sampled from the right operand, which must be a stochastic distribution.

Death operation

As soon as the death invariant is violated, i.e. if

$$\exists\, \alpha \in 1, \ldots, N(t) : t - t_\alpha^0 < t_\alpha^* \text{ and } \tau_\alpha(t) = T_{\dagger,\alpha}\,, \tag{3.42}$$

or, in words, "as soon as the death clock of a cell in G1 phase reaches the death threshold", apoptosis is triggered.

The population is updated as

$$S(t) \leftarrow S(t) \setminus \{\, \mathbf{y}_\alpha(t) \,\}\,, \tag{3.43}$$

followed by a reindexing operation to reflect the decreased population count. We also decrement $N(t)$.

Numerical implementation

We implemented the well-mixed model in Python using the class `WellMixedSimulator`. This class is responsible for checking the validity of the initial conditions, simulating the system forward in time by applying the continuous and discrete rule sets, and terminating the simulation. In addition, the class records the state of the system over time and returns raw simulation data when the simulation finishes.

ODE solver The simulation proceeds by numerically solving Equation (3.33) for all cells until an invariant is violated, which triggers the corresponding discrete operation. By default, we use the SciPy [199] ODE solver, `solve_ivp`, to integrate Equation (3.33). The `events` argument to the solver lets us specify conditions on the ODE system for which the solver terminates before the simulation end time is reached. Concretely, `events` is an array of functions that map the ODE system state to scalar values. The solver then evaluates these functions as it solves the ODE and triggers an event when a function value crosses zero. We use this feature to exit the ODE solver when an invariant is violated so that we can perform the corresponding discrete operation outside of the solver. After restoring the invariant, numerical integration of Equation (3.33) is resumed. The discrete operations are implemented by methods of the `WellMixedSimulator` class. In short, the standard mode of operation is numerical integration within the SciPy ODE solver, punctuated by division and death events implemented in the simulator class.

Optimised ODE solver Although the ODE solver is relatively efficient, we can dispense with numerical integration entirely in special cases in order to gain a massive performance boost. In particular, if the death signal is constant for all cells in the intervals between discrete events, then Equation (3.33) can be solved exactly. As a result, the time until the next event can be computed by simple arithmetic operations. For instance, if the death signal depends only on the total cell count, then the death signal only changes through division and death events so that we can apply this optimisation.

In addition to the division and death events, we added the transition from G1 phase to G2 phase as a third type of discrete event in the simulator. This event is not strictly necessary for simulating the model, but it forces the ODE solver to terminate when a cell transitions from G1 phase to G2 phase. With this extended set of event types, death signals that depend on the distribution of G1 and G2 phases are constant in between discrete events, such that explicit numerical integration can be avoided in this case as well. The simulator class by itself cannot infer whether the death signal is constant between discrete events, so in practice we control this optimisation by passing the Boolean flag `f_is_stepwise_constant` to the simulator.

Pseudorandom number generator Since G1 durations are sampled from a stochastic distribution, the well-mixed model is inherently stochastic. Therefore, in order to run reproducible simulations, we need to have control over the random number generator used by the stochastic sampler. Thus, we require that the cell cycle model is implemented using the `RandomState` Mersenne Twister pseudorandom number generator from the NumPy random library [199]. Moreover, the simulator optionally accepts a `seed` to initialise `RandomState` with. If no seed is given, it is initialised instead with a fresh random seed supplied by the operating system.

Termination conditions In addition, the simulator class accepts simulation parameters that specify the simulation time and govern the termination conditions of the simulation. In particular, `tstart` and `tend` are the start and end times of the simulation, respectively, and `min_cell_count` and `max_cell_count` are the minimum and maximum cell counts, respectively. The simulation terminates when the simulation time reaches `tend`, when the cell count drops to `min_cell_count`, or when the cell count reaches `max_cell_count`; whichever event happens first. We summarise the simulation parameters in Table 3.2.

Table 3.2: Summary of simulation parameters in the well-mixed model. The type refers to the data type used to store the parameter in the Python implementation.

Name	Description	Type
`seed`	Seed for `RandomState`	integer
`tstart`	Simulation start time	float
`tend`	Simulation end time	float
`min_cell_count`	Minimum cell count	integer
`max_cell_count`	Maximum cell count	integer
`f_is_stepwise_constant`	Flag for stepwise constant death signal	Boolean

Data processing

The `WellMixedSimulator` class returns the results of the simulation in the form of a Python dictionary containing: timestamps for every discrete event; the numerical solution of the death clocks; the time of birth and sampled G1 duration of every cell; and the termination status. The termination status indicates which condition led to the termination of the simulation. Although the data gives a complete record of the simulation, this format is difficult to parse directly.

Hence, we implemented the `WellMixedSimulationData` class to process the raw data and represent it in a useful manner. There are many different data processing methods and plotting techniques we can implement in the class to analyse our data. However, oftentimes we only want to use a subset of those methods. In addition, some of the data processing methods can be relatively costly to run. Therefore, we want to avoid running all the data processing routines when initialising the class. We achieve this using **lazy initialisation**, where the class initially does not process any data at all, and any individual data processing method is only called when the class is queried for the result of that method. Once a data processing result has been queried and computed, it is retained for future access.

3.3.2 Vertex model

As a cell-based model, the vertex model is a framework that already has a concrete representation of cells and their cell cycles, as well as implementations for cell division and cell death. In addition, the two-phase cell cycle model proposed in Chapter 2 is already congruent with the cell cycle model of the death clock framework. Therefore, as opposed to the well-mixed model, which we constructed *de novo*, our design strategy here is to augment the vertex model from Chapter 2 with components that implement the death clock

framework. In particular, we discuss how we equip cells with a death clock that evolves according to Equation (3.4), how we compute the death signal, and the application of the apoptosis rule given by Equation (3.7).

The main difference between the vertex model and the well-mixed model with respect to the death clock framework is computation of the death signal. In the well-mixed model, each cell observes the same global system state because the tissue is well-mixed. However, in the vertex model, cells are restricted to sensing their local environment. Practically, this means that cells receive different inputs to their death signal function based on their location in the tissue. Therefore, cells perceive a spatially constrained subset of the overall system state. We represent the local system state perceived by cell α using the "input vector" $\mathbf{x}_\alpha(t)$. This enables us to write the death clock ODE in the vertex model as

$$\frac{\mathrm{d}\tau_\alpha}{\mathrm{d}t} = f_\alpha(\mathbf{x}_\alpha(t))\,, \tag{3.44}$$

where the death signal function $f_\alpha(\cdot)$ converts the *local* system state to a concrete death signal. Note that the death signal function defined here is similar to the death signal function in the well-mixed model, which converts the *global* system state to a concrete death signal.

Chaste implementation

For the reasons discussed in Chapter 2, we use the Chaste implementation of the vertex model in our work. As an object-oriented library for cell-based simulations, Chaste was designed to be customisable and provides abstractions for modelling biologically relevant processes [160], such as apoptosis and subcellular reaction networks. Therefore, we implemented the death clock framework by adding custom concrete classes conforming to the abstract interface built into Chaste, rather than making structural changes to the codebase.

`CellData` One of Chaste's abstractions is a `CellData` object owned by each cell that holds an arbitrary amount of floating-point data. We use the `CellData` abstraction to store the cell state associated with the death clock framework. This includes the value of the death clock, death clock parameters such as the death threshold, as well as the simulation data populating the input vector. An example of simulation data is the number of neighbouring cells in the case of a death signal based on short-range cell interactions.

`DeathClockModifier` These inputs are updated by the `DeathClockModifier` class that implements Chaste's `SimulationModifier` interface. This is a general-purpose interface that provides read-and-write access to the entire simulation instance as it is being executed.

For our purposes, the simulation modifier extracts any pieces of simulation data relevant to the death clock at every timestep, specifically data that goes into the input vector, and stores them in `CellData` objects.

`DeathClockSrnModel` and `DeathClockOdeSystem` Chaste also offers functionality for modelling subcellular reaction networks with ODEs in tandem with a cell-based simulation. Once again, this is achieved through an object-oriented interface that lets the user specify a custom ODE system associated with each cell. In addition, the Chaste library includes a suite of numerical ODE solvers. However, the user can also supply their own solver if they wish.

We use this infrastructure to numerically integrate the death clock, i.e. Equation (3.44), on a per-cell basis. Concretely, we created the class `DeathClockOdeSystem`, which computes the right-hand side of Equation (3.44), and the class `DeathClockSrnModel`[4], which interfaces with both `DeathClockOdeSystem` and `CellData` to perform the logistical tasks involved with providing the ODE access to the input vector in `CellData`, setting up and running the numerical solver, and updating `CellData` with the death clock values thus computed.

`DeathClockCellKiller` Finally, we apply the apoptosis rule, given by Equation (3.7), with the `DeathClockCellKiller` class using Chaste's abstractions for killing cells during simulation. Iterating over all cells at every timestep, the cell killer class accesses each cell's death clock, as well as its cell cycle phase, to determine whether the cell should undergo apoptosis. We note that the application of the apoptosis rule is procedurally decoupled from the numerical integration of the death clock, in contrast to the well-mixed model, because cell death and numerical integration are performed as distinct steps in the broader cell-based simulation scheme. As a result, the precise timing of apoptosis is constrained by the simulation timestep. Assuming that the simulation timestep is small compared to the timescales relevant to the death clock, however, the numerical error resulting from this discretisation should be limited.

Another difference with the well-mixed model is that cells are not instantaneously removed from the tissue when apoptosis is triggered. Instead, an apoptotic state is induced, in which the cell is marked as dead and its target cell area linearly decreases to zero over a fixed period of time, parameterised by the apoptosis time $T_{\text{apoptosis}}$. As discussed in Chapter 2, the vertex model defines forces that try to restore any discrepancies between a cell's area and its target cell area. Therefore, an apoptotic cell will shrink in cell area until it is removed

[4] Srn is short for "subcellular reaction network".

by a T2 swap when its cell area falls below the area threshold $S_{\min}$. The exact timing of the T2 swap is determined by the mechanical interactions of the dying cell and its surroundings. Also, the T2 swap does not distinguish between apoptotic and non-apoptotic cells. It is therefore possible for a non-apoptotic cell to be removed from the tissue if it is "squeezed" out through mechanical forces. This cell fate, "death by extrusion", does not exist in the well-mixed model, where cells can only die through apoptosis. We record every T2 swap as an "apoptosis" event if the cell is apoptotic at the time of the T2 swap and as an "extrusion" event if not.

Summary We used Chaste's object-oriented interface to extend the vertex model with the death clock framework. Specifically, we associated a `DeathClockSrnModel` with each cell to embed an internal death clock, implemented a `DeathClockModifier` to extract relevant simulation data to compute the death signal, and added a `DeathClockCellKiller` to apply the apoptosis rule. Tying everything together is the `CellData`, which serves as the central point of communication. See Figure 3.3 for a diagram summarising the interactions between these components.

Chaste contributions

As mentioned in Section 2.1.4, standard vertex dynamics permit the mesh to become invalid due to cells intersecting each other. Mesh rearrangements, such as the T1, T2, and T3 transitions, prevent invalid meshes in a large number of cases. However, because of the complexity of the vertex model, there exist many more cases that these transitions cannot resolve, leading to simulation crashes. In most vertex-based simulations, this is not a problem because such invalid mesh configurations are "edge cases", in the sense that they are infrequent under normal conditions.

However, when the death clock framework is implemented in the vertex model, these edge cases become more commonplace. In particular, the death clock mechanism entails that all cells are susceptible to death by apoptosis, whereas most vertex-based studies are limited to death by extrusion. As a result, our early attempts at running the vertex model with the death clock resulted in up to 30% of simulations crashing.

Upon closer inspection, most of the errors related to apoptotic cells at the boundary of the tissue. Because boundary cells are not crowded out in the same manner as interior cells, they are more likely to self-intersect before being removed from the tissue. Chaste was not capable of resolving this type of mesh invalidation, so we identified and solved the underlying issues, and contributed the fixes back to the official Chaste repository. Some edge cases remain unresolved because they require significant architectural changes to

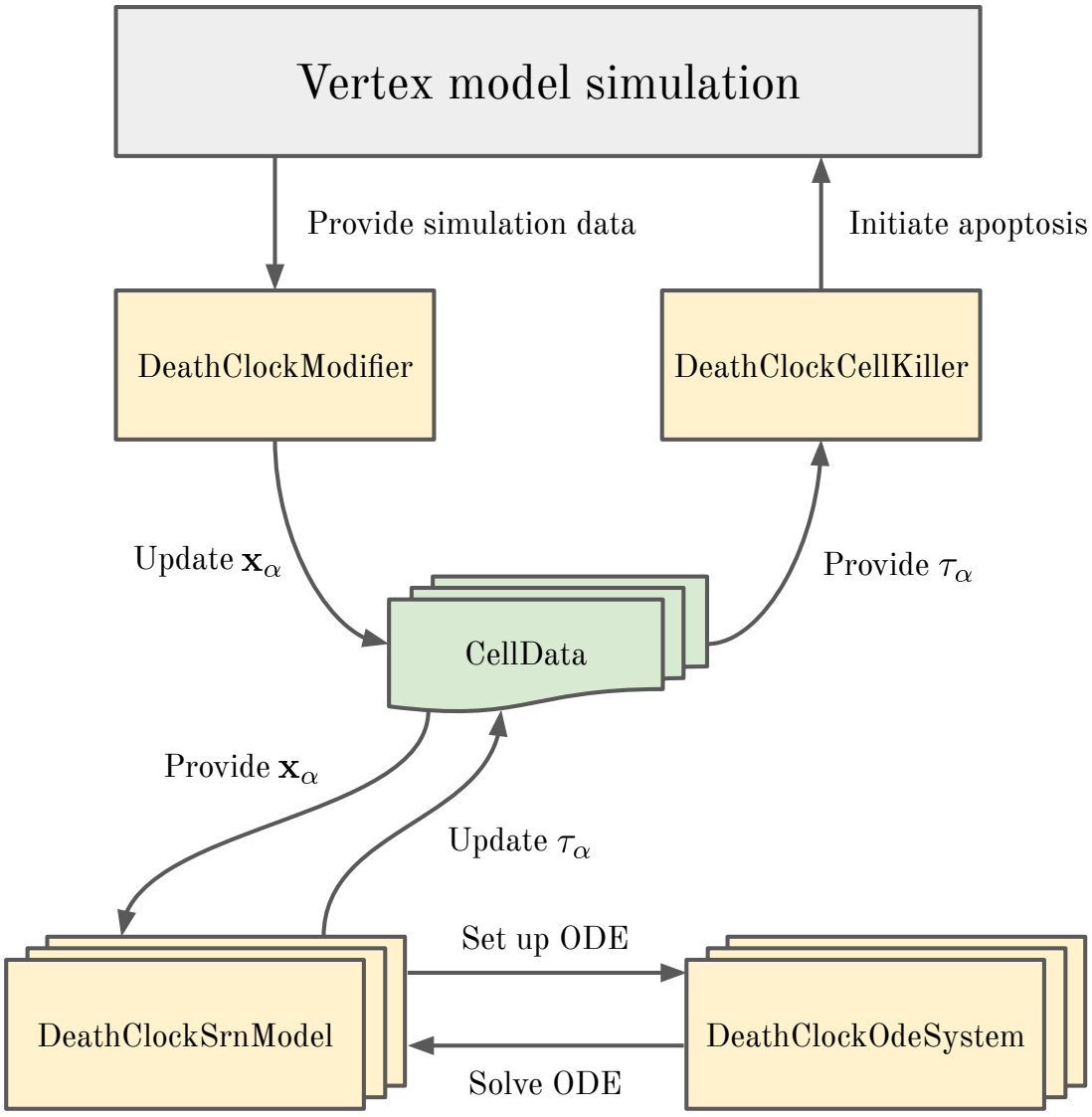

Figure 3.3: Diagram showing the components implementing the death clock framework in the vertex model and their interactions.

Chaste to fix, but fortunately they were relatively rare in the simulations that we ran. In addition to these "official" fixes, we also implemented a few *ad hoc* fixes (see the GitHub repository `https://github.com/ThomasPak/cell-competition` for details). Altogether, we reduced the simulation failure rate from 30% to roughly 1–2%. For an overview of our contributions to Chaste, see Appendix A.

3.4 Discussion

The failure in Chapter 2 to reproduce cell competition by varying mechanical parameters suggested that we need to implement an active mechanism of cell death to model cell competition. This led us to propose the novel modelling framework in this chapter based on the principle that apoptosis is triggered by the accumulation of death signals reaching a threshold. Biologically, apoptosis is regulated by multiple complex signalling networks. In this chapter, however, we abstracted away these details and described the accumulation mechanism using a simple ODE model that aggregates the death signal into an abstract quantity that we call the death clock.

In addition, we assume that the death clock is coupled to the cell cycle. In particular, the cell can only trigger apoptosis in G1 phase. The duration of G1 phase is stochastic, such that each cell has a different window in which it is vulnerable to apoptosis. We consider the cell cycle an autonomous process, since it does not depend on external influences, and we consider the death clock mechanism a non-autonomous process because it is driven by interactions with other cells. Hence, the death clock framework can be viewed as a model of autonomous and non-autonomous processes interacting to give rise to competition.

The main advantage of the death clock framework is that it does not impose an *a priori* cell status, but that competition arises from interactions among all cells. A corollary of this is that we can use the framework to model competition in homotypic cell populations just as naturally as in heterotypic cell populations. Moreover, the relatively simple formulation of the death clock ODE allows us to make theoretical predictions and run efficient computational simulations, while still capturing the basic principle of death signal accumulation. Furthermore, the death clock framework is very flexible because the death signal can take on different forms to represent different modes of competitive interactions.

We implemented our framework in two distinct cell-based models: the well-mixed model and the vertex model. The well-mixed model, which we constructed *de novo*, can be considered a minimal cell-based model implementing the death clock framework. On the other hand, we implemented the vertex model by extending Chaste with custom classes. These cell-based models enabled us to investigate the behaviour of the death clock using

simulations. The difference between the well-mixed model and the vertex model is that the latter is more realistic because it is a spatial model that incorporates the mechanical properties of an epithelial tissue. However, this comes at a great cost in terms of complexity and computational efficiency. In contrast, the well-mixed model lacks a spatial representation, but is much cheaper to run.

We also developed a theoretical tool for studying the death clock framework, namely the survival probability. It is difficult to make predictions on the cell fate outcome when both the cell cycle duration and the death signal are uncertain. For the cell cycle duration, uncertainty comes from stochasticity in the duration of G1 phase. For the death signal, uncertainty comes from the noisy extracellular environment. However, if we assume a fixed death signal, thus eliminating one source of uncertainty, we showed that it is possible to derive an exact expression for the survival probability. The survival probability is useful both theoretically to predict cell fates, and practically to quantify the fit with experiments by comparing the survival probability with experimental division and apoptosis rates.

The survival probability and the cell-based models provide analytical and computational tools, respectively, for studying the death clock modelling framework. On the one hand, we can make theoretical predictions using the survival probability, and, on the other hand, we can run simulations with cell-based models to verify the predictions. In the next chapter, we apply these tools to investigate the simplest death clock model: the constant death signal model.

Chapter 4

The constant death signal model

In the previous chapter, we introduced the death clock modelling framework and briefly discussed the constant death signal in Section 3.2.1. In particular, we predicted the survival probability of a single cell experiencing a constant death signal. In this chapter, we explore the implications of the constant death signal at the population level. Concretely, we study the behaviour of a cell population where every cell is subject to the same constant death signal, $f(t) = c$.

Since the death signal is constant, the induction of apoptosis in any given cell is not influenced by the surrounding cells. Therefore, cell death in the constant death signal model is an autonomous process. Cell competition, however, is a non-autonomous phenomenon, so the constant death signal model cannot represent cell competition. Instead, we interpret this model biologically as a cell population where every cell has the same propensity for apoptosis. This could be the result of an intrinsic propensity for apoptosis, or an external stressor[1] that affects all cells equally, such as UV irradiation or the presence of cytotoxic compounds.

The constant death signal model is the simplest possible death clock model, which makes it a suitable model for exploring the death clock framework before moving on to more complex death signals. In addition, we will use the results in this chapter as a foundation for the theoretical framework that we develop in Chapters 5 and 6.

The rest of this chapter is organised as follows: in Section 4.1 we relate the behaviour of the constant death signal model to that of a discrete-time birth–death Markov chain. We then use this relationship to make predictions about the proliferative behaviour of the population in Section 4.2. We also study the impact of the death signal on the time that cells spend in G1 phase in Section 4.3. In these two sections, we make concrete, quantitative predictions,

[1] We do not consider such stress factors as non-autonomous effects because they are not mediated by other cells in the population under consideration.

which we validate with simulations of the well-mixed model and the vertex model. Finally, we end with a discussion in Section 4.4.

4.1 Discrete-time birth–death Markov chain

As we have established in Section 3.2.1, in the constant death signal model every cell has the same survival probability θ because the death signal is identical for all cells. Hence, cell fates (survival versus death) are statistically independent and identically distributed. We can therefore express the behaviour of the population with a probabilistic scheme simulating the survival outcome of each cell one by one. This approach reduces the processes of the death clock and the cell cycle model to a single stochastic event, with probabilities determined by the survival probability.

Concretely, we consider a stochastic procedure where, given an initial cell population, we iteratively pick a cell at random, and simulate its cell fate, with the probability of survival given by θ, and the probability of death given by $1 - \theta$. In the case of survival, the cell divides into two daughter cells, so we increment the cell population. In the case of death, we decrement the cell population. Since every cell's fate is independent and identically distributed, cell identities are not important for the stochastic process and only the number of cells is significant.

The random process thus constructed is known in the literature as a **discrete-time birth–death Markov chain** [200]. Formally, this is a Markov chain on the state space $I = \{0, 1, 2, \ldots\}$, where each state corresponds to a particular population size. Furthermore, the transition probabilities are given by

$$p_{0,0} = 1\,, \tag{4.1a}$$

$$p_{i,i-1} = 1 - \theta \qquad \text{for } i \geq 1\,, \tag{4.1b}$$

$$p_{i,i+1} = \theta \qquad \text{for } i \geq 1\,. \tag{4.1c}$$

To arrive at this simplified description of the constant death signal model, we ignored the dynamical nature of the system by considering cell fates one by one, whereas in reality cells are progressing towards their fates in tandem. Therefore, the Markov chain does not respect the chronological order of events as they unfold in the constant death signal model.

As a result, the only properties that are transferable from the Markov chain to the constant death signal model are those that are independent of the chronological order of events. In particular, the asymptotic behaviour, such as the population's propensity to go extinct or proliferate indefinitely, is the same in both the Markov chain and the constant

death signal model, since the precise order of events is not important in this case. Another limitation of the Markov chain model is that it does not account for death by extrusion. Therefore, its ability to predict the behaviour of the vertex-based implementation of the constant death signal model is impacted by the rate of extrusion.

4.2 Proliferation regimes

In this section, we characterise the proliferation regimes of the constant death signal models, i.e. population extinction versus explosion, using the Markov chain model, and validate the predictions with simulations.

Since each cell has a nonzero probability of dying, it is possible for the entire cell population to go extinct at any time. Once this occurs, there are no cells left to give birth to new cells. Hence, population "extinction" is an irreversible event. This is reflected in the Markov chain model by the fact that the state 0 is an **absorbing state**; once the Markov chain reaches 0, it can never escape from it.

Furthermore, the literature on birth–death Markov chains [201] teaches us that only two asymptotic outcomes are possible in the Markov chain model: either the chain is absorbed by the state 0 (extinction), or it is not and the population "escapes to infinity" (explosion). We define the **extinction probability** h_i as the probability that the Markov chain reaches the state 0 starting from state i. It can be shown [201] that the extinction probability is given by

$$h_i = \begin{cases} 1 & \text{for} \quad \theta \leq 1/2 \\ \left(\frac{1-\theta}{\theta}\right)^i & \text{for} \quad \theta > 1/2 \end{cases} . \tag{4.2}$$

In particular, extinction is guaranteed if the survival probability θ is less than or equal to $1/2$. Intuitively, if individuals are more likely to die than to survive, then the population is doomed to extinction. On the other hand, if survival is more likely than death, the population has a nonzero probability of thriving indefinitely. Equation (4.2) also shows us that the extinction probability decreases rapidly as the population size i increases.

The implication for the constant death signal model is that the cell population should similarly go extinct or explode with the probabilities given by Equation (4.2). Notably, this eliminates the possibility of more complex behaviours such as oscillations or non-trivial steady states. Conversely, we conclude that the emergence of such behaviours requires a death signal that is *not* constant. In other words, oscillations or steady states require that the death signal varies temporally and/or across the cell population.

It can also be shown that for $\theta < 1/2$, the Markov chain reaches the state 0 in a finite

Table 4.1: Model and simulation parameter values used to estimate the extinction probability for the well-mixed model.

Parameter	Regime 1	Regime 3
t_{G1}		50
t_{G2}		50
c		1
θ	$\frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{3}{7}$	$\frac{4}{7}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}$
Initial cell count	$10, 20, 30, 40, 50$	$1, 2, 3, 4, 5$
Simulation end time		∞
Maximum cell count	∞	50
N_{sim}		1 000

number of events on average, while the average number of events until extinction is reached for $\theta = 1/2$ is unbounded [201]. Hence, we define three distinct proliferation regimes:

Regime 1 $\left\{\theta < \frac{1}{2}\right\}$ Extinction is guaranteed after a finite number of events on average.

Regime 2 $\left\{\theta = \frac{1}{2}\right\}$ Extinction is guaranteed after an infinite number of events on average.

Regime 3 $\left\{\theta > \frac{1}{2}\right\}$ Extinction is not guaranteed.

4.2.1 Computational validation

In this section, we verify whether the extinction probability predicted by Equation (4.2) matches the behaviour of the constant death signal model in simulations. We used a Monte Carlo method where we ran repeated simulations of our model for different values of θ and initial population sizes to sample the extinction frequency. We provided a unique seed for the random number generator in each simulation.

Parameter choice We chose θ values in Regimes 1 and 3, as listed in Table 4.1 for the well-mixed model and Table 4.2 for the vertex model. We excluded Regime 2 ($\theta = 1/2$) because the Markov chain model predicts it would take infinite time to verify whether the population goes extinct in this case.

In Regime 3, we chose relatively small initial cell counts (see Tables 4.1 and 4.2) such that the predicted extinction probabilities are not negligible. In theory, the population size should not matter in Regime 1. Nonetheless, we also varied the population size in Regime 1 to verify that it is insensitive to population size (see Tables 4.1 and 4.2). In addition, we

Table 4.2: Model and simulation parameter values used to estimate the extinction probability for the vertex model.

Parameter	Regime 1	Regime 3
t_{G1}		50
t_{G2}		50
c		1
θ	$\frac{1}{5}, \frac{1}{3}, \frac{3}{7}$	$\frac{4}{7}, \frac{2}{3}, \frac{4}{5}$
Initial cell count	10, 30, 50	1, 3, 5
Simulation end time		100 000
Maximum cell count	∞	50
N_{sim}		100

chose larger population sizes than in Regime 3 so that extinction is less likely to be the result of a small population size, as is the case in Regime 3.

Because the extinction probability is not dependent on the choice of cell cycle model, we used the exponential cell cycle model to validate the extinction probability because it is the simpler cell cycle model. In order to compute the death threshold, $T_\dagger$, we fix $t_{G1} = t_{G2} = 50, c = 1$ and invert Equation (3.20) to give the death threshold as a function of the survival probability:

$$T_\dagger = -ct_{G1} \ln(1 - \theta) . \tag{4.3}$$

The mechanical parameters used in the vertex model for all simulations in this chapter are set to their default values as given in Chapter 2.

For the well-mixed model, we ran 1 000 simulations for every unique parameter set, so the total number of simulations is $(5 \times 5 + 5 \times 5) \times 1\,000 = 50\,000$. For the vertex model, we ran 100 simulations for every unique parameter set. Hence, the total number of simulations is $(3 \times 3 + 3 \times 3) \times 100 = 1\,800$.

Initial conditions When we derived the survival probability, θ, for a cell experiencing a constant death signal in Section 3.2.1, we assumed that i) the cell is subject to the constant death signal starting from birth, ii) the initial death clock is set to zero, and iii) the G1 duration is sampled from the G1 duration distribution. This is true for all cells born during the simulation, but not necessarily so for the initial cell population.

To specify the initial conditions, we need to assign a birth time, t_α^0, an initial death clock value, τ_α, and a G1 duration, t_α^*, to every cell α. Hence, only a subset of possible initial conditions conform to the assumptions listed above. To ensure that the initial cells have the

same survival probability as the cells that appear afterwards in the simulation, we chose the initial conditions such that each cell mirrored the state of a newly-born cell.

Therefore, for each initial cell we set the birth time to the start of the simulation, initialised the death clock to zero, and sampled the initial G1 duration from the exponential cell cycle model described by Equation (3.9). In addition, for the vertex model we used a honeycomb pattern for the initial spatial configuration, as in Figure 2.3.

The initial conditions for all simulations in this chapter were determined in this manner. The only difference when we use the uniform cell cycle model in Section 4.3 is that we sample initial G1 durations from the uniform distribution described by Equation (3.11).

Termination conditions If the population goes extinct in either the well-mixed or the vertex model, i.e. the population size reaches zero, the simulation terminates.

For simulations in Regime 3, it is impossible to verify population explosion through computation, since it is defined as indefinite population growth. Therefore, we specified a maximum cell count, $N_{\max} > 0$, which means that the simulation terminates as soon as the population size reaches $N_{\max}$. Simulations that end by hitting the maximum cell count are classified as "non-extinct". However, imposing such a maximum population size modifies the properties of the Markov model. In particular, the extinction probability is not the same as in Equation (4.2), so we need to account for this when comparing computational results to theoretical predictions.

To be exact, by imposing a maximum cell count for the population we obtain a birth–death Markov chain on the finite state space $I = \{0, 1, 2, \ldots, N_{\max}\}$ with transition probabilities

$$p_{0,0} = 1\,, \tag{4.4a}$$

$$p_{i,i-1} = 1 - \theta \qquad \text{for } 1 \leq i \leq N_{\max} - 1\,, \tag{4.4b}$$

$$p_{i,i+1} = \theta \qquad \text{for } 1 \leq i \leq N_{\max} - 1\,, \tag{4.4c}$$

$$p_{N_{\max},N_{\max}} = 1\,. \tag{4.4d}$$

Denoting the extinction probability of a population with initial size i for this Markov chain as $\tilde{h}_i$, it can be shown that $\tilde{h}_i$ is given by [201]

$$\tilde{h}_i = \begin{cases} 1 - \frac{i}{N_{\max}} & \text{for} \quad \theta = 1/2 \\ \dfrac{\left(\frac{1-\theta}{\theta}\right)^i - \left(\frac{1-\theta}{\theta}\right)^{N_{\max}}}{1 - \left(\frac{1-\theta}{\theta}\right)^{N_{\max}}} & \text{for} \quad \theta \neq 1/2 \end{cases}\,. \tag{4.5}$$

From Equation (4.5), it follows that for $\theta > 1/2$, we have $\tilde{h}_i \to h_i$ as $N_{\max} \to \infty$. In other words, if the maximum cell count is sufficiently large relative to the initial population size in Regime 3, then the difference between $\tilde{h}_i$ and h_i is negligible. We chose $N_{\max} = 50$ as the maximum cell count, such that the maximum absolute error $|\tilde{h}_i - h_i|$ over all the parameter values listed in Tables 4.1 and 4.2 is approximately $4.32 \cdot 10^{-7}$.

Because the Markov chain predicts that simulations in Regimes 1 and 3 should always terminate with these termination conditions, we set an unbounded simulation time for the well-mixed model. For the vertex model, since there is no functionality in Chaste for an unbounded simulation time, we set a relatively long simulation time of 100 000, which is equal to the average cell cycle time multiplied by a thousand.

Data processing For every simulation, we recorded the termination status, indicating whether the simulation ended in extinction or by hitting the maximum cell count. In addition, for the vertex model there is also the possibility that the simulation terminated by hitting the simulation end time. However, this did not occur for any vertex simulation, so we do not consider this case.

For each parameter set in Regime 3, we index the simulations using $k = 1, \ldots, N_{\text{sim}}$, where N_{sim} is the number of simulations per unique parameter set, and denote the termination status as

$$☠_k = \begin{cases} 1 & \text{if simulation } k \text{ went extinct} \\ 0 & \text{otherwise} \end{cases} . \tag{4.6}$$

We then estimated the extinction probability by its extinction frequency, defined as

$$\hat{h} = \frac{1}{N_{\text{sim}}} \sum_{k=1}^{N_{\text{sim}}} ☠_k . \tag{4.7}$$

In addition to computing the extinction frequency, we constructed 95% confidence intervals using a bootstrapping method. For each parameter set, we resampled the data by sampling with replacement from $☠_k$ to obtain a "bootstrap sample" of size N_{sim}. We repeated this process 10 000 times and computed the extinction frequency in each bootstrap sample. The resulting bootstrap extinction frequencies formed a bootstrapped distribution approximating the distribution of $\hat{h}$. We then used the quantiles of the bootstrapped distribution to construct confidence intervals.

Data visualisation We plot the extinction frequencies for Regime 3 in Figure 4.1 for the well-mixed model and Figure 4.2 for the vertex model, alongside the extinction probabilities

predicted in Equation (4.5). In addition, we plot the bootstrapped confidence intervals as error bars.

Computational materials and methods See Computational materials and methods in Section 2.3.1 (page 36) for the vertex model.

Well-mixed simulations are significantly cheaper to run, so we did not run them on a computational cluster. Instead, we used a regular multi-core computer to run well-mixed simulations in parallel. We wrote a script that executes a batch of simulations, with the range specified using command-line arguments. We parallelised simulation batches using GNU Parallel[2], an open-source command-line tool for parallelising shell commands [202]. The computer used to parallelise simulations had an Intel Core i5-9500T CPU with six cores and 16 GB of DDR4 random-access memory. All the well-mixed simulation suites in this thesis were executed in this manner.

4.2.2 Results

Well-mixed model We report that all well-mixed simulations in Regime 1 ended in extinction, thus validating the predictions in Regime 1 for the well-mixed model. Figure 4.1 shows that the observed extinction frequencies and predicted extinction probabilities are in very close agreement for all simulated parameter sets in Regime 3. Together with the observation that all simulations in Regime 1 ended in extinction, we conclude that the extinction probabilities in Equation (4.2) accurately predict the proliferative behaviour of the well-mixed constant death signal model.

Vertex model The Regime 1 simulation suite took 53 CPU hours to finish. Out of 900 vertex simulations in Regime 1, 18 simulations failed, which is a failure rate of 2%. The simulations failed because the mesh was invalidated, and not because the system ran out of memory (which would indicate population explosion), so we disregard the failed simulations. In addition, all successful simulations ended in extinction, which validates the predictions in Regime 1 for the vertex model.

The Regime 3 simulation suite took 69 CPU hours to finish. Out of 900 vertex simulations in Regime 1, eight simulations failed, which is a failure rate of 0.89%. After discarding the failed simulations, we plot the extinction frequency on the left-hand side of Figure 4.2. These plots show that the observed extinction frequency is consistently slightly higher than the predicted extinction frequency. This is because cells in the vertex model can die by

[2] `https://www.gnu.org/software/parallel/`

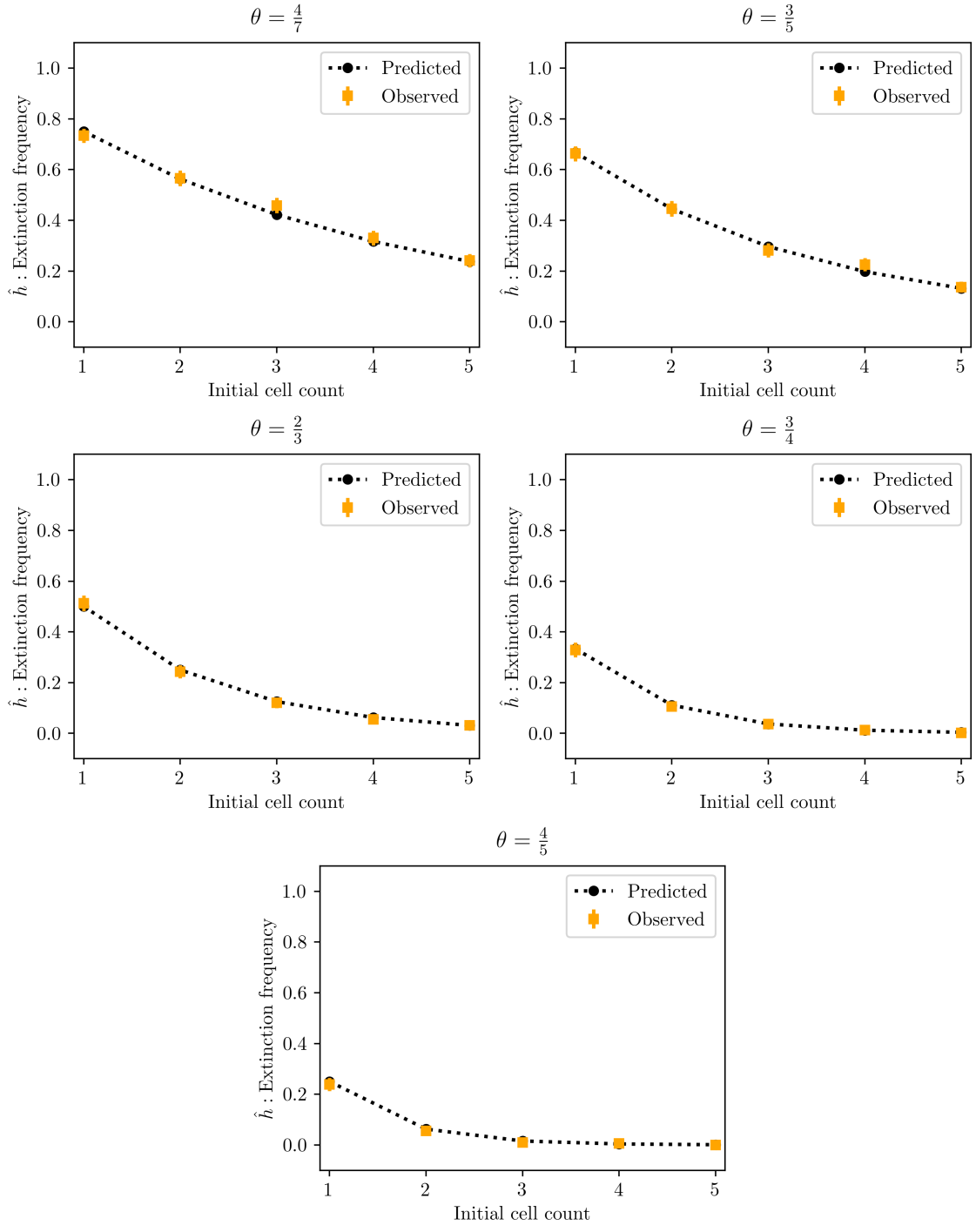

Figure 4.1: Estimated extinction frequency, $\hat{h}$, defined in Equation (4.7) for the well-mixed model. Error bars denote 95% confidence intervals obtained by bootstrapping. The predicted extinction probability, $\tilde{h}_i$, defined in Equation (4.5), is plotted for comparison.

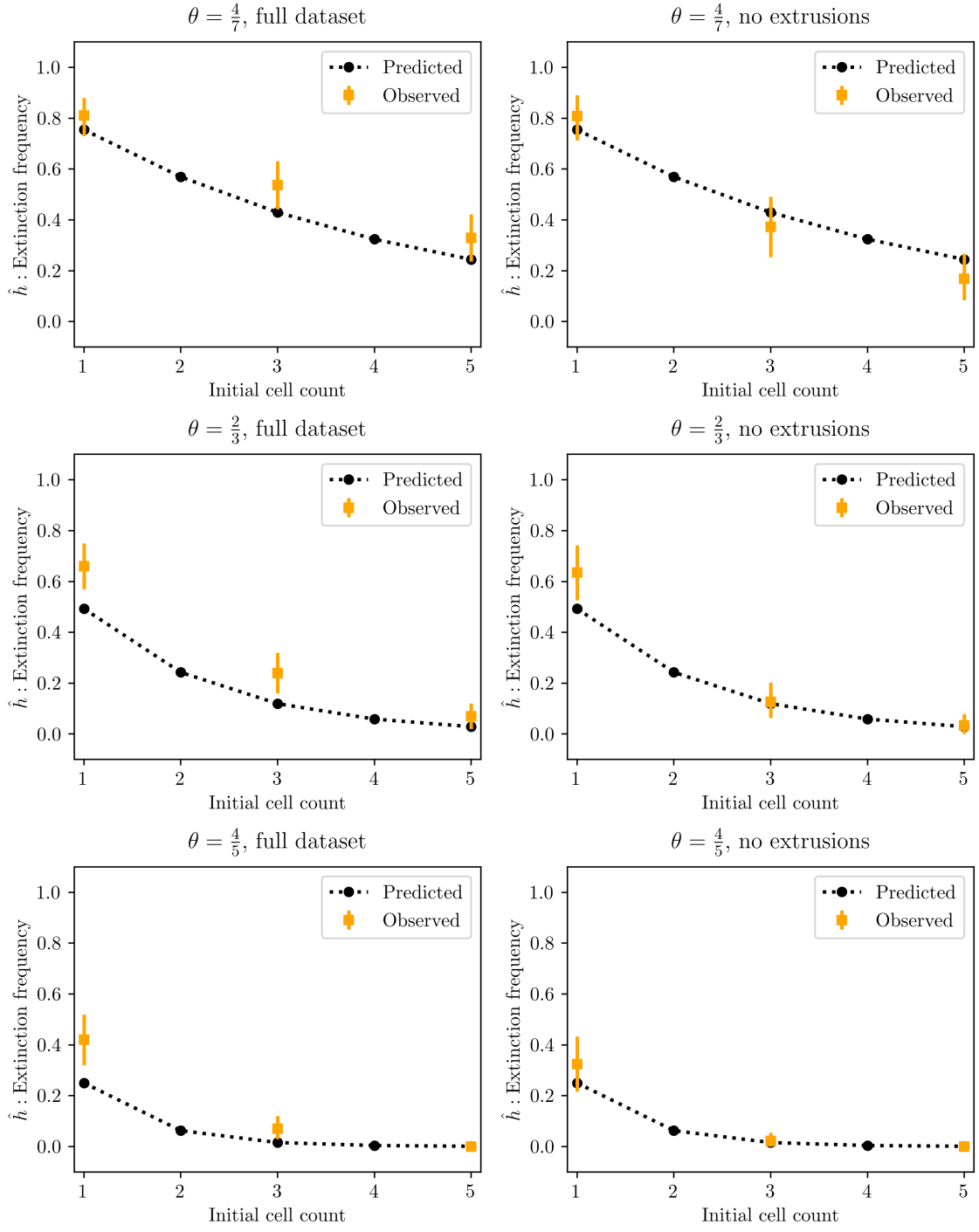

Figure 4.2: Estimated extinction frequency, $\hat{h}$, defined in Equation (4.7) for the vertex model including extrusions (left) and excluding extrusions (right). Error bars denote 95% confidence intervals obtained by bootstrapping. The predicted extinction probability, $\tilde{h}_i$, defined in Equation (4.5), is plotted for comparison.

extrusion in addition to apoptosis induced by the death clock, so the propensity for extinction is strictly higher than predicted by the Markov chain model.

We can demonstrate this effect by excluding simulations in which extrusions took place in order to separate out the impact of extrusions on the extinction frequency. Aggregating over all Regime 3 simulations, we see that extrusion events were far less common than apoptosis events, with roughly 1.2% of all deaths caused by extrusion. Out of 892 successful simulations, 184 simulations recorded at least one extrusion event. The right-hand plots in Figure 4.2 show the extinction frequency after excluding these simulations, and clearly indicate a better agreement with theoretical predictions.

4.3 Effective G1 duration

In Section 3.1.1, where we stated the biological assumptions of the death clock framework, we assumed that the death clock mechanism only initiates apoptosis in G1 phase. As a result, when a cell initiates apoptosis, the time it spends in G1 phase is truncated. The "effective G1 duration", defined as the time actually spent by the cell in G1 phase, is therefore shorter than the G1 duration, t^*, that was sampled autonomously from C. On the population level, the result is that the average effective G1 duration is shorter than t_{G1}. This discrepancy between the effective and the autonomous G1 durations depends on the death signal intensity; greater death signals trigger apoptosis sooner. Generally speaking, the death signal represents a non-autonomous process, so the truncation of the effective G1 duration is a non-autonomous effect.

In the general case, the death signal is not known *a priori*, making it difficult to find the distribution of the effective G1 duration. However, in this chapter we assume a constant death signal, which lets us predict the effective G1 duration. In this section, we first derive a general expression for the expected effective G1 duration, and then particular expressions for the exponential (Section 4.3.1) and uniform (Section 4.3.2) cell cycle models. Finally, we validate the analytical predictions with simulations.

In Section 3.2.1, we solved the death clock ODE for a constant death signal, resulting in Equation (3.16). We substitute the death threshold $T_\dagger$ in Equation (3.16) and rearrange to obtain the death time for a constant death signal:

$$t_\dagger = \frac{T_\dagger}{c}\,. \tag{4.8}$$

The death time represents a hard barrier that cells cannot cross while in G1 phase, and is thus an upper limit on the effective G1 duration.

The survival condition (Equation (3.15)) can be reformulated in terms of the death time $t_\dagger$; if $t^* < t_\dagger$, then the cell transitions to G2 phase before the death time and survives. In this case, the effective G1 duration is equal to the sampled G1 duration. However, if $t^* \geq t_\dagger$, the cell dies and the effective G1 duration is exactly equal to the death time. Given these are the only possible outcomes (ignoring death by extrusion), we express the effective G1 duration, denoted by t_{eff}, as

$$t_{\text{eff}} = \begin{cases} t^* & \text{for} \quad t^* < t_\dagger \\ t_\dagger & \text{for} \quad t^* \geq t_\dagger \end{cases} = \min(t^*, t_\dagger)\,. \tag{4.9}$$

In the case of survival (i.e. $t^* < t_\dagger$), the effective G1 duration t_{eff} is distributed as t^*, subject to the condition that $t^* < t_\dagger$. In other words, the effective G1 duration t_{eff} for surviving cells follows a truncated G1 duration distribution. Formally, if $\psi(t)$ and $\Psi(t)$ are the probability density function and cumulative distribution function of t^*, respectively, then the distribution of t_{eff} is given by the following probability density function:

$$\psi_{\text{eff}}(t) = \frac{\psi(t) \cdot \mathbb{1}(t < t_\dagger)}{\Psi(t_\dagger) - \Psi(0)} = \frac{\psi(t) \cdot \mathbb{1}(t < t_\dagger)}{\Psi(t_\dagger)}\,, \tag{4.10}$$

where $\mathbb{1}(\cdot)$ is the indicator function. In Equation (4.10), we also used the fact that $\Psi(0) = 0$, which is a result of Equation (3.2), which states that the support of $\Psi(t)$ is $[0, \infty)$. The expected value of the truncated G1 duration is then

$$E(t_{\text{eff}} \,|\, t^* < t_\dagger) = \int_0^\infty t' \psi_{\text{eff}}(t')\,\mathrm{d}t' = \frac{1}{\Psi(t_\dagger)} \int_0^{t_\dagger} t' \psi(t')\,\mathrm{d}t'\,. \tag{4.11}$$

In the case of death (i.e. $t^* \geq t_\dagger$), the effective G1 duration is exactly $t_\dagger$. In other words,

$$E(t_{\text{eff}} \,|\, t^* \geq t_\dagger) = t_\dagger\,. \tag{4.12}$$

We use the law of total expectation to find an expression for $E(t_{\text{eff}})$:

$$E(t_{\text{eff}}) = E(t_{\text{eff}} \,|\, t^* < t_\dagger) P(t^* < t_\dagger) + E(t_{\text{eff}} \,|\, t^* \geq t_\dagger) P(t^* \geq t_\dagger)\,. \tag{4.13}$$

Substituting Equation (4.11) and Equation (4.12), we have

$$E(t_{\text{eff}}) = \int_0^{t_\dagger} t' \psi(t')\,\mathrm{d}t' + t_\dagger(1 - \Psi(t_\dagger))\,, \tag{4.14}$$

where we used $P(t^* < t_\dagger) = \Psi(t_\dagger)$ and $P(t^* \geq t_\dagger) = 1 - \Psi(t_\dagger)$.

4.3.1 Exponential cell cycle model

For the exponential cell cycle model, we substitute Equations (3.9) and (3.10) in Equation (4.14),

$$E(t_{\text{eff}}) = \int_0^{t_\dagger} \frac{t'}{t_{\text{G1}}} \exp\left(-\frac{t'}{t_{\text{G1}}}\right) \mathrm{d}t' + t_\dagger \exp\left(-\frac{t_\dagger}{t_{\text{G1}}}\right), \tag{4.15}$$

and use integration by parts to find

$$\begin{aligned} E(t_{\text{eff}}) &= t_{\text{G1}}\left[1 - \exp\left(-\frac{t_\dagger}{t_{\text{G1}}}\right)\right] - t_\dagger \exp\left(-\frac{t_\dagger}{t_{\text{G1}}}\right) + t_\dagger \exp\left(-\frac{t_\dagger}{t_{\text{G1}}}\right) \\ &= t_{\text{G1}}\left[1 - \exp\left(-\frac{t_\dagger}{t_{\text{G1}}}\right)\right]. \end{aligned} \tag{4.16}$$

We plot the relationship between $E(t_{\text{eff}})$ and $t_\dagger$ in Figure 4.3. Treating $E(t_{\text{eff}})$ as a function of $t_\dagger$, we can approximate $E(t_{\text{eff}})$ for $t_\dagger \ll t_{\text{G1}}$ as

$$E(t_{\text{eff}}) \approx t_{\text{G1}} \frac{t_\dagger}{t_{\text{G1}}} = t_\dagger . \tag{4.17}$$

Hence, for small $t_\dagger$ the expected effective G1 duration closely follows the upper bound imposed by $t_\dagger$. On the other hand, in the limit of large $t_\dagger$, we have

$$\lim_{t_\dagger \to \infty} E(t_{\text{eff}}) = t_{\text{G1}} , \tag{4.18}$$

which is the same average effective G1 duration as a cell with no active death clock mechanism. We can conclude from these limiting regimes that the effect of the death clock is greatest when the death time is small, which corresponds to a large death signal. In this case, the effective G1 duration is dominated by the death time. On the other hand, when the death time is large, the distribution of the effective G1 duration approximates the autonomous G1 duration distribution.

4.3.2 Uniform cell cycle model

Because the support of the uniform distribution is limited to a finite interval, we distinguish between three distinct parameter regions based on the position of the death time relative to the uniform interval. If the death time is smaller than the lower bound of the uniform interval, then all sampled G1 durations are larger than the death time. Therefore, every cell is guaranteed to die at exactly $t_\dagger$. In other words,

$$t_\dagger < t_{\text{G1}} - \frac{1}{2} r \;\Rightarrow\; E(t_{\text{eff}}) = t_\dagger . \tag{4.19}$$

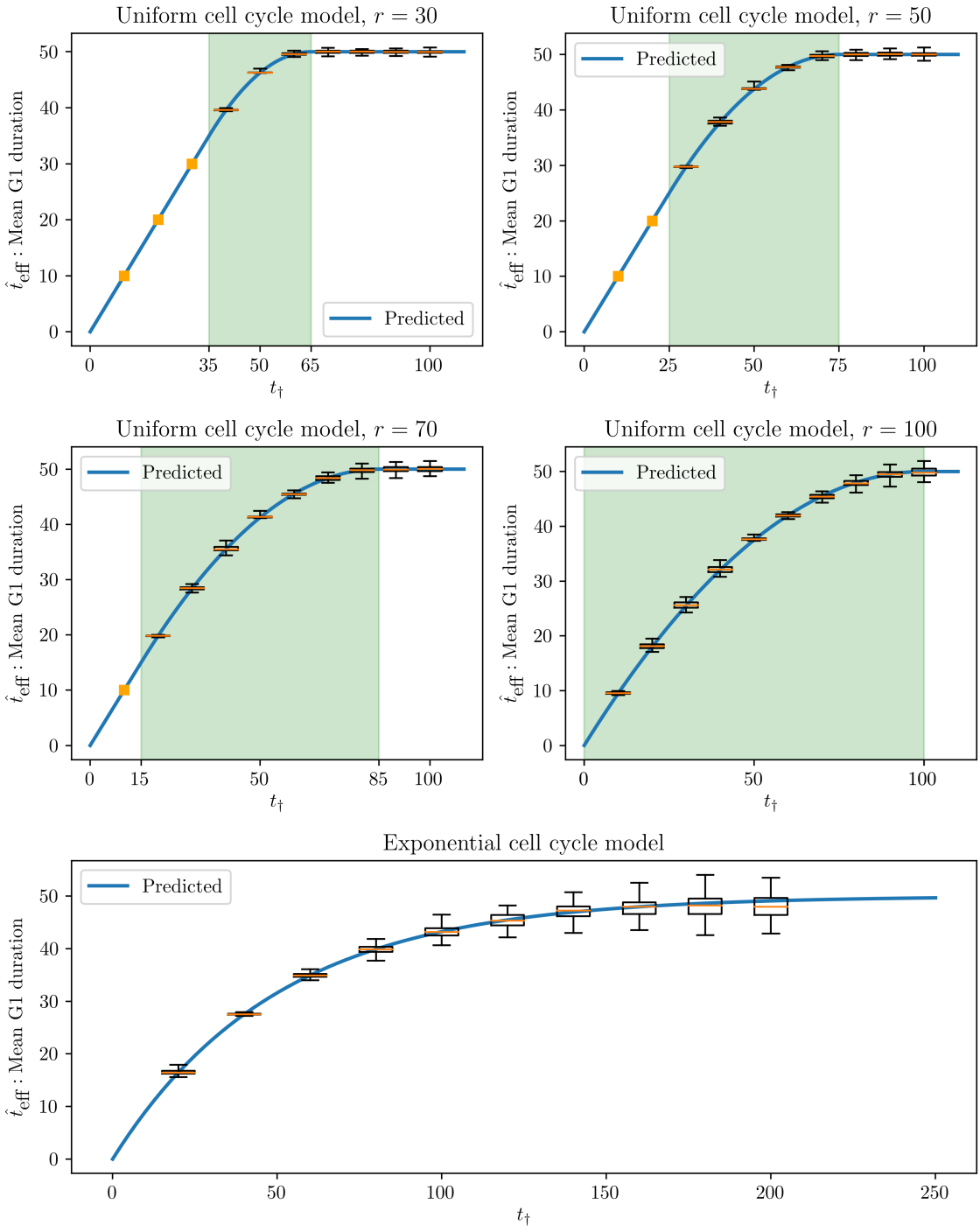

Figure 4.3: The predicted mean effective G1 duration is plotted using a solid line for the exponential (Equation (4.16)) and uniform (Equation (4.22)) cell cycle models. The distribution of the mean effective G1 duration, denoted $\hat{t}_{\text{eff}}$ and defined in Section 4.3.3, is plotted using box plots with whiskers extending to the minimum and maximum values. An orange square indicates that all observed mean effective G1 durations are the same. The green shaded area shows the range of the uniform interval.

On the other hand, if the death time is larger than the upper bound of the uniform interval, then cells are unaffected by the death clock because they never spend enough time in G1 phase to hit the death time. Hence, in this case the effective G1 duration is distributed identically as the sampled G1 duration, and its expected value is equal to t_{G1}. Formally,

$$t_\dagger > t_{\mathrm{G1}} + \frac{1}{2}r \;\Rightarrow\; E(t_{\mathrm{eff}}) = E(t^*) = t_{\mathrm{G1}}\,. \tag{4.20}$$

The non-trivial case occurs when the death time is within the uniform interval. We substitute the probability density function (Equation (3.11)) and the cumulative distribution function (Equation (3.13)) in Equation (4.14) to obtain

$$\begin{aligned} E(t_{\mathrm{eff}}) &= \int_{t_{\mathrm{G1}}-\frac{1}{2}r}^{t_\dagger} \frac{t'}{r}\,\mathrm{d}t' + t_\dagger\left[1 - \frac{t_\dagger - \left(t_{\mathrm{G1}} - \frac{1}{2}r\right)}{r}\right] \\ &= t_{\mathrm{G1}} - \frac{1}{2r}\left[t_\dagger - \left(t_{\mathrm{G1}} + \frac{1}{2}r\right)\right]^2\,. \end{aligned} \tag{4.21}$$

Combining all three cases, we have

$$E(t_{\mathrm{eff}}) = \begin{cases} t_\dagger & \text{for} \quad t_\dagger < t_{\mathrm{G1}} - \frac{1}{2}r \\ t_{\mathrm{G1}} - \frac{1}{2r}\left[t_\dagger - \left(t_{\mathrm{G1}} + \frac{1}{2}r\right)\right]^2 & \text{for} \quad t_\dagger \in \left[t_{\mathrm{G1}} - \frac{1}{2}r, t_{\mathrm{G1}} + \frac{1}{2}r\right] \\ t_{\mathrm{G1}} & \text{for} \quad t_\dagger > t_{\mathrm{G1}} - \frac{1}{2}r \end{cases}\,. \tag{4.22}$$

In Figure 4.3, we plot this relationship for selected values of r. Analysing the behaviour of $E(t_{\mathrm{eff}})$ as a function of $t_\dagger$, we find that the limiting regimes $t_\dagger \ll t_{\mathrm{G1}}$ and $t_\dagger \to \infty$ in the exponential cell cycle model are analogous to the cases $t_\dagger < t_{\mathrm{G1}} - r/2$ and $t_\dagger > t_{\mathrm{G1}} + r/2$ in the uniform cell cycle model, respectively. In the case of $t_\dagger < t_{\mathrm{G1}} - r/2$, the effective G1 duration is determined solely by the death time, whereas in the case of $t_\dagger > t_{\mathrm{G1}} + r/2$ the behaviour of cells is identical to cells without the death clock mechanism. In the interval bridging these two outer cases, i.e. for $t_\dagger \in [t_{\mathrm{G1}} - r/2, t_{\mathrm{G1}} + r/2]$, the expected effective G1 duration takes on the form of a quadratic function that smoothly connects the linear function in $t_\dagger < t_{\mathrm{G1}} - r/2$ with the constant function in $t_\dagger > t_{\mathrm{G1}} + r/2$.

4.3.3 Computational validation

In this section, we validate our predictions for the expected effective G1 duration for the exponential (Equation (4.16)) and the uniform (Equation (4.22)) cell cycle models. In particular, we used a Monte Carlo method to estimate $E(t_{\mathrm{eff}})$ by running repeated simulations

Table 4.3: Model and simulation parameter values used to estimate the mean effective G1 duration for the well-mixed model.

Parameter	Exponential	Uniform
t_{G1}		50
t_{G2}		50
c		1
$t_\dagger$	$20, 40, \ldots, 200$	$10, 20, \ldots, 100$
r	-	$30, 50, 70, 100$
Initial cell count		64
Simulation end time		∞
Maximum cell count		1 000
N_{sim}		100

for different values of $t_\dagger$ with a unique seed in each simulation.

Parameter choice We used the parameter values for $t_\dagger$ listed in Table 4.3 for the exponential and uniform cell cycle models. In addition, for the uniform cell cycle model, we also varied the uniform interval width r and took the Cartesian product over all parameters. The death threshold is determined by the other parameters as follows:

$$T_\dagger = t_\dagger c \,. \tag{4.23}$$

We ran 100 simulations for each unique parameter set, hence the total number of simulations is $(10 + 10 \times 4) \times 100 = 5\,000$. The parameters used span Regime 1 ($\theta < 1/2$) and Regime 3 ($\theta > 1/2$) for the exponential cell cycle model and all proliferation regimes (including $\theta = 1/2$) for the uniform cell cycle model. Hence, population extinction and explosion are both possible proliferation outcomes.

Initial conditions We set the initial population size to 64 to obtain a sufficiently large sample of effective G1 durations in the case of population extinction, and initialise the cells as described in the Initial conditions subsection of Section 4.2.1 (page 74).

Termination conditions We chose a large maximum cell count of 1 000 to get a sufficiently large sample of effective G1 durations in the case of population explosion.

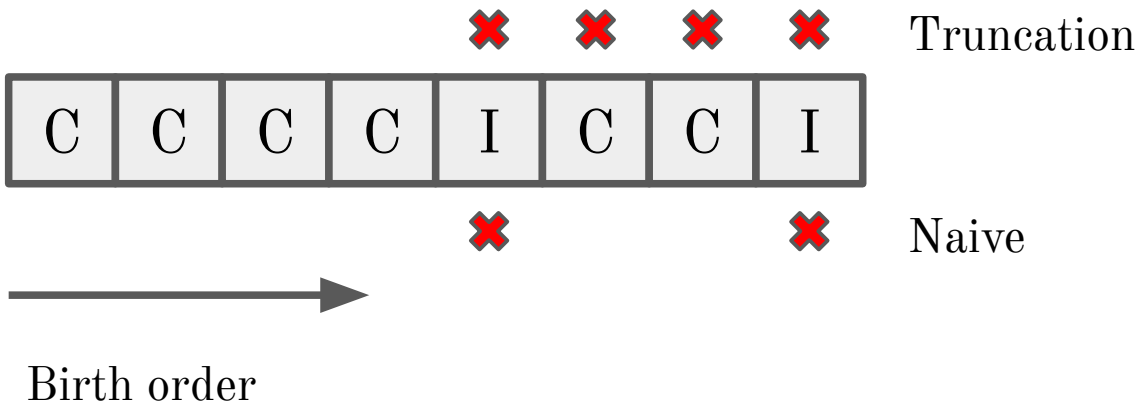

Figure 4.4: Effective G1 duration sampling method. The observed G1 durations are ordered by birth order of the corresponding cells. At the end of the simulation, the observed G1 durations can either be complete ("C") or incomplete ("I"). In the naive sampling method, we discard only incomplete G1 durations, which biases the sample towards smaller G1 durations. In the truncation method, we discard all observations starting from the first incomplete G1 duration. This leaves us with a smaller, but unbiased sample of G1 durations.

Data processing At the moment of termination there is generally a subset of cells that are still in G1 phase. These cells will have spent some time in G1 phase, but that is only a part of the total time that the cells would have spent in G1 phase if their cell cycles were allowed to complete. Therefore, the recorded times spent in G1 phase by these cells are not representative of the mean effective G1 duration and we should not include them in our sample.

However, we need to be careful when throwing out observations not to introduce biases. A naive sampling method would be to simply discard the incomplete G1 durations and retain all completed G1 durations. This has the potential, however, to bias the sample towards shorter G1 durations. The issue with the naive sampling method is that cells with relatively long effective G1 durations are more likely to still be in G1 phase at the end of the simulation, and are therefore overrepresented among the discarded observations.

We avoid this sampling bias by sorting the observed G1 durations by the birth order of the corresponding cells and including only completed G1 durations up to and excluding the first incomplete G1 duration in our sample. The remaining observations constitute a smaller, but pristine sample of effective G1 durations. For a visual comparison of these sampling methods, see Figure 4.4. Finally, to estimate $E(t_{\text{eff}})$, we take the mean of the sample thus obtained, which we denote as $\hat{t}_{\text{eff}}$.

Data visualisation The distribution of these $\hat{t}_{\text{eff}}$ is plotted using box plots in Figure 4.3, together with the theoretical predictions.

4.3.4 Results

Figure 4.3 shows a very close agreement between predictions and experiments, validating Equations (4.16) and (4.22). We note that for the uniform cell cycle model, there is no variation in the effective G1 duration for $t_\dagger < t_{\mathrm{G1}} - r/2$, as every cell dies at exactly $t_\dagger$ in this regime.

4.4 Discussion

In this chapter, we studied the simplest possible death clock model: the constant death signal model. Because the death signal is constant, the survival probability is the same for all cells in the population. We used this property to relate the asymptotic behaviour of the constant death signal model to that of a discrete-time birth–death Markov chain model. This enabled us to predict the extinction probability, a population-level property, from the survival probability, a property of individual cells, and the population size. In addition, we used the properties of the Markov chain to define distinct proliferation regimes, and validated the predictions in these regimes with well-mixed and vertex simulations. For the well-mixed model, the observed extinction frequencies were very closely aligned with predictions. For the vertex model on the other hand, the extinction frequency was slightly higher because mechanical extrusions resulted in a higher rate of cell death than predicted by the survival probability. We also noted the lack of other proliferative behaviours from the Markov chain description, such as oscillations or steady states, which are therefore not possible with a constant death signal.

As mentioned in the introduction, the constant death signal model does not represent intercellular interactions. We did not discuss heterotypic populations in this chapter for this reason, since two distinct cell types with a different constant death signal would also not interact with each other. Nonetheless, the simplicity of the constant death signal model enabled us to make theoretical predictions that we will use in the following two chapters. In the next chapter, we consider a death signal that is locally emitted by cells in G2 phase. Even though this is a non-autonomous death signal, we will see that it can usefully be approximated as constant, allowing us to build upon the results of this chapter.

Chapter 5

The G2 death signal model for homotypic populations

To recap, the overarching aim of this thesis is to study the mechanisms of cell competition without asserting *a priori* winner/loser status. In Chapter 2, we generated winners and losers in a vertex model by varying the mechanical parameters of heterotypic populations. However, homotypic simulations revealed that loser cells were eliminated due to their intrinsic nonviability, rather than through competitive interactions. A defining feature of cell competition, however, is that prospective loser and winner cells are both viable in homotypic environments.

A common observation in cell competition experiments is some form of programmed cell death, such as apoptosis, occurring in loser cells. This was missing from the model explored in Chapter 2, so we conjecture that an active mechanism of cell death is necessary for modelling cell competition. To test this hypothesis, we constructed the death clock modelling framework in Chapter 3, which proposes that cells accumulate death signals into an abstract quantity called the death clock and that they trigger apoptosis when the death clock reaches the death threshold.

One of the advantages of the death clock modelling framework is its flexibility. The death signal can take on many forms, representing different modes of intercellular communication. In Chapter 4, we examined the effects of a constant death signal because it is the simplest possible model. However, a constant death signal cannot represent cell competition because it is not based on intercellular interactions. Hence, in this chapter and the next, we examine the effects of a death signal based on short-range intercellular interactions: the G2 death signal (Section 5.1).

The ultimate goal is to investigate whether the G2 death signal model can produce emergent winners and losers in heterotypic populations. However, the precondition for cell

competition is that prospective winners and losers are homotypically viable. In this chapter we investigate the conditions for viability in homotypic populations. We leave the analysis of competition in heterotypic populations to Chapter 6.

The rest of this chapter is organised as follows: first, we introduce the G2 death signal in Section 5.1. Then, we propose that the system is ergodic in Section 5.2. This leads to a constant approximation for the G2 death signal that enables us to use the results of Chapter 4 to i) define the homotypic survival probability, ii) define distinct proliferation regimes for homotypic populations, and iii) map these proliferation regimes in parameter space. We characterise the survival probability and proliferation regimes for the exponential cell cycle model in Sections 5.3 and 5.4, and for the uniform cell cycle model in Sections 5.5 and 5.6. In Section 5.7 we summarise and discuss the results.

5.1 G2 death signal

We introduce our first death signal model for cell competition. It has been observed that elimination of loser cells occurs at a short range from winner cells [43, 53, 203]. Thus, in the G2 death signal model, we assume that cells communicate via short-range cell–cell interactions, for instance through binding surface receptors. This is an example of biochemical competition because signals are mediated by binding receptors, as opposed to mechanical competition where signals are transduced mechanically. Moreover, based on indications that cell competition may be linked to the cell cycle [186, 187], as discussed in Section 3.1.1, we assume that cells only emit death signals to their neighbours in G2 phase. The reason for choosing G2 phase is that cell competition is often manifested as proliferating cells inducing apoptosis in neighbouring cells in order to make room for themselves. In our framework, cells in G2 phase have already committed to division, which makes them a suitable candidate for generating death signals.

Concretely, we define the death signal identically for each cell as

$$f(t) = cg(t)\,, \tag{5.1}$$

where $g(t)$ is the proportion of neighbouring cells in G2 phase, i.e.

$$g(t) = \begin{cases} \dfrac{\#\text{ neighbours in G2}}{\#\text{ neighbours}} & \text{if \# neighbours} > 0 \\ 0 & \text{otherwise} \end{cases}\,. \tag{5.2}$$

In the well-mixed model, all cells are assumed neighbours. In the vertex model, on the other

hand, only cells that share a cell–cell boundary are considered neighbours. We normalise the number of neighbours in G2 phase in Equation (5.2) so that cells at the boundary of the tissue in the vertex model, who have fewer neighbours, do not receive an artificially lower death signal. Cells located in the interior of the tissue usually have a relatively constant number of neighbours, so the normalising factor does not fundamentally change the dynamics of the G2 death signal away from the boundary. Moreover, the normalisation also makes it easier to make useful approximations, as we will see in Section 5.2.

Assuming that cells only emit death signals in G2 phase creates a trade-off; cells in G1 phase are vulnerable to death signals and do not generate them, whereas cells in G2 phase are impervious to death signals but do generate them. This raises a question: what is the effect of changing the proportion of the cell cycle that is spent in G1 or G2 phase, given a fixed total cell cycle duration? In order to investigate this, we define t_G as the total cell cycle duration, and define $0 < \beta < 1$ as the fraction of t_G that is spent, on average, in the G1 phase. Formally,

$$t_{G1} = \beta t_G \,, \tag{5.3}$$

and

$$t_{G2} = (1 - \beta) t_G \,. \tag{5.4}$$

5.2 Ergodic approximation

Even without considering spatial structure, the G2 death signal model describes a highly unpredictable system. The cell cycle is stochastic because of the stochastic G1 duration, and therefore the death signal, which is based on G2 phase, is stochastic as well. Since the death signal is variable, the decision to commit apoptosis is also uncertain, and this uncertainty is fed back into the death signal. It is therefore difficult to characterise the death signal exactly, let alone the survival probability.

Despite this complexity, it turns out that, in many cases, the death signal is not only relatively stable, but even predictable. In particular, the system can often be regarded as **ergodic**, in the sense that the average proportion of cells in G2 phase relative to the population approximates the average proportion of the cell cycle spent in G2 phase. In other words, the spatial distribution of cell cycle phases in the population at any given moment is approximately the same as the distribution of cell cycle phases in the individual cell cycle over time. Conceptually, we state that the system is ergodic if, on average,

$$\frac{\text{\# cells in G2 phase}}{\text{\# cells in total}} = \frac{\text{G2 duration}}{\text{cell cycle duration}} \,. \tag{5.5}$$

If, in addition, we assume that the local cell neighbourhood is representative of the system as a whole, then we can approximate $g(t)$ as

$$g(t) \approx \frac{\text{\# cells in G2 phase}}{\text{\# cells in total}}\,. \tag{5.6}$$

Equation (5.6) expresses that the system is **well-mixed**. If cells are spatially segregated by cell cycle phase, for instance, then most cells are surrounded only by cells in the same phase, so Equation (5.6) does not hold. Following the assumptions in Equations (5.5) and (5.6), we find that

$$g(t) \approx \frac{\text{G2 duration}}{\text{cell cycle duration}} = \frac{t_{\mathrm{G2}}}{t_{\mathrm{G}}} = \frac{(1-\beta)t_{\mathrm{G}}}{t_{\mathrm{G}}} = 1-\beta\,. \tag{5.7}$$

This is the **ergodic approximation**. The death signal is then approximated as

$$f(t) = cg(t) \approx c(1-\beta)\,, \tag{5.8}$$

giving a constant death signal. By analogy with the constant death signal model, we define a death time,

$$t_\dagger = \frac{T_\dagger}{c(1-\beta)}\,, \tag{5.9}$$

and find an expression for the **homotypic survival probability** as

$$\lambda = \Psi(t_\dagger) = \Psi\left(\frac{T_\dagger}{c(1-\beta)}\right), \tag{5.10}$$

where we use λ to denote the homotypic survival probability and $\Psi(t)$ is the cumulative distribution function of the G1 duration distribution. We note here that varying β changes both the G2 and G1 durations, hence it affects both the death time and the distribution $\Psi(t)$ simultaneously. In the rest of this chapter, we use the terms "survival probability" and "homotypic survival probability" interchangeably.

A precondition for the ergodic assumption to hold is that the population size is sufficiently large for a population average to be meaningful. In a population of two cells, for instance, each cell senses only the other cell's death signal, such that the death signal can only be zero or unity. We also note that in the limit of a single cell the death signal is zero. In addition, the ergodic approximation assumes that cell cycles are uncorrelated across cells. The more that cells synchronise their cell cycles, the less that the population approximates an ergodic system. When these conditions are not met, we expect that Equation (5.10) is a poor predictor of cell behaviour.

Starting from an ideal ergodic cell population, i.e. a sufficiently large population with

completely uncorrelated cell cycles, we note that there are two important processes which may perturb the population's ergodicity. Firstly, when a cell divides, it produces two daughter cells in G1 phase. The cell cycles of sister cells are therefore always correlated following division. Secondly, when a cell dies, it does so in G1 phase. As a result, every cell death tips the population balance of cell cycle phases away from the G1 phase.

In the case of correlated sister cells, the G1 phase is overrepresented immediately after division, which pushes the death signal below its ergodic value. On the other hand, cell deaths result in an underrepresentation of the G1 phase, biasing the death signal upwards. Hence, divisions and deaths have opposing effects on the death signal. In the best-case scenario, divisions and deaths balance each other out, resulting in a stable ergodic population. In the worst-case scenario, divisions and deaths become separated in time such that the death signal varies wildly from extreme to extreme. As we will see in the rest of this chapter, the exponential and uniform cell cycle models give rise to the former and latter scenarios, respectively.

5.3 Homotypic survival probability for exponential cell cycle model

In this section, we derive an expression for the homotypic survival probability with the exponential cell cycle model under the ergodic approximation, and validate the predictions with simulations. First, we substitute Equation (5.3) into the cumulative distribution function for the exponential distribution, Equation (3.10), to get

$$\Psi(t) = 1 - \exp\left(-\frac{t}{\beta t_{\mathrm{G}}}\right). \tag{5.11}$$

We then substitute the death time from Equation (5.9) into Equation (5.11) to get the homotypic survival probability

$$\lambda(\beta) = 1 - \exp\left(-\frac{T_{\dagger}}{c t_{\mathrm{G}} \beta(1-\beta)}\right). \tag{5.12}$$

In order to simplify the notation, we introduce the dimensionless[1] parameter η,

$$\eta \equiv \frac{T_\dagger}{ct_\mathrm{G}}\,, \tag{5.13}$$

which is a non-negative constant that can be interpreted as a normalised death time or death threshold. We can thus write the survival probability as a function of two dimensionless parameters:

$$\lambda(\beta, \eta) = 1 - \exp\left(-\frac{\eta}{\beta(1-\beta)}\right). \tag{5.14}$$

As Figure 5.1 indicates, the survival probability with respect to β is symmetric and U-shaped. In particular, the survival probability is at its minimum for $\beta = 1/2$, and increases as β approaches zero or one. This illustrates the balance of two opposing effects taking place when varying β. When the G1 phase is long, i.e. for large β, there is a large window in which cells are vulnerable to death signals, but fewer cells in G2 phase to broadcast death signals. On the other hand, when the G2 phase is relatively long, i.e. for small β, there are more cells in G2 phase and a stronger death signal is felt, but the G1 phase is shorter, enabling cells to escape apoptosis. When the G1 and G2 durations are equal, these effects combine to create the fiercest competition.

5.3.1 Computational validation

We now explore whether the homotypic survival probability predicted by Equation (5.14) matches the behaviour of the model in simulations. We used a Monte Carlo method where we ran repeated simulations of our model for different values of η and β to sample the resulting survival frequencies. As our model is stochastic, we provided a unique seed for the random number generator in each simulation.

Parameter choice We used the well-mixed model with the parameters listed in Table 5.1 to validate the predicted homotypic survival probability in Equation (5.14). We chose six different values of η in order to span a large range of survival probabilities, both above and below $1/2$. For each value of η, we varied β from 0.1 to 0.9 in increments of 0.1. The death threshold was determined as a function of the other parameters, specifically by rewriting

[1] To see that η is dimensionless, we note that the death threshold, $T_\dagger$, has the same dimension as the death clock $\tau(t)$. Moreover, from the definition of the G2 death signal, Equation (5.1), it follows that c has the same dimension as the death signal $f(t)$. Furthermore, the death clock ODE, Equation (3.4), shows that the death clock divided by the death signal has the dimension of time. Finally, the total cell cycle time, t_G, also has the dimension of time, so η is dimensionless.

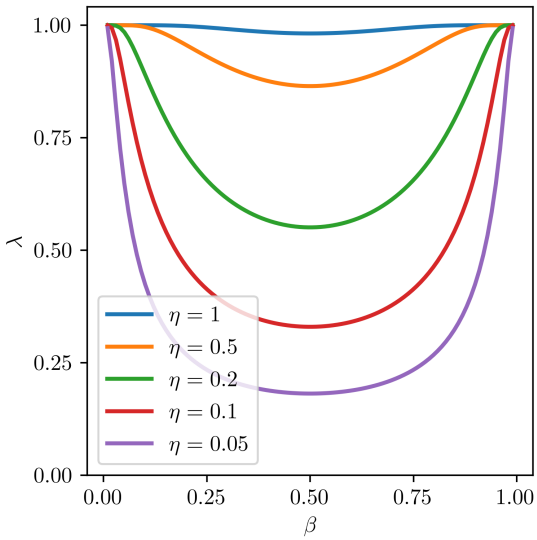

Figure 5.1: Homotypic survival probability with the exponential cell cycle model under the ergodic approximation as predicted by Equation (5.14).

Table 5.1: Model and simulation parameter values used to estimate the homotypic survival probability for the exponential cell cycle model.

Parameter	Well-mixed	Vertex
t_G		100
c		1
η		$0.05, 0.1, 0.2, 0.5, 1$
β	$0.1, 0.2, \ldots, 0.9$	$0.2, 0.5, 0.8$
Initial cell count		100
Simulation end time	∞	100 000
Minimum cell count		10
Maximum cell count		1 000
N_{sim}	100	20

Equation (5.13) as

$$T_\dagger = \eta c t_\mathrm{G} \,. \tag{5.15}$$

For the vertex model, we used the parameters listed in Table 5.1. Since the vertex model is computationally more expensive, we limited ourselves to a small sample of β values, specifically 0.2, 0.5, and 0.8. The mechanical parameters used in the vertex model for all simulations in this chapter are set to the default values given in Chapter 2.

For the well-mixed model, we ran 100 simulations for every unique parameter set, so the total number of simulations is $5 \times 9 \times 100 = 4\,500$. For the vertex model, we ran 20 simulations for every unique parameter set. Hence, the total number of simulations is $5 \times 3 \times 20 = 300$.

Initial conditions For the initial conditions, we need to assign an initial value for the death clock, τ_α, a birth time, t_α^0, and a G1 duration, t_α^*, to every cell α. For each cell, we initialised the death clock to zero and sampled the birth time uniformly from the interval $[-t_{\mathrm{G1}} - t_{\mathrm{G2}}, 0]$. To assign a G1 duration, ideally we would simply sample the G1 duration from the exponential distribution described by Equation (3.9), since this is how G1 durations are sampled during the simulation. However, in conjunction with the previously assigned birth time, this may result in a violation of the division invariant in Equation (3.34). Concretely, this happens when the total cell cycle duration (G1 duration plus G2 duration) is less than the age of the cell at the start of the simulation. To avoid this problem, for each cell we iteratively sampled G1 durations from the exponential distribution until the division invariant was respected.

The initial conditions are determined similarly for the vertex model, with the exception of the G1 durations. Unlike the well-mixed model, the vertex model does not have a division invariant. Instead, Chaste checks at every timestep whether the cell age is greater than its cell cycle duration, and performs a division if so. As a result, the cell cycle duration is allowed to be smaller than the cell age at the start of the simulation, it just means that the cell is overdue for division and will divide as soon as the simulation starts (see top simulation in Figure 5.2). Hence, the G1 duration is simply sampled from the exponential distribution, without any checks comparing the cell cycle duration and the cell age. Furthermore, we use a honeycomb pattern for the initial spatial configuration as shown in Figure 5.2. Because we are simulating a homotypic population, all cells are of the same type and we do not need to label them.

The biological interpretation of these initial conditions is that the cells were reproducing without exchanging death signals prior to the simulated time period, and that we "turn on" the exchange of death signals at the start of the simulation. The initial conditions for all

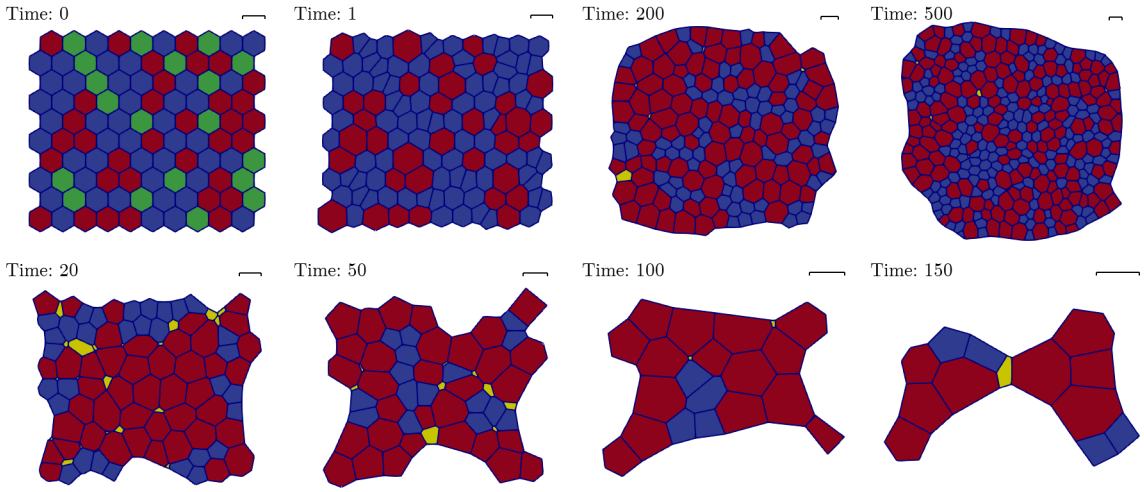

Figure 5.2: Snapshots from simulations of the vertex model with the exponential cell cycle model. Each row is one simulation and cells are coloured according to their state. ■: G1 phase. ■: G2 phase. ■: mitotic. ■: apoptotic. **Top** ($\eta = 0.2$, $\beta = 0.5$): at the start of the simulation, cells are either in G1 phase, G2 phase, or mitotic. The mitotic state indicates a cell that is overdue for division. The snapshot $t = 1$ shows that mitotic cells divide at the start of the simulation. The cell population grew with few apoptotic cells until the maximum cell count was reached at $t = 1032$. **Bottom** ($\eta = 0.05$, $\beta = 0.5$): the cell population declined due to a high rate of apoptosis until the minimum cell count was reached at $t = 158$.

simulations in this chapter were determined in this manner. The only difference when we use the uniform cell cycle model in later sections is that we sample initial G1 durations from the uniform distribution described by Equation (3.11).

Termination conditions As mentioned before, the population cannot go extinct in the G2 death signal model. Therefore, in order to ensure that the simulation terminates, we set a minimum cell count of 10. In addition, we impose a maximum cell count of 1 000 so that the simulation also terminates in the case of population explosion. When the population reaches the minimum or maximum cell count, the simulation immediately terminates.

For the well-mixed model we set an unbounded simulation time. Therefore, the simulation can only terminate by hitting the minimum or maximum cell count. Figure 5.3 shows the cell counts of two well-mixed model simulations that terminated by hitting the minimum and maximum cell count, respectively. On the other hand, there is no functionality in Chaste for an unbounded simulation time, so we set a simulation time of 100 000, which is long relative to the total cell cycle time of $t_G = 100$.

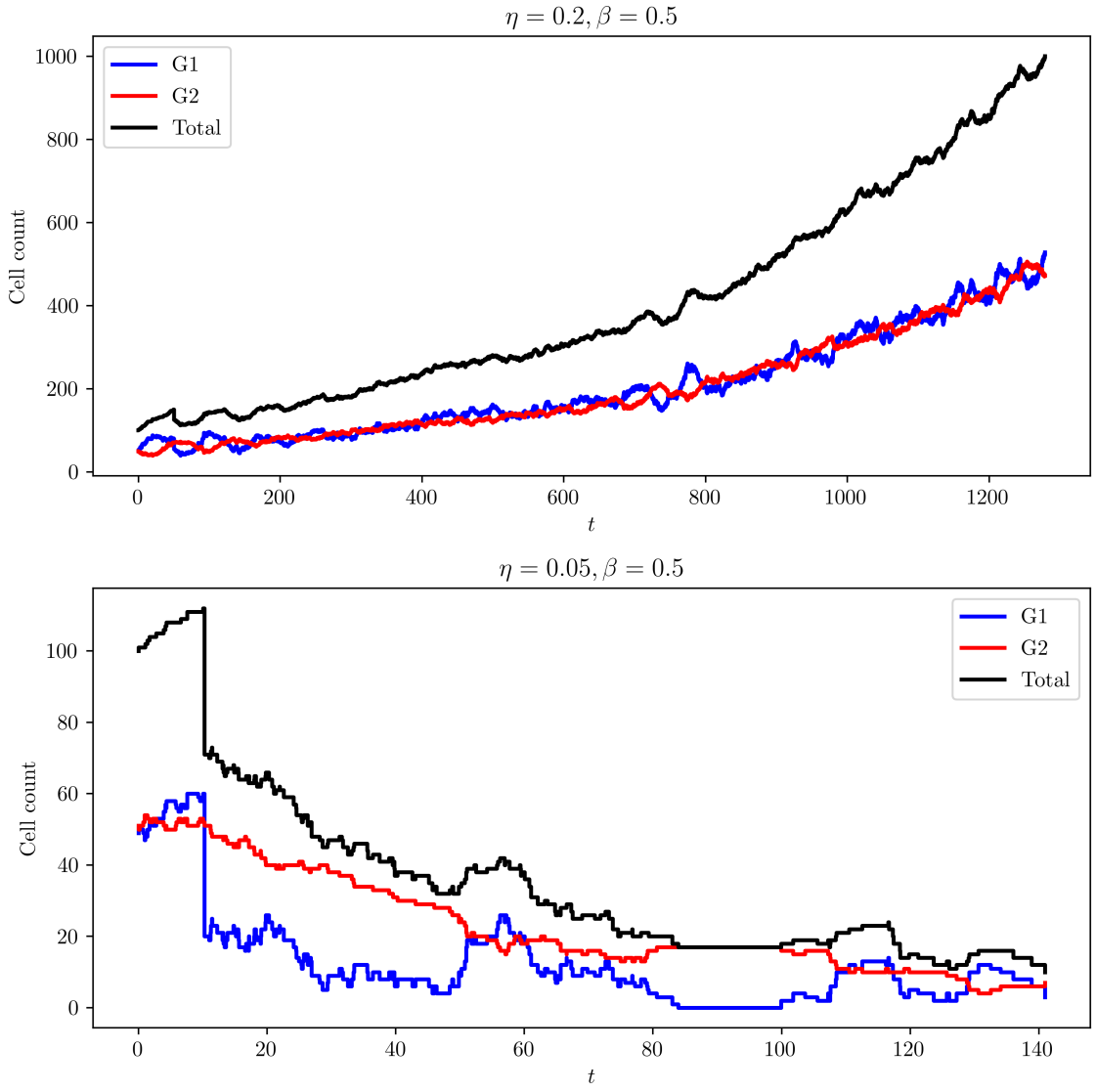

Figure 5.3: Cell counts for simulations of the well-mixed model with the exponential cell cycle model. The two sets of parameters η and β are the same as in Figure 5.2 for the respective rows. The top and bottom figures show simulations that terminated by hitting the maximum and minimum cell counts, respectively.

Data processing For each simulation, we computed the survival frequency, denoted by $\hat{\lambda}$, using Equation (2.12). For the well-mixed model, cells can only die through apoptosis, so all death events are apoptosis events. However, cells in the vertex model can be eliminated by a T2 swap despite not being in an apoptotic state. If a cell is apoptotic when it is extruded, we classify it as an apoptosis event, and if a cell is not apoptotic, we classify it as an "extrusion" event. We count both apoptosis and extrusion events as death events.

Data visualisation The distribution of observed survival frequencies is plotted against theoretical predictions in Figure 5.4 for the well-mixed model and Figure 5.5 for the vertex model.

Computational materials and methods See Computational materials and methods in Section 2.3.1 (page 36) for the vertex model and in Section 4.4 (page 77) for the well-mixed model.

5.3.2 Results

Well-mixed model Since our predictions are based on an ergodic approximation, we do not expect the computational results to match the predictions exactly. In particular, when there are relatively many deaths (i.e. at low survival probabilities), the number of cells in G1 phase will decrease, since they are susceptible to apoptosis. As a result, the proportion of cells in G1 phase will be skewed, and the ergodic assumption violated. Therefore, we expect that the simulations will deviate from the predictions when the survival frequency is low. In addition, for low cell counts the death signal fluctuates very strongly, making the ergodic approximation less tenable.

Figure 5.4 generally shows a good agreement between theory and observations, validating the ergodic approximation. The deviation from predictions is greatest when the survival frequency is relatively small, particularly when it drops below $1/2$. This is because the increased rate of apoptosis decreases the population size, such that the ergodic assumption no longer holds.

Vertex model Figure 5.5 generally shows a good agreement between theory and observations, validating the ergodic approximation for the vertex model. For high η values, there are some noticeable discrepancies from predictions. In particular, the survival frequency is lower than expected and shows a decreasing trend with respect to β for $\eta = 1.0$. We can explain this as follows. A high survival frequency means a high proliferation rate, which causes cell crowding and cell extrusion. Moreover, because extrusion events are counted

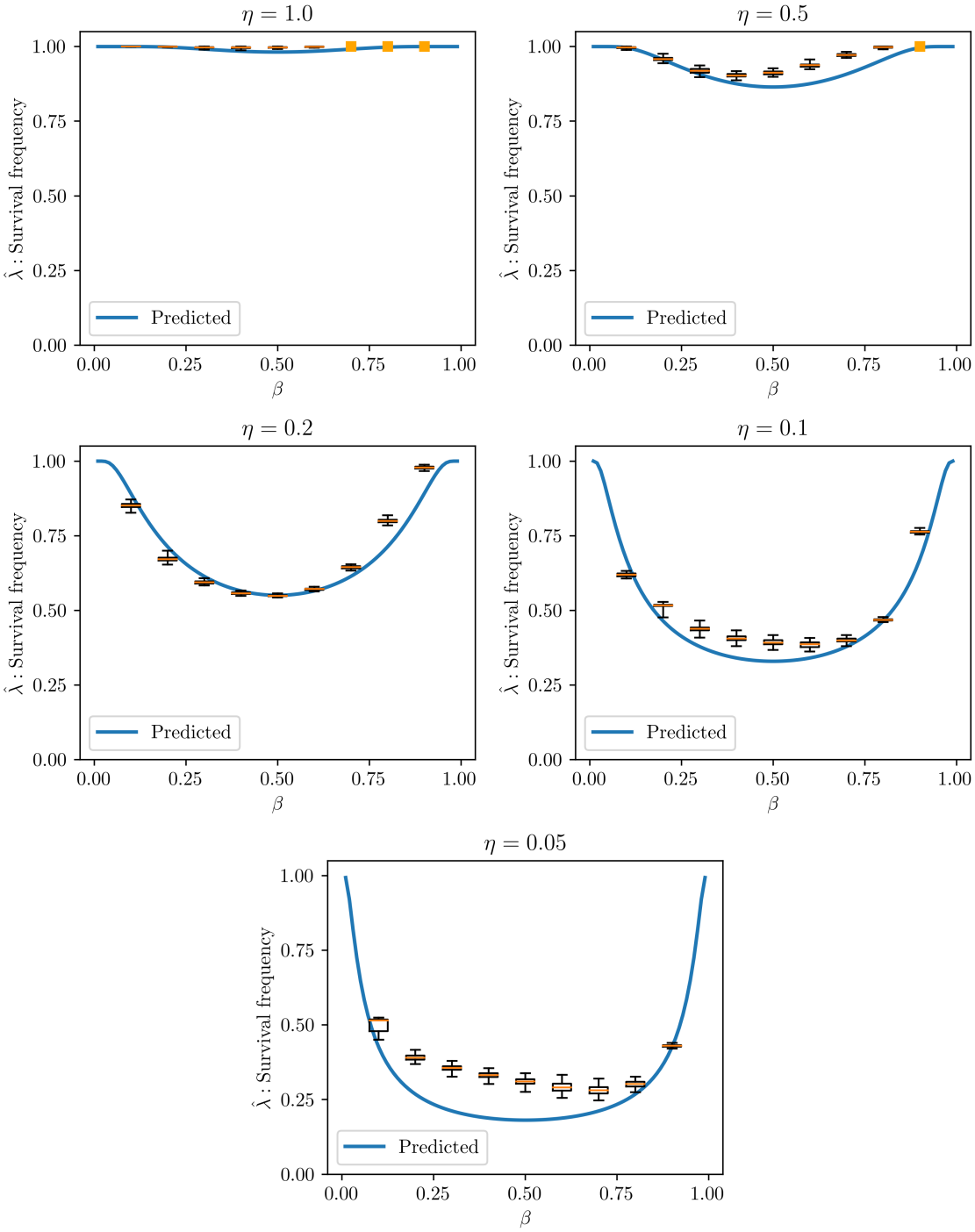

Figure 5.4: Estimated homotypic survival frequency, $\hat{\lambda}$, defined in Equation (2.12), for well-mixed simulations with the exponential cell cycle model. The distribution of the survival frequency is plotted using box plots with whiskers extending to the minimum and maximum values. An orange square indicates that all observed survival frequencies are the same. The predicted homotypic survival probability, defined in Equation (5.14), is plotted for comparison.

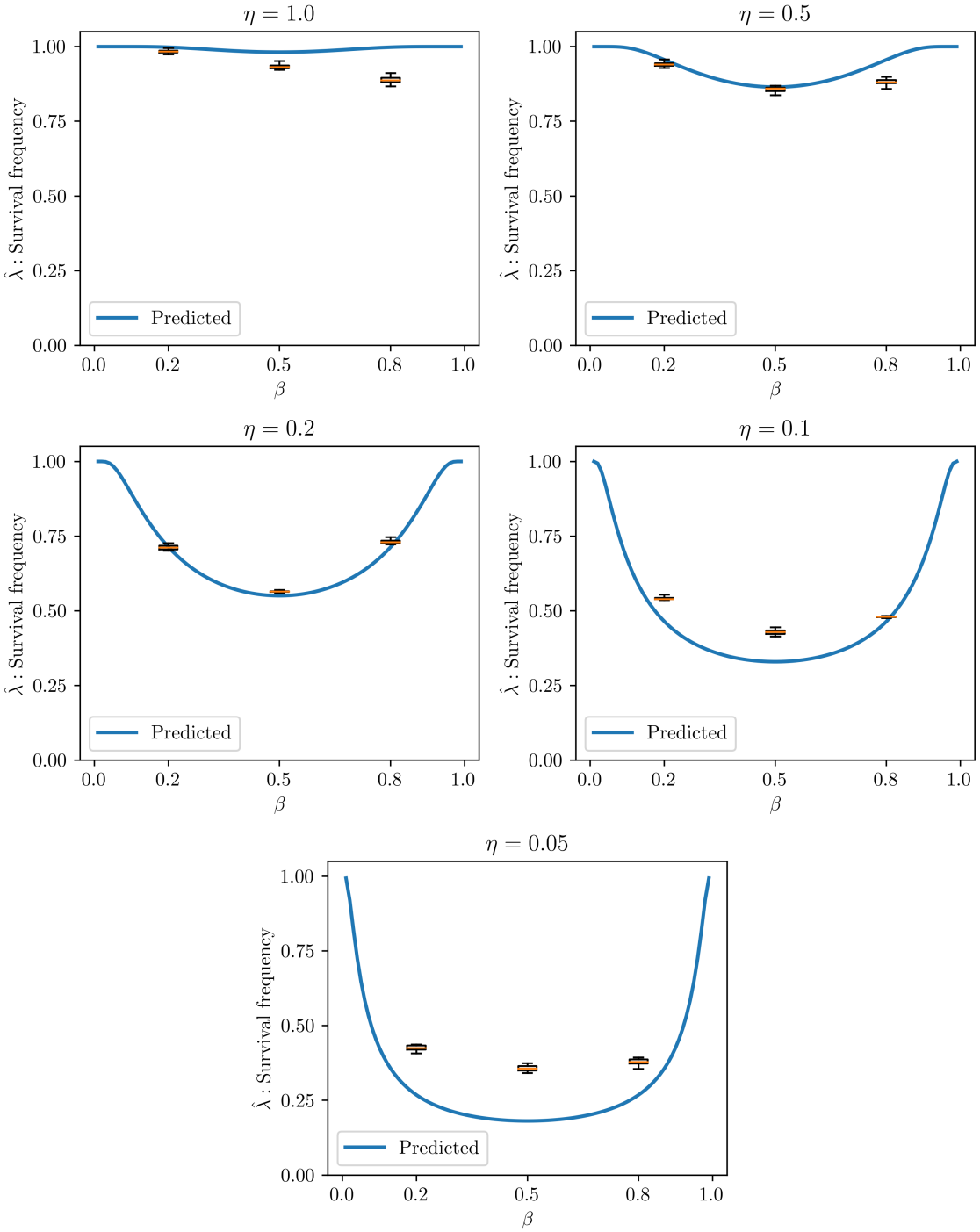

Figure 5.5: Estimated homotypic survival frequency, $\hat{\lambda}$, defined in Equation (2.12), for vertex simulations with the exponential cell cycle model. The distribution of the survival frequency is plotted using box plots with whiskers extending to the minimum and maximum values. An orange square indicates that all observed survival frequencies are the same. The predicted homotypic survival probability, defined in Equation (5.14), is plotted for comparison.

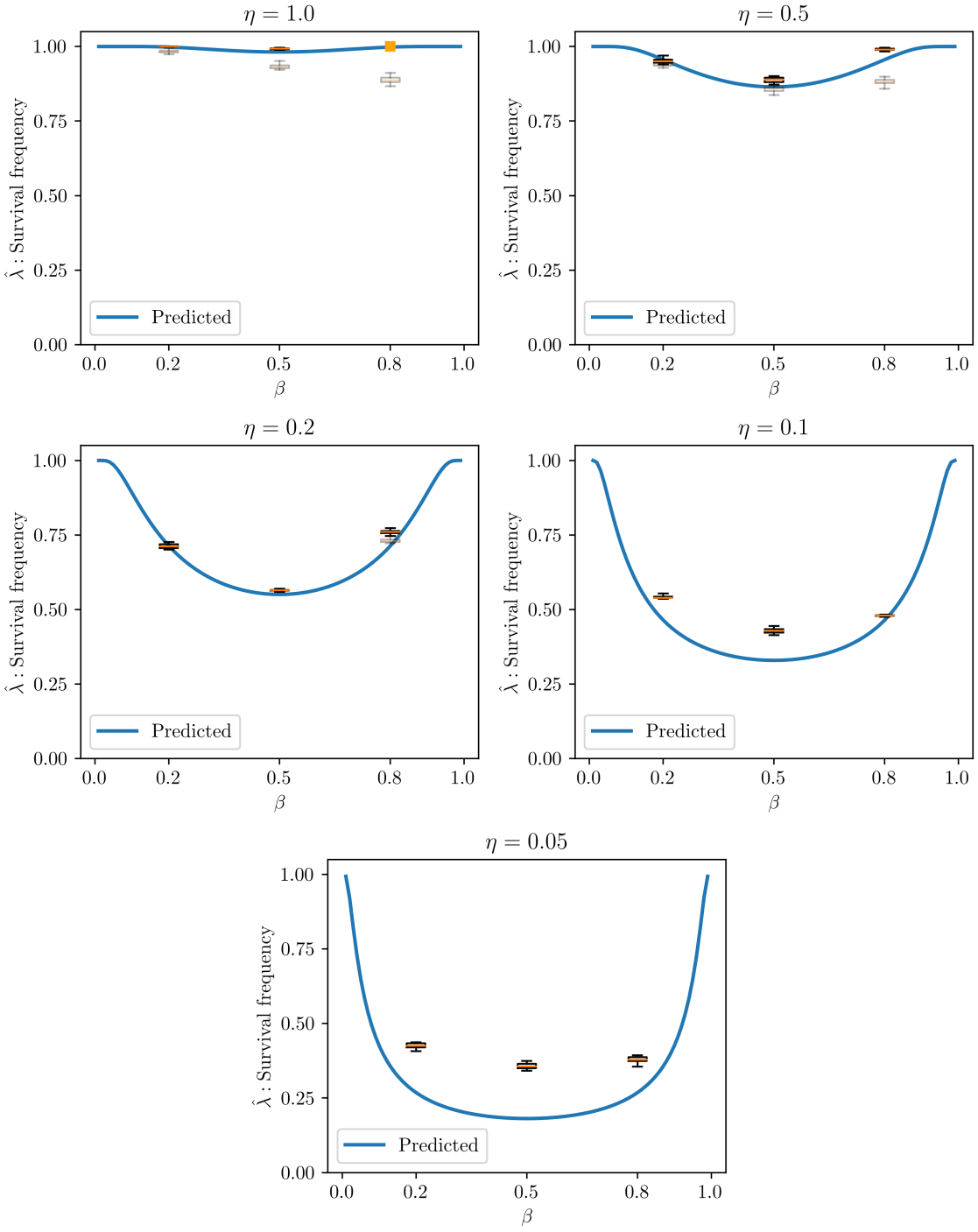

Figure 5.6: Estimated homotypic survival frequency, $\hat{\lambda}$, defined in Equation (2.12), for vertex simulations with the exponential cell cycle model without counting extrusions as death events. The results when counting extrusions as death events have been included for reference but made partially transparent.

as death events, the survival frequency is lower than what we expect from the death clock mechanism alone. This is demonstrated in Figure 5.6, where we presented the survival frequency calculated *without* extrusion events. The figure shows that the observed and predicted survival frequencies are in better agreement when we exclude extrusion events. Nonetheless, we will continue counting extrusion events as death events because the effect only occurs in cell crowding conditions and hence does not affect the overall viability of the population.

5.4 Proliferation regimes for exponential cell cycle model

In the previous chapter, we predicted different proliferation regimes based on the value of the survival probability for the constant death signal model (Section 4.2). This suggests that we can similarly use the homotypic survival probability to identify the proliferation regimes for a homotypic population with a G2 death signal. In this section, we define the proliferation regimes for the G2 death signal, characterise them for the exponential cell cycle model in particular, and validate the predictions with computational experiments.

We define two proliferation regimes based on the homotypic survival probability and the proliferation regimes for a constant death signal (Section 4.2):

Nonviable Regime $\left\{\lambda \leq \frac{1}{2}\right\}$ Cells are equally or more likely to die than to proliferate, hence the population declines. We say that cell types in this regime are **nonviable**.

Viable Regime $\left\{\lambda > \frac{1}{2}\right\}$ Cells are more likely to proliferate than to die, hence the population grows. We say that cell types in this regime are **viable**.

We collapse Regimes 1 and 2 into the Nonviable Regime because they both predict population decline. The Viable Regime corresponds to Regime 3, where we expect populations to grow unboundedly.

The proliferative behaviour in the Nonviable Regime deserves closer attention. In particular, as mentioned in Section 5.2, the ergodic approximation requires that there are enough cells to obtain a representative sample of the death signal. However, as a population shrinks, this stops being the case. Take the limiting case of a single cell for instance; it has no neighbours and therefore there is no G2 death signal to induce cell death. This also means that population extinction is impossible, in contrast to the constant death signal model. As the population grows back from a single cell, eventually the population will reach a size where the ergodic approximation holds again, so we expect the population to decline afresh. Therefore, we expect oscillatory behaviour for small population sizes in

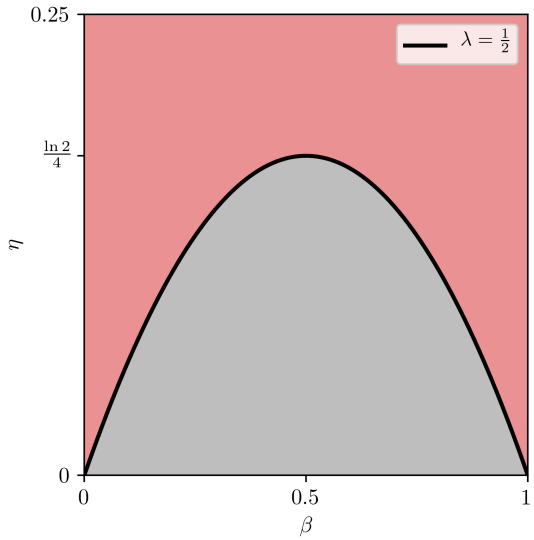

Figure 5.7: Diagram of proliferation regimes for the exponential cell cycle model. The homotypic viability curve is given by Equation (5.17). : Viable Regime. : Nonviable Regime.

the Nonviable Regime. We consider these oscillations an artefact of our model, however, and not representative of a biological process. Instead, this regime corresponds to a tissue that cannot grow or maintain its population size, i.e. the tissue is nonviable, as the name suggests.

We define the **homotypic viability curve** as the curve satisfying

$$\lambda = \frac{1}{2}\,. \tag{5.16}$$

This is the curve that separates the Nonviable Regime from the Viable Regime. For the exponential cell cycle model in particular, we substitute Equation (5.14) in Equation (5.16) to obtain the following homotypic viability curve:

$$\eta = \ln(2)\beta(1-\beta)\,. \tag{5.17}$$

Using this curve, we demarcate the proliferation regimes for the exponential cell cycle model in Figure 5.7. We see that for $\eta > \ln(2)/4$ all cell types are viable. Moreover, for $\eta \leq \ln(2)/4$ cell types are viable for extreme values of β and nonviable otherwise.

Table 5.2: Model and simulation parameter values used to characterise the proliferation regimes for the exponential cell cycle model.

Parameter	Well-mixed	Vertex
t_G		100
c		1
η	$0.01, 0.02, \ldots, 0.25$	$0.02, 0.04, \ldots, 0.24$
β	$0.05, 0.10, \ldots, 0.95$	$0.1, 0.2, \ldots, 0.9$
Initial cell count		100
Simulation end time		10 000
Minimum cell count		10
Maximum cell count		1 000
N_{sim}	50	20

5.4.1 Computational validation

In order to validate our predictions of the proliferation regimes for the exponential cell cycle model, we ran a parameter sweep over a regular grid in (β, η)-space and estimated the average survival frequency. For every unique parameter set we ran repeated Monte Carlo simulations using a unique seed in each simulation.

Parameter choice We sampled parameters from the area displayed in Figure 5.7, i.e. $\eta \in [0, 0.25]$ and $\beta \in [0, 1]$, which covers the entire homotypic viability curve. We discretised the parameter space into a regular grid of η- and β-values. Concrete parameter values for the well-mixed and vertex models are listed in Table 5.2. We used a coarser grid for the vertex model than the well-mixed model because of the difference in computational cost. In addition, we fixed $t_G = 100$ and $c = 1$, and computed the death threshold using Equation (5.15).

For the well-mixed model, we ran 50 simulations for every unique parameter set, so the total number of simulations is $25 \times 19 \times 50 = 23\,750$. For the vertex model, we ran 20 simulations for every unique parameter set. Hence, the total number of simulations is $12 \times 9 \times 20 = 2\,160$.

Initial conditions See Initial conditions in Section 5.3.1 (page 95).

Termination conditions We set minimum and maximum cell counts of 10 and 1 000, respectively. In addition, because the parameter sweep is relatively large, we set a simulation

end time of 10 000 to limit the computational time that each simulation could take.

Data processing As described in Section 5.3.1 (Data processing, page 98), we computed the survival frequency for each simulation k, denoted as $\hat{\lambda}_k$, using Equation (2.12). However, here we are not interested in the distribution of survival frequencies for a given parameter set, but instead we want to characterise the variation in survival frequency across (β, η)-space. Hence, for every unique parameter set, we averaged the homotypic survival frequency as

$$\overline{\lambda} = \frac{1}{N_{\text{sim}}} \sum_{k=1}^{N_{\text{sim}}} \hat{\lambda}_k \,, \tag{5.18}$$

where N_{sim} is the number of simulations for the given parameter set.

Data visualisation For $\hat{\lambda}_k < 1/2$, there were fewer division than death events in the simulation, which means that the population declined. Conversely, if $\hat{\lambda}_k > 1/2$, the population grew. Nonviable populations are more likely to have a survival frequency below $1/2$, and vice versa for viable populations. Figure 5.7 predicts that cell types below the homotypic viability curve are nonviable and cell types above the homotypic viability curve are viable.

In order to verify these predictions, we visualised $\overline{\lambda}$ using heat maps with β along the horizontal axis and η along the vertical axis. Since the critical value $1/2$ separates nonviable from viable populations, we chose a diverging colour map[2] with range $[0, 1]$ and central value $1/2$. In particular, we use the *seismic* colour map from Matplotlib [162], such that blue regions correspond to $\overline{\lambda} < 1/2$, red regions correspond to $\overline{\lambda} > 1/2$, and the border is white. In addition, we plot the theoretical homotypic viability curve to compare our results to our predictions. The results are shown in Figure 5.8 for the well-mixed and vertex models.

Computational materials and methods See Computational materials and methods in Section 2.3.1 (page 36) for the vertex model and in Section 4.4 (page 77) for the well-mixed model.

5.4.2 Results

Well-mixed model Figure 5.8 shows that the observed border between nonviable and viable regimes follows predictions very well. For small η values, the results are asymmetrical with respect to β, with lower survival frequencies for $\beta > 1/2$ than $\beta < 1/2$. The bottom plot

[2] A diverging colour map accentuates the central value by a peak in lightness and a colour transition [204].

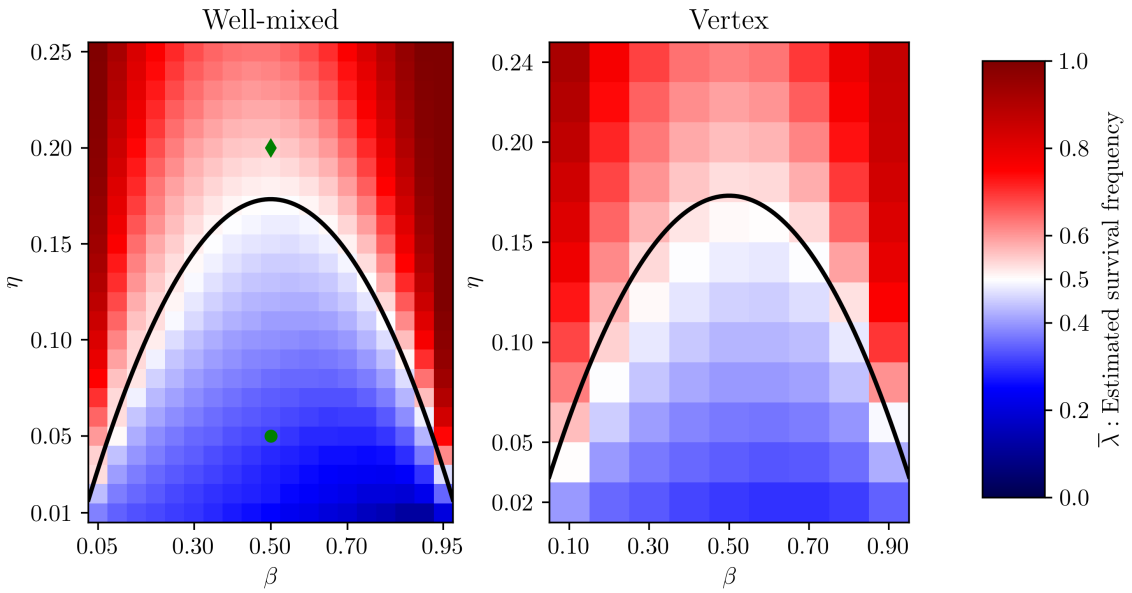

Figure 5.8: Estimated survival frequency, $\overline{\lambda}$, defined in Equation (5.18), for the well-mixed (left) and vertex (right) models with the exponential cell cycle model. All curves are the same as in Figure 5.7. The markers ♦ and ● in the left-hand plot correspond to the top and bottom simulations, respectively, in Figures 5.2 and 5.3.

in Figure 5.4 shows that this region corresponds to a deviation from the predicted homotypic survival probability more generally.

The reason for this discrepancy is that for very low η the rate of apoptosis becomes sufficiently high so that the limiting factor is no longer the survival probability, but rather the number of cells that are susceptible to apoptosis. For high β cell populations, there is a large fraction of cells in G1 phase at any one time, that are, in addition, vulnerable to apoptosis for a longer time. Contrast this with low β cell populations, where cells spend less time in G1 phase and thus have a higher chance of evading apoptosis, which results in a comparatively higher survival frequency.

Vertex model Even though the border is less finely resolved than in the well-mixed case, Figure 5.8 similarly shows a good agreement between theory and simulations. Moreover, we observe the same asymmetry for small η as we did in the well-mixed case. Compared to the well-mixed results, the observed survival frequencies appear less extreme, although this could also be because we sampled a smaller range of parameters for the vertex model.

5.5 Homotypic survival probability for uniform cell cycle model

In this section, we derive an expression for the homotypic survival probability with the uniform cell cycle model under the ergodic approximation, and validate the predictions with simulations. First, we rewrite the cumulative distribution function for the uniform distribution, Equation (3.13), using Equation (5.3) to give

$$\Psi(t) = \begin{cases} 0 & \text{for} \quad t < \beta t_{\mathrm{G}} - \frac{1}{2} r \\ \frac{t - \left(\beta t_{\mathrm{G}} - \frac{1}{2} r\right)}{r} & \text{for} \quad t \in \left[\beta t_{\mathrm{G}} - \frac{1}{2} r, \beta t_{\mathrm{G}} + \frac{1}{2} r\right] \\ 1 & \text{for} \quad t > \beta t_{\mathrm{G}} + \frac{1}{2} r \end{cases} . \tag{5.19}$$

We substitute the death time, defined in Equation (5.9), into Equation (5.19) to derive the homotypic survival probability

$$\lambda(\beta) = \begin{cases} 0 & \text{for} \quad \frac{T_\dagger}{c(1-\beta)} < \beta t_{\mathrm{G}} - \frac{1}{2} r \\ \frac{\frac{T_\dagger}{c(1-\beta)} - \left(\beta t_{\mathrm{G}} - \frac{1}{2} r\right)}{r} & \text{for} \quad \frac{T_\dagger}{c(1-\beta)} \in \left[\beta t_{\mathrm{G}} - \frac{1}{2} r, \beta t_{\mathrm{G}} + \frac{1}{2} r\right] \\ 1 & \text{for} \quad \frac{T_\dagger}{c(1-\beta)} > \beta t_{\mathrm{G}} + \frac{1}{2} r \end{cases} . \tag{5.20}$$

In contrast to the exponential cell cycle model, not all values of β in the interval $[0, 1]$ are valid. In particular, we need to satisfy the requirement that the lower bound of the uniform interval is non-negative. We substitute Equation (5.3) into Equation (3.12) to give

$$0 < r \leq 2\beta t_{\mathrm{G}} . \tag{5.21}$$

Rearranging Equation (5.21) yields

$$0 < \frac{1}{2} \frac{r}{t_{\mathrm{G}}} \leq \beta . \tag{5.22}$$

Equation (5.22) suggests that we can simplify notation by defining the non-negative, dimensionless parameter

$$\rho \equiv \frac{1}{2} \frac{r}{t_{\mathrm{G}}} , \tag{5.23}$$

which can be interpreted as a normalised range parameter. This enables us to rewrite Equation (5.22) as

$$0 < \rho \leq \beta. \tag{5.24}$$

In addition, we can also rewrite Equation (5.20) using the dimensionless parameter ρ. We show this by first dividing the lower bound condition of Equation (5.20) by t_G,

$$\frac{T_\dagger}{ct_G(1-\beta)} < \beta - \frac{1}{2}\frac{r}{t_G}, \tag{5.25}$$

and substituting Equations (5.13) and (5.23) to give

$$\frac{\eta}{1-\beta} < \beta - \rho. \tag{5.26}$$

Rearranging Equation (5.26), we have

$$\eta < (1-\beta)(\beta-\rho). \tag{5.27}$$

We repeat this process on the upper bound and rewrite Equation (5.20) as

$$\lambda(\beta,\eta,\rho) = \begin{cases} 0 & \text{for} \quad \eta < (1-\beta)(\beta-\rho) \\ \frac{1}{2} + \frac{1}{2\rho}\left(\frac{\eta}{1-\beta} - \beta\right) & \text{for} \quad \eta \in [(1-\beta)(\beta-\rho), (1-\beta)(\beta+\rho)] \\ 1 & \text{for} \quad \eta > (1-\beta)(\beta+\rho) \end{cases}. \tag{5.28}$$

We have thus derived an expression for the homotypic survival probability in terms of three dimensionless parameters, ρ, η, and β. We are particularly interested in the shape of the homotypic survival probability with respect to β when ρ and η are fixed, and will use the notation $\lambda(\beta)$ whenever we want to emphasise this. The function $\lambda(\beta)$ must span at least one and at most all of the three cases in Equation (5.28) (these are $\lambda = 0$, $0 < \lambda < 1$, and $\lambda = 1$), depending on the values of ρ and η. The qualitative shape of $\lambda(\beta)$ depends on which cases are traversed by varying β, hence we use these cases to classify distinct parameter regions[3].

5.5.1 Case $\{\lambda = 0\}$

First, we characterise the conditions for which the survival probability is zero. We define the **lower function** $l(\beta)$ as

$$l(\beta) = (1-\beta)(\beta-\rho), \tag{5.29}$$

[3] We use "parameter regions" instead of "parameter regimes" to avoid confusion with proliferation regimes.

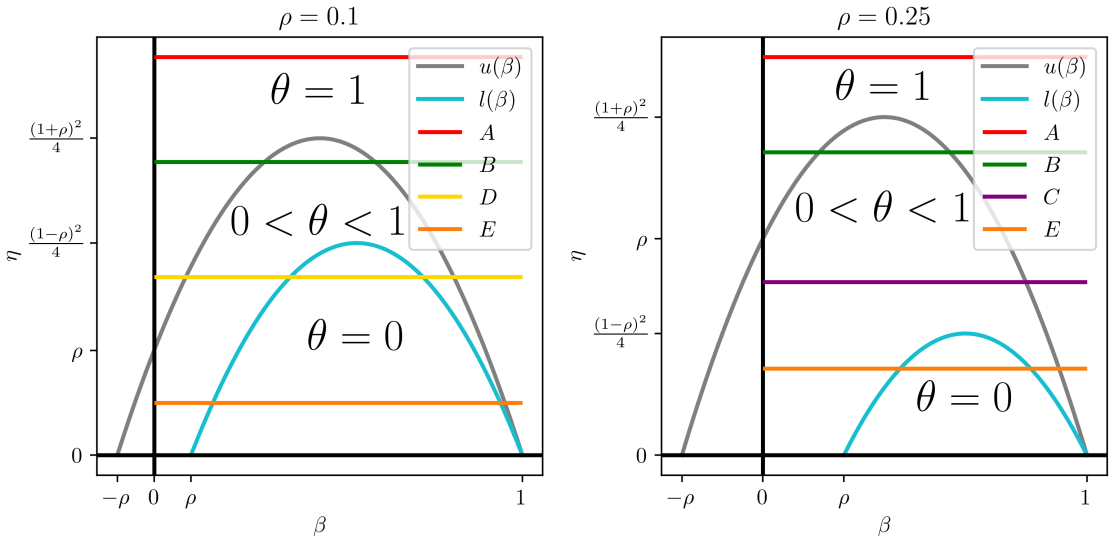

Figure 5.9: Plots of $l(\beta)$ and $u(\beta)$ defined by Equations (5.29) and (5.34), respectively. The horizontal lines correspond to particular values of η, which are classified in different regions depending on how they intersect $l(\beta)$ and $u(\beta)$. A: intersects neither $l(\beta)$ nor $u(\beta)$. B: intersects $u(\beta)$ twice and does not intersect $l(\beta)$. C: intersects $u(\beta)$ once and does not intersect $l(\beta)$. D: intersects $u(\beta)$ twice and intersects $l(\beta)$ twice. E: intersects $u(\beta)$ once and intersects $l(\beta)$ twice. **Left**: the peak of $l(\beta)$ is above the intersection of $u(\beta)$ with the y-axis, hence there is no horizontal line that can intersect $u(\beta)$ once without intersecting $l(\beta)$, i.e. parameter region C is excluded. **Right**: the peak of $l(\beta)$ is below the intersection of $u(\beta)$ with the y-axis, hence there is no horizontal line that can intersect $u(\beta)$ twice while also intersecting $l(\beta)$, i.e. parameter region D is excluded.

so that we can write the lower bound condition in Equation (5.28) as

$$\eta < l(\beta)\,. \tag{5.30}$$

The lower function is a concave quadratic function with roots at $\beta = \rho$ and $\beta = 1$, and has a maximum value of $(1-\rho)^2/4$. We can visually interpret Equation (5.30) by plotting $l(\beta)$ and drawing a horizontal line corresponding to the value of η, as demonstrated in Figure 5.9. Where the line is below the quadratic function, the condition in Equation (5.30) is satisfied and thus $\lambda = 0$.

The qualitative shape of $\lambda(\beta)$ depends on whether η intersects $l(\beta)$ or not. If η does not intersect $l(\beta)$, i.e. η is greater than the maximum value of $l(\beta)$, then $\lambda > 0$ for all β. However, if η intersects $l(\beta)$, i.e. η is less than or equal to the maximum value of $l(\beta)$, then $\lambda = 0$ for

some β. Formally,

$$\eta > \frac{(1-\rho)^2}{4} \Rightarrow \quad \lambda(\beta, \eta, \rho) > 0 \quad \text{for all} \quad \beta \ , \tag{5.31a}$$

$$\eta \leq \frac{(1-\rho)^2}{4} \Rightarrow \begin{cases} \lambda(\beta, \eta, \rho) = 0 & \text{for} \quad \beta \in [\beta^l_-, \beta^l_+] \\ \lambda(\beta, \eta, \rho) > 0 & \text{for} \quad \beta \notin [\beta^l_-, \beta^l_+] \end{cases} , \tag{5.31b}$$

where $\beta^l_\pm$ are the values of β where η intersects $l(\beta)$, i.e. the solutions to

$$\eta = l(\beta) \, , \tag{5.32}$$

which are

$$\beta^l_\pm = \frac{1+\rho}{2} \pm \sqrt{\frac{(1-\rho)^2}{4} - \eta} \, . \tag{5.33}$$

Since $l(\beta)$ is a quadratic function with roots at $\beta = \rho$ and $\beta = 1$, and ρ and η are both non-negative parameters, the solutions $\beta^l_\pm$ must lie in the interval $[\rho, 1]$.

5.5.2 Case $\{\lambda = 1\}$

Next, we characterise the conditions for which the survival probability is equal to one. We define the **upper function** $u(\beta)$ as

$$u(\beta) = (1-\beta)(\beta+\rho) \, , \tag{5.34}$$

so that we can write the upper bound condition of Equation (5.28) as

$$\eta > u(\beta) \, . \tag{5.35}$$

The upper function is, similar to $l(\beta)$, a concave quadratic function. However, its roots are at $\beta = -\rho$ and $\beta = 1$, and its maximum value is $(1+\rho)^2/4$. The visual interpretation of Equation (5.35) is similar as well; we plot $u(\beta)$ and a horizontal line representing η in Figure 5.9, and note that where the line is above the quadratic function, the condition in Equation (5.35) is satisfied and thus $\lambda = 1$.

The qualitative shape of $\lambda(\beta)$ depends on whether η intersects $u(\beta)$ or not. If η does not intersect $u(\beta)$, i.e. η is greater than the maximum value of $u(\beta)$, then $\lambda = 1$ for all β. However, if η intersects $u(\beta)$, i.e. η is less than or equal to the maximum value of $u(\beta)$, then

$\lambda < 1$ for some β. Formally,

$$\eta \geq \frac{(1+\rho)^2}{4} \Rightarrow \quad \lambda(\beta, \eta, \rho) = 1 \quad \text{for all} \quad \beta \ , \tag{5.36a}$$

$$\eta < \frac{(1+\rho)^2}{4} \Rightarrow \begin{cases} \lambda(\beta, \eta, \rho) < 1 & \text{for} \quad \beta \in [\max(\beta_-^u, 0), \beta_+^u] \\ \lambda(\beta, \eta, \rho) = 1 & \text{for} \quad \beta \notin [\max(\beta_-^u, 0), \beta_+^u] \end{cases} , \tag{5.36b}$$

where $\beta_\pm^u$ are the values of β where η intersects $u(\beta)$, i.e. the solutions to

$$\eta = u(\beta) \, , \tag{5.37}$$

which are

$$\beta_\pm^u = \frac{1-\rho}{2} \pm \sqrt{\frac{(1+\rho)^2}{4} - \eta} \, . \tag{5.38}$$

Since $u(\beta)$ is a quadratic function with roots at $\beta = -\rho$ and $\beta = 1$, and ρ and η are both non-negative parameters, the solutions $\beta_\pm^u$ must lie in the interval $[-\rho, 1]$.

The reason for using $\max(\beta_-^u, 0)$ in Equation (5.36b) is that the solution β_-^u can be negative. When this is the case, the condition in Equation (5.35) is not satisfied for $\beta = 0$. The condition that β_-^u is negative, i.e.

$$\beta_-^u = \frac{1-\rho}{2} - \sqrt{\frac{(1+\rho)^2}{4} - \eta} < 0 \, , \tag{5.39}$$

can be rewritten by rearranging and raising to the power of two to give

$$\frac{(1-\rho)^2}{4} < \frac{(1+\rho)^2}{4} - \eta \, , \tag{5.40}$$

which amounts to

$$\eta < \rho \, . \tag{5.41}$$

Visually, this translates to the horizontal line corresponding to η being below the intersection of $u(\beta)$ with the y-axis, since $u(0) = \rho$, as seen in Figure 5.9.

Thus, in addition to the relative position of η with respect to the peak of $u(\beta)$, the shape of $\lambda(\beta)$ also depends on whether η is above the intersection of $u(\beta)$ with the y-axis or not. If η is below the intersection of $u(\beta)$ with the y-axis, i.e. η is less than ρ, then $\lambda < 1$ for $\beta = 0$. However, if η is above the intersection of $u(\beta)$ with the y-axis, i.e. η is greater than or equal

to ρ, then $\lambda = 1$ for $\beta = 0$. Formally,

$$\eta \geq \rho \Rightarrow \lambda(0, \eta, \rho) = 1\,, \tag{5.42a}$$

$$\eta < \rho \Rightarrow \lambda(0, \eta, \rho) < 1\,. \tag{5.42b}$$

5.5.3 Parameter regions

Taken together, Equations (5.31), (5.36), and (5.42) form three conditions on the values of η and ρ that determine the shape of $\lambda(\beta)$. In total, we can distinguish five different parameter regions. We classify them as A through E, roughly starting from the highest η value to the lowest value. For a given ρ value, the parameter regions C and D are mutually exclusive. This is because the peak of $l(\beta)$ can either be below or above the intersection of $u(\beta)$ with the y-axis, depending on the value of ρ. The left-hand plot in Figure 5.9 shows the case with the peak above the intersection, while the right-hand plot shows the opposite case.

Region A This region is defined through the constraint

$$\eta \geq \frac{(1+\rho)^2}{4}\,, \tag{5.43}$$

is satisfied, such that Equation (5.36a) is true. Equation (5.43) also implies that the conditions in Equations (5.31a) and (5.42a) are satisfied. In this parameter region, the value of η is sufficiently large that the survival probability is equal to one for all β. Hence, when plotted, $\lambda(\beta)$ is a straight line.

Region B This region is defined through the constraint

$$\frac{(1-\rho)^2}{4} < \eta < \frac{(1+\rho)^2}{4} \quad \text{and} \quad \eta \geq \rho\,, \tag{5.44}$$

are satisfied, such that the conditions in Equations (5.36b), (5.31a), and (5.42a) apply. In this region, η intersects $u(\beta)$, but not $l(\beta)$, and is greater than or equal to ρ. Therefore, λ is less than one for $\beta \in [\beta^u_-, \beta^u_+]$, equal to one outside of this interval, and never reaches zero.

Region C This region is defined through the constraint

$$\frac{(1-\rho)^2}{4} < \eta < \frac{(1+\rho)^2}{4} \quad \text{and} \quad \eta < \rho\,, \tag{5.45}$$

are satisfied, such that the conditions in Equations (5.36b), (5.31a), and (5.42b) apply. In this region, η intersects $u(\beta)$, but not $l(\beta)$, and is less than ρ. Therefore, λ is less than one for $\beta \in [0, \beta_+^u]$, equal to one outside of this interval, and never reaches zero.

Region D This region is defined through the constraint

$$\eta \leq \frac{(1-\rho)^2}{4} \quad \text{and} \quad \eta \geq \rho\,, \tag{5.46}$$

are satisfied, such that the conditions in Equations (5.31b), and (5.42a) apply. Equation (5.46) also implies that the condition in Equation (5.36b) applies. In this region, η intersects both $u(\beta)$ and $u(\beta)$, and is greater than ρ. Therefore, λ is less than one for $\beta \in [\beta_-^u, \beta_+^u]$, equal to one outside of this interval, and equal to zero for $\beta \in [\beta_-^l, \beta_+^l]$.

Region E This region is defined through the constraint

$$\eta \leq \frac{(1-\rho)^2}{4} \quad \text{and} \quad \eta < \rho\,, \tag{5.47}$$

are satisfied, such that the conditions in Equations (5.36b), and (5.42b) apply. Equation (5.46) also implies that the condition in (5.31b) applies. In this region, η intersects both $u(\beta)$ and $l(\beta)$, and is less than ρ. Therefore, λ is less than one for $\beta \in [0, \beta_+^u]$, equal to one outside of this interval, and equal to zero for $\beta \in [\beta_-^l, \beta_+^l]$.

The parameter regions in (ρ, η)-space are visualised in Figure 5.10. In addition, a representative plot of $\lambda(\beta)$ for each parameter region is plotted in Figure 5.11.

5.5.4 Minima

In addition to the shape of $\lambda(\beta)$, we are also interested in the minimum homotypic survival probability. As we have established, in parameter regions D and E, the survival probability is equal to zero in the interval $[\beta_-^l, \beta_+^l]$, hence the survival probability is minimised in this interval. On the other hand, in parameter region A the survival probability is always equal to one, so the survival probability is "minimised" in the interval $[0, 1]$.

In the remaining parameter regions, B and C, the minimum survival probability is between zero and one. We find the minimiser, denoted by β^*, by first taking the derivative of the survival probability in Equation (5.28) with respect to β:

$$\frac{\partial \lambda}{\partial \beta} = \frac{1}{2\rho}\left(\frac{\eta}{(1-\beta)^2} - 1\right). \tag{5.48}$$

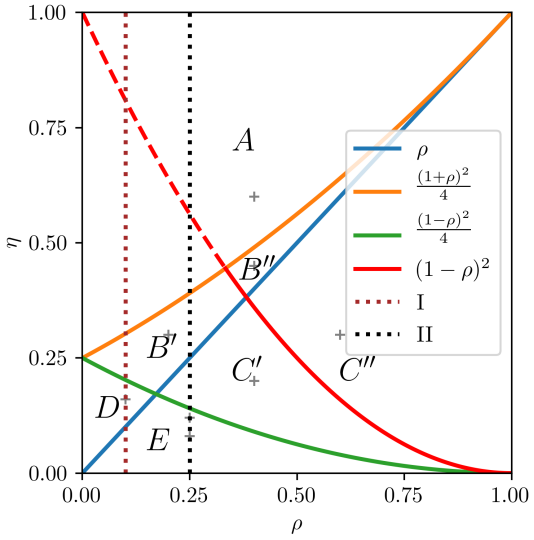

Figure 5.10: Diagram of the parameter regions that determine the shape of $\lambda(\beta)$ in (ρ, η)-space. The curve $(1-\rho)^2$ is only relevant to parameter regions B and C (see Section 5.5.4), hence it is dashed within parameter region A. In each parameter region, representative parameter sets are marked with crosses. These crosses correspond to the values in Figure 5.11 and Table 5.4. The lines $\rho = 0.1$ and $\rho = 0.25$ are plotted, corresponding to the values in Figure 5.9 and to Cross sections I and II, as defined in Section 5.6.

We then find the values of β such that Equation (5.48) is zero, which yields two solutions

$$\beta^*_{\pm} = 1 \pm \sqrt{\eta}\,. \tag{5.49}$$

We can eliminate $1 + \sqrt{\eta}$ because β is a proportion and cannot be larger than one. The unique minimiser is therefore

$$\beta^* = 1 - \sqrt{\eta}\,. \tag{5.50}$$

However, it is not guaranteed that β^* satisfies the condition in Equation (5.24). If we impose this condition, i.e.

$$\rho \leq \beta^* = 1 - \sqrt{\eta}\,, \tag{5.51}$$

we have the following condition on η:

$$\eta \leq (1-\rho)^2\,. \tag{5.52}$$

If this condition is satisfied, then the minimiser is β^*. If it is not satisfied, however, then β^* is not a valid value for β. In that case, the survival probability is instead minimised by the smallest valid value for β, which is ρ. Based on this distinction, we subdivide parameter

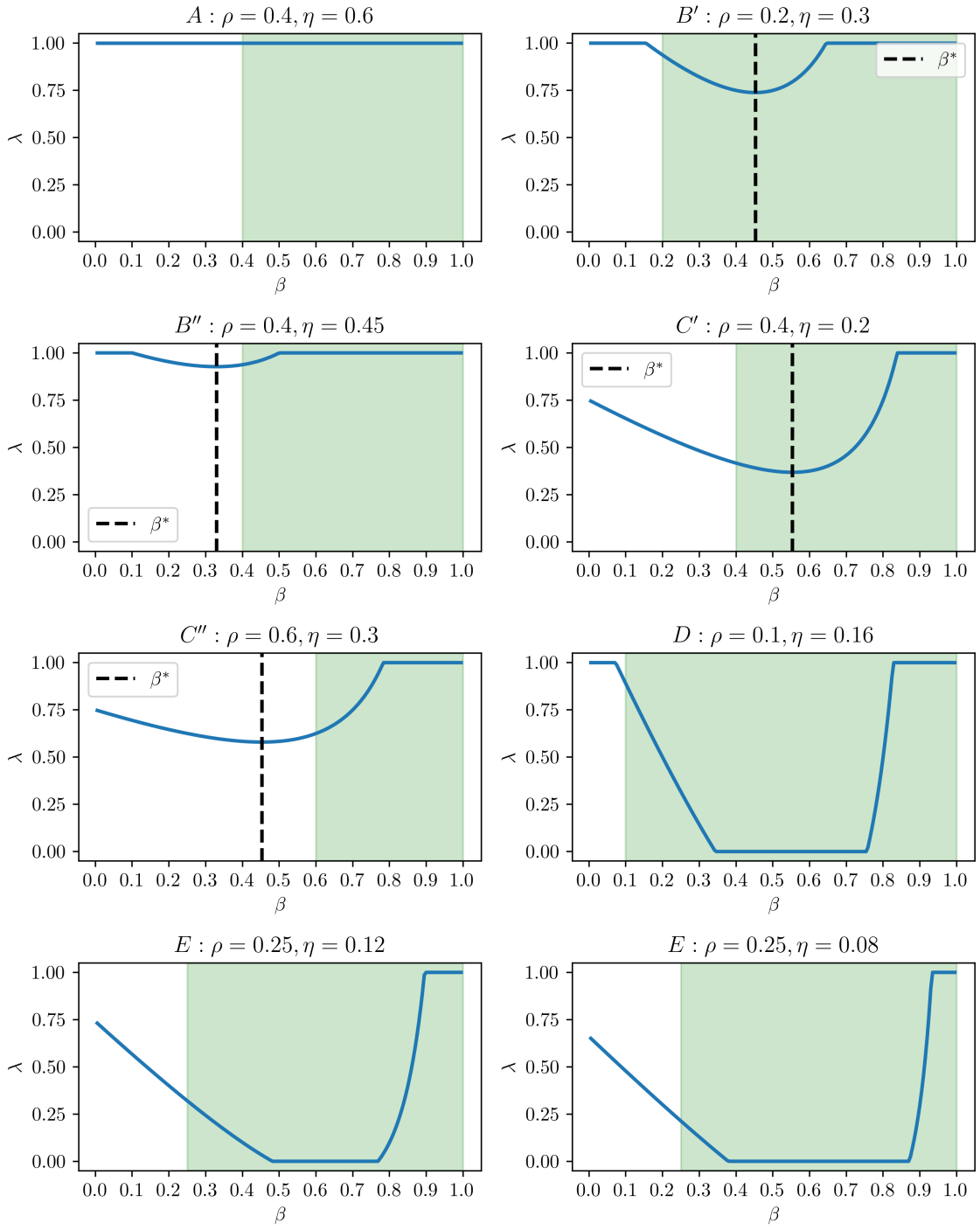

Figure 5.11: Homotypic survival probability with the uniform cell cycle model under the ergodic approximation as predicted by Equation (5.28). The black dashed lines indicate the minimum survival probability for Regions B and C. The green shaded area shows the range of valid β values.

region B into B' and B'', corresponding to β^* being a valid and invalid parameter, respectively. We split parameter region C into C' and C'' using the same criteria.

5.5.5 Computational validation

We check whether the homotypic survival probability predicted in Equation (5.28) is supported by simulations of the G2 death signal model. Concretely, we used a Monte Carlo method where we ran repeated simulations for various values of ρ, η, and β, and sampled the resulting survival frequencies. The survival frequency was computed using Equation (2.12). We supplied a different random seed to each simulation, and set the minimum cell count to 10.

Similarly to the exponential cell cycle model, we expect that the ergodic assumption is violated at low survival probabilities and low cell counts. A major difference, however, from the exponential cell cycle model, is that we can independently control the variance of the uniform cell cycle model. For small ρ, which can be seen as a normalised range parameter, G1 durations are sampled from a narrow distribution, potentially resulting in synchronised cell cycles and low ergodicity. In the limit of $\rho = 0$, the cell cycle is entirely deterministic, in which case we do not expect ergodicity.

Parameter choice We used the well-mixed model with the parameters listed in Tables 5.3 and 5.4 to validate the predicted homotypic survival probabilities in Equation (5.28). We chose eight pairs of values for ρ and η, as listed in Table 5.4, each corresponding to a different parameter region. In each parameter region, we sampled ten values for β ranging from ρ to $1 - (1 - \rho)/10$ in increments of $(1 - \rho)/10$. We also added the minimiser β^* to the parameter values in parameter regions C' and B'. The death threshold was determined as a function of the other parameters according to Equation (5.15).

For the vertex model, we used the parameters listed in Tables 5.3 and 5.4. Because the vertex model uses significantly more computational resources to run, we only ran simulations for selected β values, as listed in Table 5.4. The mechanical parameters of the vertex model are set to their default values as defined in Chapter 2.

For the well-mixed model, we ran 100 simulations for every unique parameter set, so the total number of simulations is $(8 \times 10 + 2) \times 100 = 8\,200$. For the vertex model, we ran 20 simulations for every unique parameter set. Hence, the total number of simulations is $20 \times 20 = 400$.

Initial conditions See Initial conditions in Section 5.3.1 (page 95).

Table 5.3: Model and simulation parameter values used to estimate the homotypic survival probability for the uniform cell cycle model. The values of ρ and η for the well-mixed and vertex models, and of β for the vertex model are listed separately in Table 5.4.

Parameter	Well-mixed	Vertex
t_G		100
c		1
Initial cell count		100
Simulation end time	∞	100 000
Minimum cell count		10
Maximum cell count		1 000
N_{sim}	100	20

Table 5.4: Parameter values of ρ and η used to estimate the homotypic survival probability for the uniform cell cycle model. In addition, the last column contains the values of β used for the vertex model.

Region	ρ	η	Vertex β-values
A	0.4	0.6	0.5
B'	0.2	0.3	0.2, $1-\sqrt{0.3}$, 0.8
B''	0.4	0.45	0.4, 0.8
C'	0.4	0.2	0.4, $1-\sqrt{0.2}$, 0.9
C''	0.6	0.3	0.6, 0.9
D	0.1	0.16	0.1, 0.5, 0.9
E	0.25	0.12, 0.08	0.25, 0.5, 0.95

Termination conditions See Termination conditions in Section 5.3.1 (page 96).

Data processing See Data processing in Section 5.3.1 (page 98)

Data visualisation The distribution of observed survival frequencies is plotted against theoretical predictions in Figure 5.12 for the well-mixed model and in Figure 5.14 for the vertex model.

Computational materials and methods See Computational materials and methods in Section 2.3.1 (page 36) for the vertex model and in Section 4.4 (page 77) for the well-mixed model.

5.5.6 Results

Well-mixed model Figure 5.12 shows a good qualitative agreement between theory and experiment for the parameter regions A, B, and C. However, in parameter regions D and E the survival frequency is significantly higher than expected. For parameter region E, the observed survival frequency hovers at around $1/2$, in contrast to the predicted survival probability of zero. Moreover, in parameter region D the survival frequency is significantly higher than $1/2$. Particularly surprising is the fact that all of the simulations in parameter region D hit the maximum cell count, contrary to what we would expect from the predicted survival probability.

In order to investigate this further, we plotted the cell counts of selected simulations in the D and E parameter regions in Figure 5.13. We can see that both simulations settle into an oscillatory pattern where the population alternates between G1 and G2 phases, thus violating the ergodic assumption. In the E simulation, the number of apoptosis and division events are roughly in balance, allowing the oscillation to persist over many cycles. The end time for the simulation was 7 000, meaning that the population went through approximately 70 cell cycles before finally going extinct. On the other hand, in the D simulation there are proportionally far fewer deaths occurring per cycle. As a result, the population grows significantly at every cycle and ultimately reaches the maximum cell count.

Vertex model Similar to the well-mixed model, Figure 5.14 shows qualitative agreement between the observations and predictions for parameter regions A, B, and C, but large discrepancies for parameter regions D and E. Snapshots from the vertex simulations in Figure 5.15 demonstrate that the cell population also settles into oscillatory patterns for

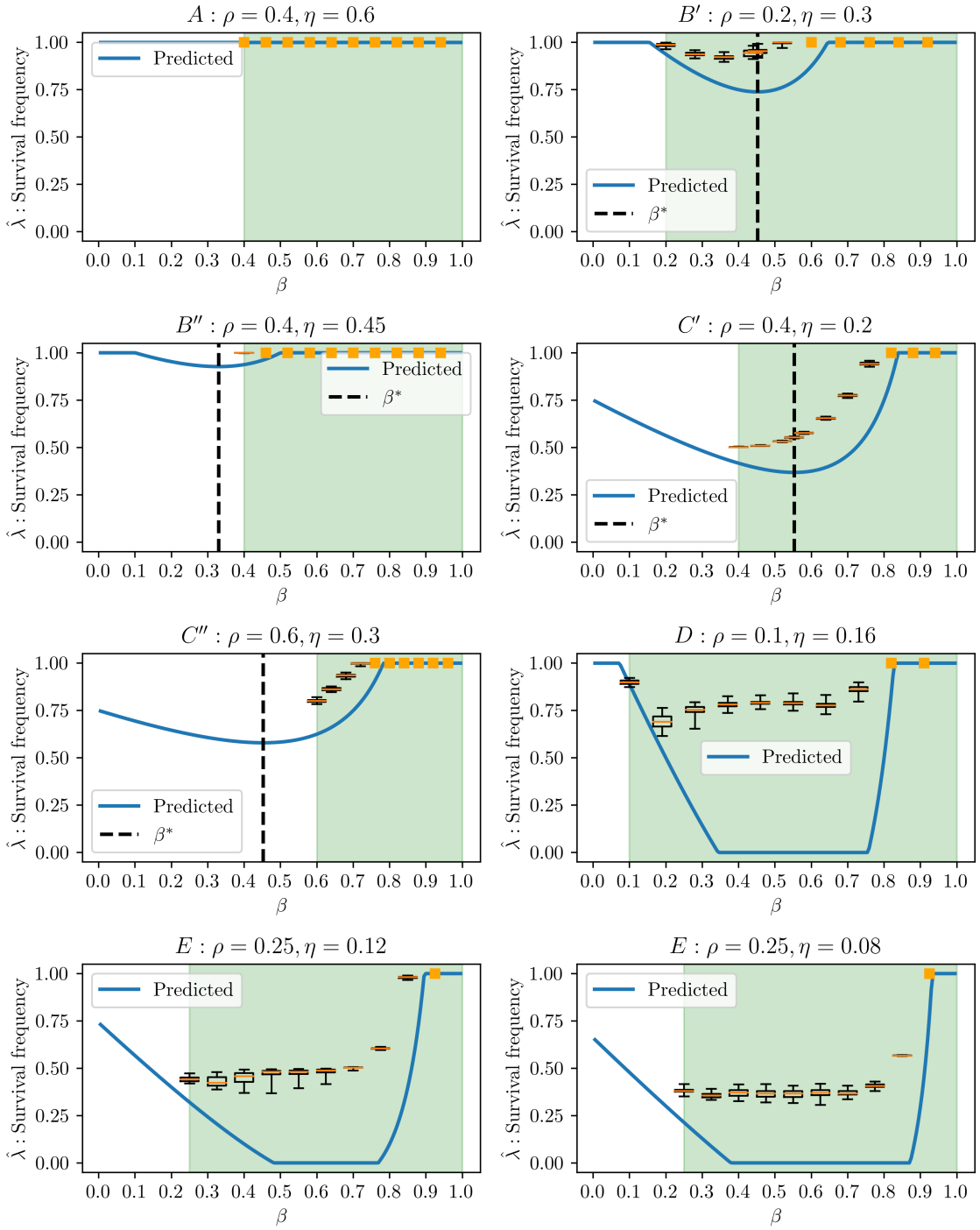

Figure 5.12: Estimated homotypic survival frequency, $\hat{\lambda}$, defined in Equation (2.12), for well-mixed simulations with the uniform cell cycle model. The distribution of the survival frequency is plotted using box plots with whiskers extending to the minimum and maximum values. An orange square indicates that all observed survival frequencies are the same. The green shaded area shows the range of valid β values. The predicted homotypic survival probability, defined in Equation (5.28), is plotted for comparison.

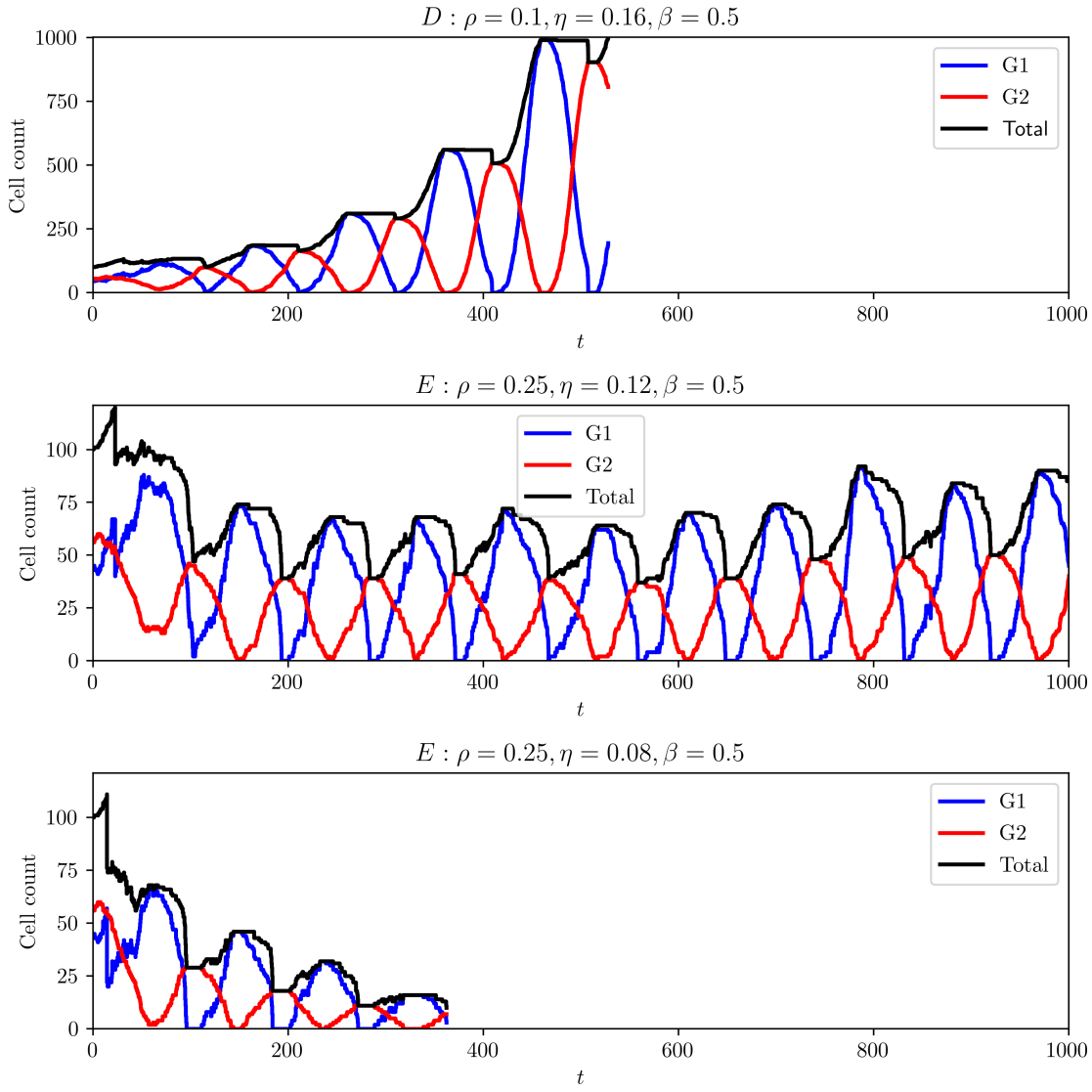

Figure 5.13: Cell counts for simulations of the well-mixed model with the uniform cell cycle model in parameter regions D and E. The top figure shows a simulation in parameter region D that terminated by hitting the maximum cell count. The middle and bottom figures show simulations in parameter region E that terminated by hitting the minimum cell count. The top, middle, and bottom simulations terminated at $t = 527.99$, $t = 11\,502.72$, and $t = 362.16$, respectively. In order to compare the cell counts across simulations, the range from $t = 0$ to $t = 1000$ is plotted.

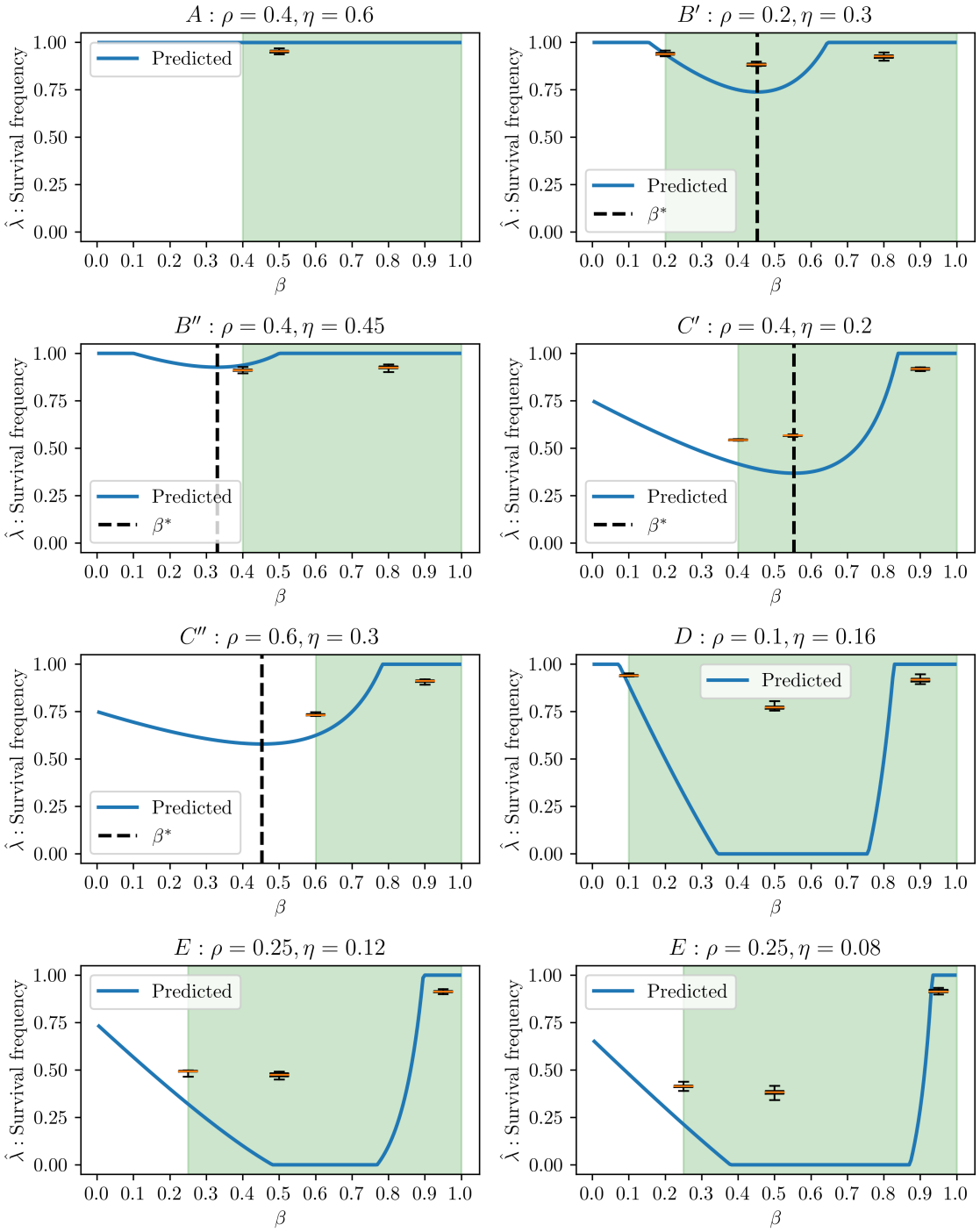

Figure 5.14: Estimated homotypic survival frequency, $\hat{\lambda}$, defined in Equation (2.12), for vertex simulations with the uniform cell cycle model. The distribution of the survival frequency is plotted using box plots with whiskers extending to the minimum and maximum values. An orange square indicates that all observed survival frequencies are the same. The green shaded area shows the range of valid β values. The predicted homotypic survival probability, defined in Equation (5.28), is plotted for comparison.

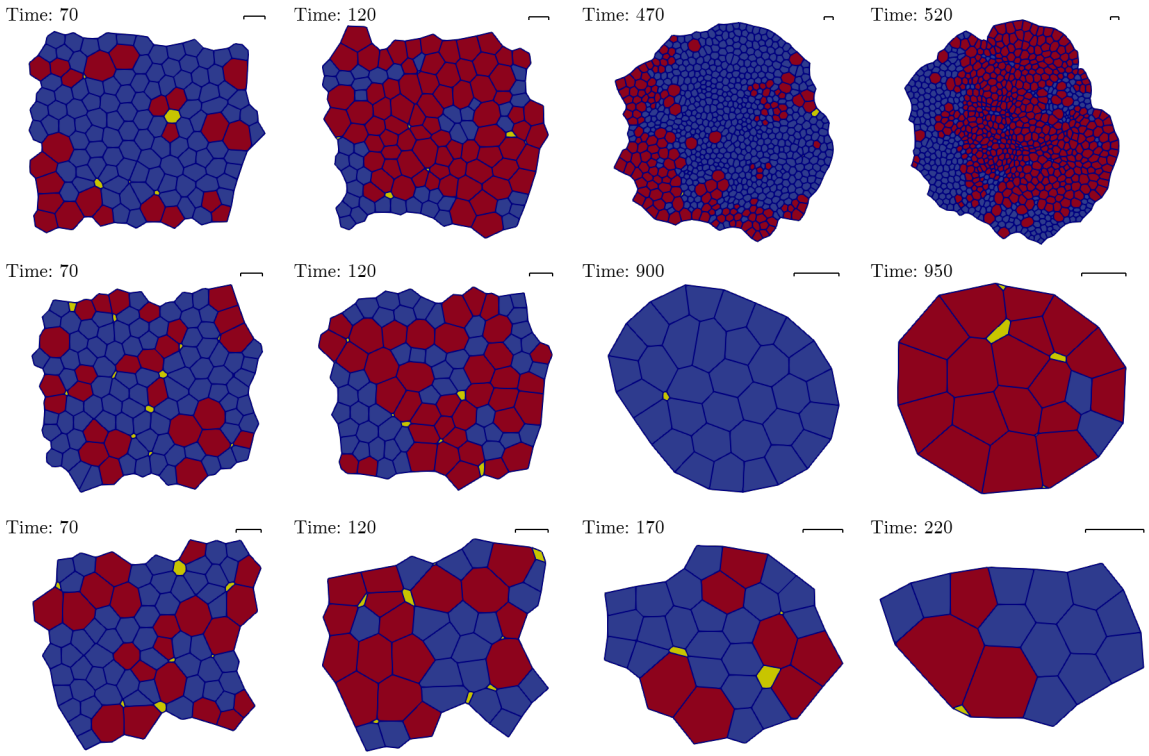

Figure 5.15: Snapshots from simulations of the vertex model with the uniform cell cycle model. Each row is one simulation and cells are coloured according to their state. ■: G1 phase. ■: G2 phase. ■: apoptotic. The sets of parameters ρ, η, and β are the same as in Figure 5.13 for the respective rows. **Top** (region $D : \rho = 0.1, \eta = 0.16, \beta = 0.5$): the cell population grew with few apoptotic cells until the maximum cell count was reached at $t = 540$. Snapshots taken at 50 time units apart (half the average cell cycle length) show local synchronisation of cell cycle phases. **Middle** (region $E : \rho = 0.25, \eta = 0.12, \beta = 0.5$): the cell population slowly declined until the minimum cell count was reached at $t = 1958$. Snapshots at $t = 900$ and $t = 950$ show strong synchronisation across the whole population. **Bottom** (region $E : \rho = 0.25, \eta = 0.08, \beta = 0.5$): the cell population rapidly declined until the minimum cell count was reached at $t = 251$.

this model. In contrast to the well-mixed model, however, the synchronisation of cell cycle phases occurs in local regions instead of across the whole population.

5.6 Proliferation regimes for uniform cell cycle model

For the uniform cell cycle model, the ergodic approximation is violated in some parameter regions where the population goes through oscillations. There are also oscillations for the exponential cell cycle model, but those occur because the population becomes too small for the ergodic approximation to hold. Hence, we classified it as "nonviable"[4]. The oscillations in the uniform cell cycle model, on the other hand, can go through cyclic growth (as evidenced by Figures 5.13 and 5.15), which makes them qualitatively different. However, so far there is no indication in the literature that oscillations and cell cycle synchrony are biologically relevant to cell competition, so we do not attempt to classify different types of oscillations, and instead class all oscillations as nonviable.

To get the homotypic viability curve for the uniform cell cycle model, we substitute Equation (5.28) into Equation (5.16) to give

$$\eta = \beta(1 - \beta)\,. \tag{5.53}$$

This expression is similar to Equation (5.17), except for a coefficient of ln(2). We note, however, that the homotypic survival probability is a function of three parameters: ρ, η, and β. Therefore, the homotypic viability curve is more precisely a surface in this case. We only consider cross sections of this parameter space for fixed ρ values, so we will continue to use the word "curve".

The fact that ρ does not appear in Equation (5.53) may give the impression that ρ does not affect cell proliferation. However, we know from the previous section that it strongly modulates the survival probability on either side of the homotypic viability curve, and also the proliferative behaviour observed. Hence, to examine the effect of varying ρ, we plot two representative cross sections at $\rho = 0.1$ and $\rho = 0.25$. The reason for choosing these values is that they intersect different parameter regions: as Figure 5.10 shows, $\rho = 0.1$ intersects parameter region D but not C, and vice versa for $\rho = 0.25$. These ρ values also correspond to the values used to plot the lower and upper functions in Figure 5.9.

In addition to viability, we also distinguish the cases where the ergodic approximation predicts $\lambda = 0$ and $\lambda = 1$. We know from Section 5.5 that these regimes are delineated by the lower and upper functions, respectively. Even though the homotypic viability curve does not

[4] See Section 5.4 for classification.

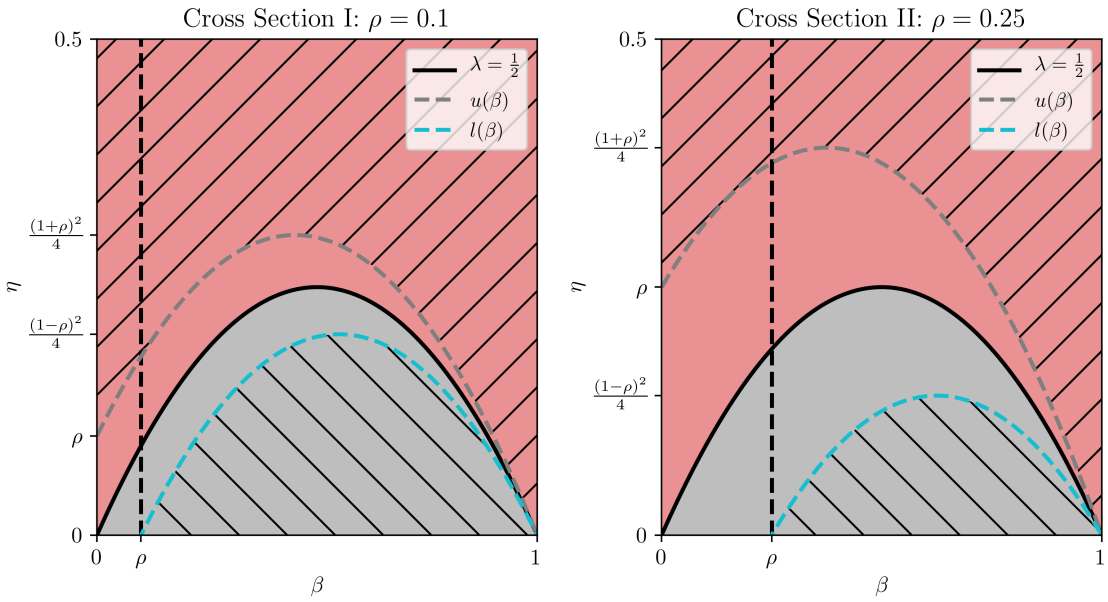

Figure 5.16: Diagrams of proliferation regimes for uniform cell cycle model for Cross Sections I and II. : Viable Regime and $\lambda < 1$. : Viable Regime and $\lambda = 1$. : Nonviable Regime and $\lambda > 0$. : Nonviable Regime and $\lambda = 0$.

depend on ρ, the lower and upper functions are dependent on ρ. As Figure 5.16 shows, the gap between the lower and upper functions increases with ρ, while the homotypic viability curve is stationary in the widening gap.

5.6.1 Computational validation

We validated the predicted proliferation regimes for the uniform cell cycle model by running a parameter sweep over a regular grid in (β, η) for different values of ρ and estimating the average survival frequency. For every unique parameter set we ran repeated Monte Carlo simulations using a unique seed in each simulation.

Parameter choice We sampled parameters from the same cross sections displayed in Figure 5.16, i.e. $\eta \in [0, 0.5]$ and $\beta \in [0, 1]$ for $\rho = 0.1$ and $\rho = 0.25$. We discretised the parameter space into a regular grid of η- and β-values. Concrete parameter values for the well-mixed and vertex models are listed in Table 5.5. Because β must be greater than or equal to ρ, we omit the β values in Cross Section II for which this is not the case. We used a coarser grid for the vertex model than the well-mixed model because of the difference in computational cost. In addition, we fixed $t_G = 100$ and $c = 1$, and computed the death threshold using Equation (5.15).

Table 5.5: Model and simulation parameter values used to estimate the proliferation regimes for the uniform cell cycle model.

Parameter	Well-mixed	Vertex
t_G		100
c		1
ρ		$0.1, 0.25$
η	$0.02, 0.04, \ldots, 0.50$	$0.04, 0.08, \ldots, 0.48$
β	$0.05, 0.10, \ldots, 0.95$	$0.1, 0.2, \ldots, 0.9$
Initial cell count		100
Simulation end time		10 000
Minimum cell count		10
Maximum cell count		1 000
N_{sim}	50	20

For the well-mixed model, we ran 50 simulations for every unique parameter set, so the total number of simulations is $(25 \times 18 + 25 \times 15) \times 50 = 41\,250$. For the vertex model, we ran 20 simulations for every unique parameter set. Hence, the total number of simulations is $(12 \times 9 + 12 \times 7) \times 20 = 3\,840$.

Initial conditions See Initial conditions in Section 5.3.1 (page 95).

Termination conditions We set minimum and maximum cell counts of 10 and 1 000, respectively. In addition, because the parameter sweep is relatively large, we set a simulation end time of 10 000 to limit the computational time that each simulation can take.

Data processing See Data processing in Section 5.4.1 (page 105).

Data visualisation We used heat maps with the *seismic* colour map, range $[0, 1]$, and central value $1/2$, similarly to Section 5.4.1 (Data visualisation, page 105), to plot the estimated survival frequency for the well-mixed and vertex models in both cross sections. In addition, we drew the homotypic viability curve, as well as the upper and lower functions, on the heat map. The results are given in Figure 5.17.

Computational materials and methods See Computational materials and methods in Section 2.3.1 (page 36) for the vertex model and in Section 4.4 (page 77) for the well-mixed

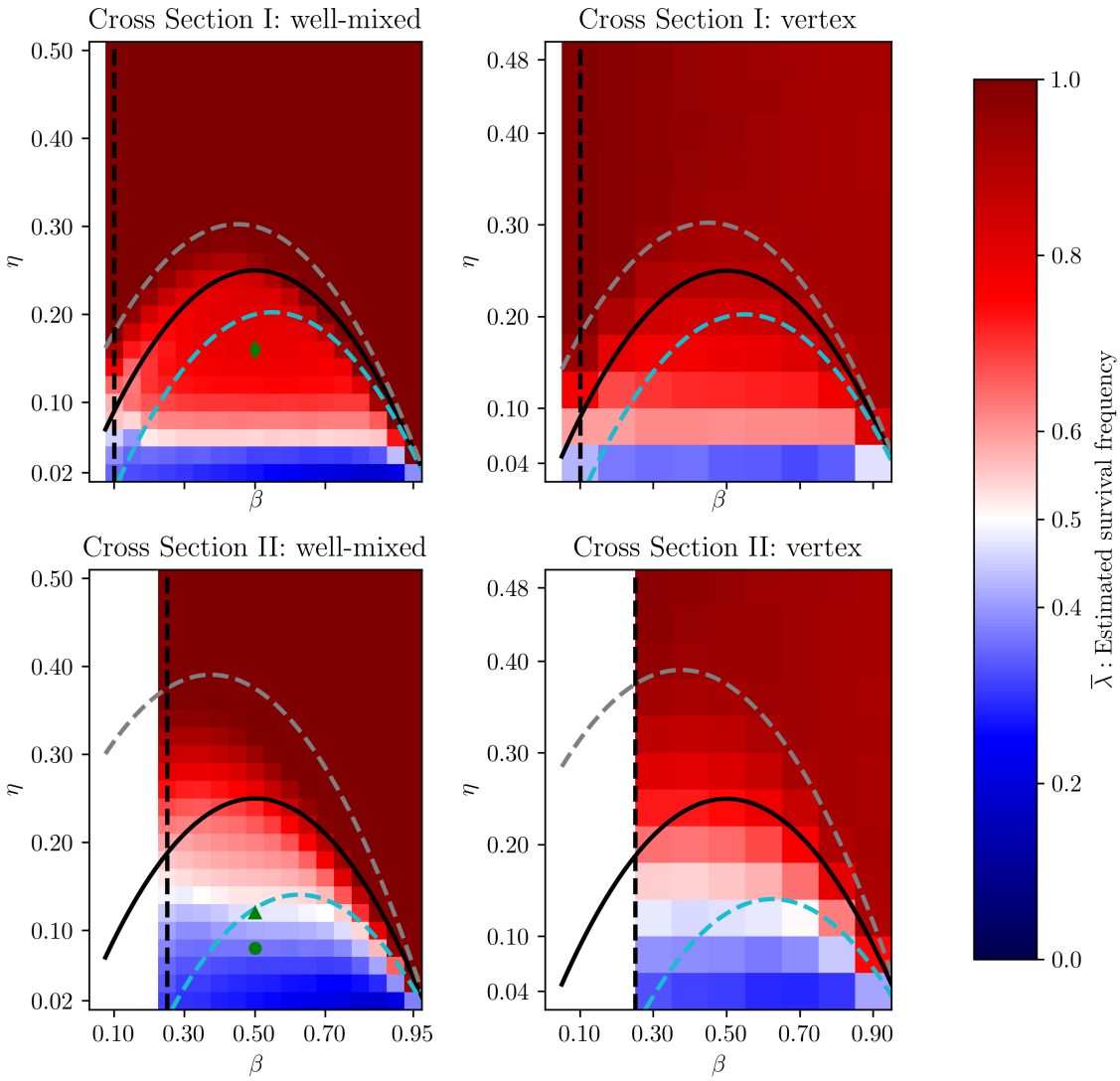

Figure 5.17: Estimated survival frequency, $\overline{\lambda}$, defined in Equation (5.18), for the well-mixed (left column) and vertex (right column) models with the uniform cell cycle model. All curves are the same as in Figure 5.16. The markers ♦, ▲, and ● in the left-hand plots correspond to the top, middle, and bottom simulations, respectively, in Figures 5.13 and 5.15.

model.

5.6.2 Results

Well-mixed model Figure 5.17 shows large discrepancies with our predictions. In particular, the red region, indicating population growth, penetrates into the predicted Nonviable Regime. This effect is more pronounced for $\rho = 0.1$ than $\rho = 0.25$. Even worse, the red region extends into the region where the ergodic approximation predicts $\lambda = 0$. In fact, the estimated survival frequency is larger than zero across all parameter sets. In contrast, the region where λ is predicted to be equal to one does produce survival frequencies close to one.

As previously observed in Section 5.5, we find large discrepancies between simulations and predictions because the uniform cell cycle model results in non-ergodicity. To illustrate this point, we have marked the parameters corresponding to the simulations in Figures 5.13 and 5.15 on the heat map. We see that oscillations and temporal segregation of cell cycles can "rescue" a cell type from extinction by minimising overlap between death signal-producing cells and death signal-sensitive cells. Even in the case of oscillations, however, increasingly lower η values will result in extinction, as evidenced by the simulations in the bottom rows of Figures 5.13 and 5.15. This explains why we can still observe a blue region for very small η in Figure 5.5.

Vertex model Figure 5.17 shows that the vertex model produces similar results as the well-mixed model. Therefore, even though the ergodic approximation does not hold for the uniform cell cycle model, it appears that local interactions can still result in cell cycle synchronisation and oscillations. One difference with the well-mixed results is that the survival frequency for high η is slightly lower than one. When we exclude extrusion events from the calculation of survival frequency (results not shown for brevity), this difference disappears.

5.7 Discussion

In this chapter, we defined the G2 death signal model and studied its properties in homotypic populations. The G2 death signal is a short-range intercellular interaction where cells in G2 phase emit death signals to their direct neighbours. In order to investigate the viability of homotypic populations, we built a theoretical framework based on the notion of **ergodicity**. This is the system-level property that the spatial distribution of cell cycle phases reflects

the temporal distribution of cell cycle phases. Using this approximation, we derived the homotypic survival probability and charted proliferation regimes in parameter space for the exponential and uniform cell cycle models. We validated the predictions using simulations with great success for the former and mixed results for the latter.

Survival is governed by two key parameters We found that the homotypic survival probability, λ, is determined by two key dimensionless parameters: β and η. The former is the proportion of the cell cycle spent in the G1 phase on average, i.e. $t_{\mathrm{G1}}/t_{\mathrm{G}}$, and the latter is the death threshold $T_\dagger$ divided by the death signal constant c and the total cell cycle time t_{G}. In addition, the homotypic survival probability for the uniform cell cycle model is also influenced by ρ, which is half the uniform interval width divided by t_{G}.

The parameter η controls overall sensitivity to death signals. Increasing η results in a higher survival probability for any value of β. Therefore, cell types with sufficiently high η can escape death signal-induced apoptosis. This is also illustrated by the fact that as we "turn off" the death signal by taking $c \to 0$, we have $\eta \to \infty$. Biologically, this limit corresponds to inhibiting apoptosis, which can be achieved experimentally in *Drosophila* [205]. Indeed, it has been demonstrated experimentally that inhibiting apoptosis prevents cell competition [42, 53].

On the other hand, we found that the homotypic survival probability is U-shaped with respect to β. This represents the balance of opposing forces inherent in the G2 death signal model: cells are vulnerable to death signals in G1 phase, but impervious to apoptosis in G2 phase, and cells only emit death signals in G2 phase, but are silent in G1 phase. Thus, increasing β increases the cell's sensitivity to death signals on the one hand, and lowers the cell's emission of death signals on the other hand. A low β value corresponds to a short window of vulnerability with a high death signal, and vice versa for a high β value. Extreme values of β allow for an escape from the effects of the G2 death signal, and intermediate values produce the highest rates of cell death.

Cell cycle model determines ergodicity We validated the predictions made using the ergodic approximation with well-mixed and vertex simulations for the exponential and uniform cell cycle models. For the exponential cell cycle model, the computational results were in excellent agreement with the theoretical results for both the homotypic survival probability and the proliferation regimes. Hence, we conclude that the G2 death signal model is ergodic when using the exponential cell cycle model for both the well-mixed and vertex models.

For the uniform cell cycle model, we observed large differences between the observed

and predicted survival frequencies. For low η, the ergodic approximation predicted a zero survival probability. However, in practice we observed a range of survival frequencies. For some parameter sets the survival frequency went above $1/2$, indicating a growing population. Similarly, when estimating proliferation regimes, we observed many parameter sets resulting in growing populations, despite being below the homotypic viability curve.

This indicates that the ergodic assumption is not satisfied for the uniform cell cycle model. Indeed, visual inspection of specific simulations revealed that cells settle into an oscillatory pattern, alternating between G1 and G2 phases. It appears that, instead of causing extinction, the low predicted survival probability results in the elimination of cells that are unsynchronised, leading to the temporal segregation of cell cycle phases at the population level. This breaks the ergodic assumption and lets synchronised cells escape death signal-induced apoptosis. The reason for this is presumably because the uniform cell cycle model samples G1 durations from a narrower range than the exponential cell cycle model, and is therefore more likely to produce cell cycle synchronisation.

We expected to see some discrepancies because populations cannot go extinct in the G2 death signal model. A survival probability of zero is technically impossible for that reason. However, oscillations are sustained for a large range of parameters, including regions where the survival probability is predicted to be under a half but above zero. These observations show that the G2 death signal model with the uniform cell cycle model is not ergodic, and a different approach should be taken to understand its dynamics. In conclusion, the results demonstrate that the behaviour of the G2 death signal model is sensitive to the choice of cell cycle model.

Chapter 6

The G2 death signal model for heterotypic populations

In this chapter, we apply the G2 death signal model to heterotypic populations. We focus on two cell types to keep the analysis simple, but the model can in principle be extended to more cell types. In addition, we only use the exponential cell cycle model because it conforms better to the ergodic framework that we developed in the previous chapter and that we extend in this chapter.

Importantly, the competing cell types communicate through a common mechanism and share the same cell cycle model. We impose the constraint that the cell types can vary only in their parameters (Section 6.1). Winners and losers are thus not predetermined. Instead, the aim in this chapter is to explore whether winners and losers can emerge as a result of differences in parameters, and if so, whether their behaviour reproduces the properties of cell competition.

First, we adapt the ergodic approximation to heterotypic populations (Section 6.2) to find an expression for the heterotypic survival probability (Section 6.3). Unlike the previous chapter, we cannot derive a constant survival probability. However, we can predict the sign of the heterotypic survival difference (Section 6.4) and the homotypic survival difference (Section 6.5). The former is the difference in survival probability between competing cell types in a heterotypic population, and the latter is the difference in survival probability of a cell type in a heterotypic population compared to homotypic conditions. Based on these metrics, we enumerate and classify different types of competitive interactions in Section 6.6.

A defining characteristic of cell competition is that the loser cell type is completely eliminated. It is not enough to qualify as cell competition that one cell type is more abundant than the other. Hence, we also characterise the proliferation regimes for heterotypic populations (Section 6.7).

In Section 6.8, we synthesise everything we have learned from the behaviour of the G2 death signal model to identify biologically relevant competition regimes. We find that not only can our model reproduce the defining features of cell competition without asserting *a priori* winner/loser status, but also it reveals adjacent competition regimes and their relationship to cell competition. Finally, in Section 6.9, we discuss and interpret the results.

6.1 Heterotypic populations

We consider two distinct cell types, which we name cell types A and B. Each cell type has its own cell cycle model, $\Psi_A(t)$ and $\Psi_B(t)$, death signal function, $f_A(t)$ and $f_B(t)$, and death threshold, $T_{\dagger,A}$ and $T_{\dagger,B}$. In general, the cell cycle model and death signal function can be qualitatively different. For instance, cell type A could have a uniform cell cycle model and experience a G2 death signal while cell type B has an exponential cell cycle model and experiences a constant death signal. One can also imagine a model where winner and loser identities are determined *a priori* with winners broadcasting death signals to losers, but not the other way around.

However, our aim is to produce winners and losers as an emergent phenomenon. Therefore, we will assume that the cell cycle model and death signal function are the same for both cell types, except in their parameters. In particular, we only consider cell cycle models for which[1]

$$\Psi_A(t) = \Psi(t\,; t_{\mathrm{G1},A}, t_{\mathrm{G2},A})\,, \tag{6.1}$$

$$\Psi_B(t) = \Psi(t\,; t_{\mathrm{G1},B}, t_{\mathrm{G2},B})\,, \tag{6.2}$$

where $\Psi(\cdot)$ is the shared cell cycle model. In addition, we demonstrated in Chapter 5 that the proportion of the cell cycle spent in G1 phase, β, is an important nondimensional parameter in determining the survival probability. This suggests defining β_A, $t_{\mathrm{G},A}$, β_B, and $t_{\mathrm{G},B}$ such that the following conditions are satisfied:

$$t_{\mathrm{G1},A} = \beta_A t_{\mathrm{G},A}\,, \tag{6.3}$$

$$t_{\mathrm{G2},A} = (1-\beta_A) t_{\mathrm{G},A}\,, \tag{6.4}$$

$$t_{\mathrm{G1},B} = \beta_B t_{\mathrm{G},B}\,, \tag{6.5}$$

$$t_{\mathrm{G2},B} = (1-\beta_B) t_{\mathrm{G},B}\,. \tag{6.6}$$

These equations are analogous to Equations (5.3) and (5.4). We assume that the G1 and G2

[1] In the case of the uniform cell cycle model, we also need to specify r_A and r_B.

cell cycle durations are nonzero, such that the parameters β_A and β_B lie on the interval $(0, 1)$. Moreover, we define the G2 death signal functions as

$$f_A(t) = c_A g(t)\,, \tag{6.7}$$

$$f_B(t) = c_B g(t)\,, \tag{6.8}$$

with $g(t)$ as defined in Equation (5.2). This means that both cell types receive the same G2 death signal, but process it differently based on their type-specific parameters.

6.2 Ergodic approximation

We generalise the ergodic approximation, first discussed in Section 5.2, to obtain an expression for the heterotypic survival probability of cell types A and B. First, we assume that the ergodic property holds for both cell types separately. Concretely, this means that

$$\frac{\#\text{ A cells in G2 phase}}{\#\text{ A cells in total}} \approx \frac{\text{G2 duration of A cells}}{\text{cell cycle duration of A cells}}\,, \tag{6.9}$$

$$\frac{\#\text{ B cells in G2 phase}}{\#\text{ B cells in total}} \approx \frac{\text{G2 duration of B cells}}{\text{cell cycle duration of B cells}}\,. \tag{6.10}$$

We substitute Equations (6.4) and (6.6) into Equations (6.9) and (6.10) to rewrite them as

$$\frac{\#\text{ A cells in G2 phase}}{\#\text{ A cells in total}} \approx 1 - \beta_A\,, \tag{6.11}$$

$$\frac{\#\text{ B cells in G2 phase}}{\#\text{ B cells in total}} \approx 1 - \beta_B\,. \tag{6.12}$$

We introduce the notation

$$n_A(t) = \#\text{ A cells in total}\,, \tag{6.13}$$

$$n_B(t) = \#\text{ B cells in total}\,, \tag{6.14}$$

where we have made it explicit that the subpopulation size varies over time. The fraction of cells in the G2 phase for the whole population can then be written as

$$\frac{\#\text{ cells in G2 phase}}{\#\text{ cells in total}} = \frac{\#\text{ A cells in G2 phase} + \#\text{ B cells in G2 phase}}{n_A(t) + n_B(t)}\,. \tag{6.15}$$

We substitute Equations (6.11) and (6.12) into Equation (6.15) to obtain:

$$\frac{\text{\# cells in G2 phase}}{\text{\# cells in total}} \approx \frac{n_A(t)(1-\beta_A) + n_B(t)(1-\beta_B)}{n_A(t) + n_B(t)}\,. \tag{6.16}$$

The right-hand side of Equation (6.16) is the weighted average of $1-\beta_A$ and $1-\beta_B$, with weights $n_A(t)$ and $n_B(t)$, respectively. For brevity, we introduce the following notation:

$$\langle 1-\beta\rangle(t) \equiv \frac{n_A(t)(1-\beta_A) + n_B(t)(1-\beta_B)}{n_A(t) + n_B(t)}\,. \tag{6.17}$$

Assuming that the population-level distribution of cell cycle phases is replicated at the level of the local neighbourhood of the cell, i.e. assuming Equation (5.6), we approximate the death signal as

$$g(t) \approx \frac{\text{\# cells in G2 phase}}{\text{\# cells in total}} \approx \langle 1-\beta\rangle(t)\,, \tag{6.18}$$

where we used Equations (5.6) and (6.16) in the first and second approximations, respectively. We note that Equation (5.6) expresses the condition that the population is well-mixed. In a spatially segregated heterotypic population, for example, cells mostly interact with cells of the same type, such that the local neighbourhood does *not* represent the population-level distribution of cell cycles. The death signal functions, given by Equations (6.7) and (6.8), are then approximated by

$$f_A(t) = c_A g(t) \approx c_A\langle 1-\beta\rangle(t)\,, \tag{6.19}$$

$$f_B(t) = c_B g(t) \approx c_B\langle 1-\beta\rangle(t)\,. \tag{6.20}$$

Note that the quantity $\langle 1-\beta\rangle(t)$ is not constant with respect to time because it depends on $n_A(t)$ and $n_B(t)$. This is in contrast to the homotypic case, where the death signal is approximated by the constant quantity $1-\beta$.

This means that there is no fixed death time and survival probability, even with the ergodic approximation. However, we can define the *instantaneous* death time at time t as the amount of time needed to reach the death threshold for a cell *assuming* a constant death signal of magnitude $f_A(t)$ for cell type A and $f_B(t)$ for cell type B. Similarly, the instantaneous survival probability at time t is the survival probability of a cell assuming a constant death signal of magnitude $f_A(t)$ for cell type A and $f_B(t)$ for cell type B. Concretely,

the instantaneous death times are given by

$$t_{\dagger,A}(t) = \frac{T_{\dagger,A}}{c_A\langle 1-\beta\rangle(t)}\,, \tag{6.21}$$

$$t_{\dagger,B}(t) = \frac{T_{\dagger,B}}{c_B\langle 1-\beta\rangle(t)}\,, \tag{6.22}$$

for cell types A and B, respectively. We define the instantaneous heterotypic survival probabilities as

$$\xi_{A|B}(t) = \Psi_A\left(t_{\dagger,A}(t)\right) = \Psi_A\left(\frac{T_{\dagger,A}}{c_A\langle 1-\beta\rangle(t)}\right), \tag{6.23}$$

$$\xi_{B|A}(t) = \Psi_B\left(t_{\dagger,B}(t)\right) = \Psi_B\left(\frac{T_{\dagger,B}}{c_B\langle 1-\beta\rangle(t)}\right), \tag{6.24}$$

where we use $\xi_{A|B}(t)$ to denote the instantaneous survival probability at time t of cell type A in a heterotypic population with cell type B, and vice versa for cell type B.

In summary, in this section we have derived approximate expressions for the instantaneous heterotypic survival probabilities of cell types A and B by making two assumptions. First, we assumed that the subpopulations of A-type and B-type cells are ergodic in Equations (6.9) and (6.10). This is analogous to the assumption of ergodicity for homotypic populations, as stated in Equation (5.5). Second, we assumed that the cell cycle phase distribution in the local cell neighbourhood is approximately representative of the distribution in the population as a whole, thus applying Equation (5.6) to the heterotypic case. With these two assumptions, we derived an ergodic approximation to the death signal. However, this ergodic death signal is dependent on the system state and thus varies with time, so we defined the instantaneous death time and survival probability. For brevity, hereafter we omit the word "instantaneous" and use $\langle 1-\beta\rangle$ and $\xi_{A|B}$ instead of $\langle 1-\beta\rangle(t)$ and $\xi_{A|B}(t)$, unless we wish to emphasise their time dependence.

6.3 Heterotypic survival probability for exponential cell cycle model

In this section, we derive an expression for the heterotypic survival probability with the exponential cell cycle model under the ergodic approximation. First, we substitute Equations (6.3) and (6.5) into the respective cumulative distribution functions for the exponential

distribution, Equation (3.10), to give

$$\Psi_A(t) = 1 - \exp\left(-\frac{t}{\beta_A t_{G,A}}\right), \tag{6.25}$$

$$\Psi_B(t) = 1 - \exp\left(-\frac{t}{\beta_B t_{G,B}}\right). \tag{6.26}$$

Substituting the death times, given in Equations (6.21) and (6.22), into Equations (6.25) and (6.26), we obtain the heterotypic survival probabilities

$$\xi_{A|B} = 1 - \exp\left(-\frac{T_{\dagger,A}}{c_A t_{G,A}\beta_A\langle 1-\beta\rangle}\right), \tag{6.27}$$

$$\xi_{B|A} = 1 - \exp\left(-\frac{T_{\dagger,B}}{c_B t_{G,B}\beta_B\langle 1-\beta\rangle}\right). \tag{6.28}$$

Analogously with Equation (5.13), we define the dimensionless parameters

$$\eta_A \equiv \frac{T_{\dagger,A}}{c_A t_{G,A}}, \tag{6.29}$$

$$\eta_B \equiv \frac{T_{\dagger,B}}{c_B t_{G,B}}, \tag{6.30}$$

so that we can rewrite the heterotypic survival probabilities as

$$\xi_{A|B}\left(\beta_A, \eta_A, \langle 1-\beta\rangle\right) = 1 - \exp\left(-\frac{\eta_A}{\beta_A\langle 1-\beta\rangle}\right), \tag{6.31}$$

$$\xi_{B|A}\left(\beta_B, \eta_B, \langle 1-\beta\rangle\right) = 1 - \exp\left(-\frac{\eta_B}{\beta_B\langle 1-\beta\rangle}\right). \tag{6.32}$$

These are similar to the expression obtained in Equation (5.14) for homotypic populations, except that the weighted average $\langle 1-\beta\rangle$ is used instead of $1-\beta$. Moreover, if $n_B = 0$, then $\langle 1-\beta\rangle = 1-\beta_A$ and vice versa for $n_A = 0$. Hence, when either cell type A or B is absent, we recover the homotypic survival probability of the remaining cell type, and thus Equations (6.31) and (6.32) generalise Equation (5.14).

For the sake of simplicity, we assume that $t_{\mathrm{G},A} = t_{\mathrm{G},B}$, such that differences in survival probability alone determine relative proliferative success. In the general case, however, the ratio

$$\kappa_{A/B} \equiv \frac{t_{\mathrm{G},A}}{t_{\mathrm{G},B}}, \tag{6.33}$$

is a dimensionless parameter that also modulates the dynamics of the heterotypic population. For instance, a cell type with a lower survival probability may become more abundant

than the competing cell type by having a shorter cell cycle time and dividing more quickly. However, for simplicity we do not consider such cases here. Instead, in this chapter we focus on the case $\kappa_{A/B} = 1$, and characterise population dynamics solely in terms of survival probabilities.

6.4 Heterotypic survival difference

The instantaneous heterotypic survival probabilities $\xi_{A|B}(t)$ and $\xi_{B|A}(t)$ are dependent on system state because the ergodic death signal $\langle 1 - \beta \rangle(t)$ depends on the distribution of cell types in the population. However, in this section we show that the sign of their difference is invariant with respect to system state for the exponential cell cycle model. Furthermore, we express the sign in terms of model parameters and run simulations to verify this prediction.

We define the **heterotypic survival difference** between cell types A and B as

$$\Delta^{\neq}_{A|B} \equiv \xi_{A|B} - \xi_{B|A} \,, \tag{6.34}$$

where $\xi_{A|B}$ and $\xi_{B|A}$ are the heterotypic survival probabilities of cell types A and B as given in Equations (6.27) and (6.28), respectively. The sign of the heterotypic survival difference tells us which cell type is at a proliferative advantage. If $\Delta^{\neq}_{A|B} > 0$, then we say that A-type cells are **winner cells** and B-type cells are **loser cells**, and vice versa for $\Delta^{\neq}_{A|B} < 0$. Moreover, if $\Delta^{\neq}_{A|B} = 0$, then we say that the cell types are in **coexistence**, since neither cell type has a fitness advantage over the other.

For the exponential cell cycle model in particular, we insert Equations (6.31) and (6.32) into Equation (6.34) and rearrange to obtain

$$\Delta^{\neq}_{A|B} = \exp\left(-\frac{\eta_B}{\beta_B \langle 1 - \beta \rangle}\right) - \exp\left(-\frac{\eta_A}{\beta_A \langle 1 - \beta \rangle}\right) . \tag{6.35}$$

Because $\exp(\cdot)$ is a monotonically increasing function, the following identity holds:

$$\operatorname{sgn}\left(\exp(x) - \exp(y)\right) = \operatorname{sgn}(x - y) \,. \tag{6.36}$$

Applying the sign function to Equation (6.35) thus yields

$$\begin{aligned}\operatorname{sgn}\left(\Delta^{\neq}_{A|B}\right) &= \operatorname{sgn}\left(\frac{\eta_A}{\beta_A\langle 1-\beta\rangle} - \frac{\eta_B}{\beta_B\langle 1-\beta\rangle}\right) \\ &= \operatorname{sgn}\left(\frac{1}{\langle 1-\beta\rangle}\right)\operatorname{sgn}\left(\frac{\eta_A}{\beta_A} - \frac{\eta_B}{\beta_B}\right) \\ &= \operatorname{sgn}\left(\frac{\eta_A}{\beta_A} - \frac{\eta_B}{\beta_B}\right), \end{aligned} \tag{6.37}$$

where we used the fact that $\langle 1-\beta\rangle > 0$ in the last line.

Since we factored out $\langle 1-\beta\rangle$ to derive Equation (6.37), we note that we can replace $\langle 1-\beta\rangle$ with any non-negative death signal in Equations (6.19) and (6.20), and we would still obtain the same result as in Equation (6.37). Moreover, as discussed in the previous chapter, η and β together determine sensitivity to death signals, with lower η and higher β corresponding to higher sensitivity to death signals. Hence, we can interpret η/β as a cell type's tolerance to death signals. Therefore, Equation (6.37) states that, in a heterotypic population, the relative sensitivity to death signals determines winner/loser status, with the least sensitive cell type becoming the winner.

6.4.1 Coexistence curve

Equation (6.37) is an expression for the sign of the heterotypic survival difference that is only dependent on the model parameters, and independent of $\langle 1-\beta\rangle(t)$. In other words, the winner and loser status is invariant with respect to the system state. We can therefore partition parameter space into two regions for which $\Delta^{\neq}_{A|B} > 0$ and $\Delta^{\neq}_{A|B} < 0$, respectively.

The border between these two regions is given by the hypersurface $\Delta^{\neq}_{A|B} = 0$. When taking cross sections of parameter space for fixed values of β_B and η_B, the hypersurface $\Delta^{\neq}_{A|B} = 0$ becomes a curve. We define the **coexistence curve** as the curve in (β_A, η_A)-space that satisfies $\Delta^{\neq}_{A|B} = 0$. Using Equation (6.37), we find that the coexistence curve is given by

$$\frac{\eta_A}{\beta_A} - \frac{\eta_B}{\beta_B} = 0\,. \tag{6.38}$$

We will use the coexistence curve in Section 6.6 to classify competitive interactions.

6.4.2 Computational validation

Equation (6.37) provides a concrete prediction of the sign of the heterotypic survival difference based solely on parameter values. In addition, Figure 6.7 shows how the regions

Table 6.1: Model and simulation parameter values used to estimate the heterotypic survival difference for the well-mixed model.

Parameter	Cell type A	Cell type B
t_G		100
c		1
η	$0.01, 0.02, \ldots, 0.25$	$0.05, 0.1, 0.2$
β	$0.05, 0.10, \ldots, 0.95$	$0.3, 0.5, 0.7$
Initial cell count	50	50
Simulation end time		10 000
Minimum cell count		10
Maximum cell count		1 000
N_{sim}		50

$\Delta^{\neq}_{A|B} > 0$ and $\Delta^{\neq}_{A|B} < 0$ are situated in parameter space for fixed values of β_B and η_B. In this section, we validate these theoretical results with simulations. Concretely, we used a Monte Carlo method to estimate the survival frequency of both cell types for different parameter values, and compared the difference in survival frequency with predictions.

Parameter choice Since there are two cell types, we have two times as many parameters as in a homotypic population. The parameter space is therefore much larger than in the previous chapter. In particular, we want to confirm our predictions in $(\beta_A, \eta_A, \beta_B, \eta_B)$-space. In Section 5.4.1 (Parameter choice, page 104), we used a regular grid in (β, η)-space to construct a parameter sweep. However, extending this approach here would result in a simulation suite that is 500 times as large. In addition, visualising the results would be challenging.

Therefore, we chose a limited number of values for β_B and η_B, but varied β_A and η_A along a regular grid. For the well-mixed model, we picked $\beta_B = 0.3, 0.5, 0.7$ and $\eta_B = 0.05, 0.1, 0.2$. For cell type A, we evenly sampled 19 values for β_A from the range $[0, 1]$ and 25 values for η_A from the range $[0, 0.25]$. For the vertex model, we picked $\beta_B = 0.5, \eta_B = 0.1$, 9 values for β_A, and 12 values for η_A. The mechanical parameters for both cell types are set to the default values, as defined in Chapter 2. See Tables 6.1 and 6.2 for a summary of the parameter values for the well-mixed and vertex models, respectively.

To compute the death thresholds, we applied the approach in Section 5.3.1 (Parameter choice, page 93) to both cell types; we fixed $t_{G,A}, c_A, t_{G,B}, c_B$ and computed the death

Table 6.2: Model and simulation parameter values used to estimate the heterotypic survival difference for the vertex model.

Parameter	Cell type A	Cell type B
t_G		100
c		1
η	$0.02, 0.04, \ldots, 0.24$	0.1
β	$0.1, 0.2, \ldots, 0.9$	0.5
Initial cell count	50	50
Pattern	random, segregated	
Simulation end time		10 000
Minimum cell count		10
Maximum cell count		1 000
N_{sim}		20

thresholds using the following expressions[2]:

$$T_{\dagger,A} = \eta_A c_A t_{G,A} \,, \tag{6.39}$$

$$T_{\dagger,B} = \eta_B c_B t_{G,B} \,. \tag{6.40}$$

For the well-mixed model, we ran 50 simulations for every unique parameter set, so the total number of simulations is $25 \times 19 \times 3 \times 3 \times 50 = 213\,750$. For the vertex model, we additionally combined every parameter set with two different initial conditions (described below), and ran 20 simulations per unique combination of parameter set and type of initial conditions. Hence, the total number of simulations is $12 \times 9 \times 2 \times 20 = 4\,320$.

Initial conditions In Section 5.3.1 (Initial conditions, page 95), we described the initial conditions for a homotypic population for the well-mixed and vertex models. We took a similar approach here, except that each cell was initialised according to their type. In particular, the birth times and G1 durations were assigned using the cell cycle parameters $t_{\text{G1},A}, t_{\text{G2},A}$ for cell type A, and $t_{\text{G1},B}, t_{\text{G2},B}$ for cell type B. Furthermore, we initialised each simulation with 50 A-type cells and 50 B-type cells. For the vertex model we also need to specify the initial spatial configuration. In order to test the effect of spatial segregation, we arranged the cell types in a "segregated" or "random" pattern as in Section 2.2.2. The initial conditions for all simulations in this chapter are determined in this manner.

[2]These expressions follow from the definition of η_A and η_B in Equations (6.29) and (6.30), respectively.

Termination conditions We set a minimum cell count of 10, a maximum cell count of 1 000, and a simulation end time of 10 000. In addition, we also terminated the simulation as soon as any cell type goes extinct, i.e. if either $n_A(t) = 0$ or $n_B(t) = 0$, since the simulation then becomes homotypic. Note that a homotypic population cannot go extinct because the G2 death signal is derived from other cells, but a cell type in a heterotypic population can go extinct because it can receive death signals from the surviving cell type. Hence, there are a total of four ways that the simulation can terminate: the simulation hits the minimum cell count, the simulation hits the maximum cell count, cell type A or B goes extinct, or the simulation runs until the end time.

Data processing For each simulation, we compute the survival frequencies, $\hat{\xi}_{A|B}$ and $\hat{\xi}_{B|A}$, using Equations (2.13) and (2.14), respectively. Each parameter set was simulated N_{sim} times with different random seeds. To estimate the survival frequency for a particular parameter set, we averaged the heterotypic survival frequencies as

$$\overline{\xi_{A|B}} = \frac{1}{N_{\text{sim}}} \sum_{k=1}^{N_{\text{sim}}} \hat{\xi}_{A|B,k} \,, \tag{6.41}$$

$$\overline{\xi_{B|A}} = \frac{1}{N_{\text{sim}}} \sum_{k=1}^{N_{\text{sim}}} \hat{\xi}_{B|A,k} \,. \tag{6.42}$$

The estimator for the heterotypic survival difference is then given by

$$\overline{\Delta^{\neq}_{A|B}} = \overline{\xi_{A|B}} - \overline{\xi_{B|A}} \,. \tag{6.43}$$

Data visualisation We plotted $\overline{\Delta^{\neq}_{A|B}}$ for fixed values of β_B and η_B using a heat map, with β_A on the horizontal axis and η_A on the vertical axis. In addition, for the vertex model we decomposed the results further into segregated and random initial conditions. Similarly to Section 5.4.1 (Data visualisation, page 105), we want to highlight the critical value $\overline{\Delta^{\neq}_{A|B}} = 0$ separating the regions where $\overline{\Delta^{\neq}_{A|B}} < 0$ and $\overline{\Delta^{\neq}_{A|B}} > 0$, respectively. Hence, we used the diverging *seismic* colour map. Denoting the maximum absolute value of $\overline{\Delta^{\neq}_{A|B}}$ over all parameter sets as $|\Delta^{\neq}|$, we set the range of the colour map to $[-|\Delta^{\neq}|, |\Delta^{\neq}|]$ with central value zero. Hence, red regions indicate positive values and blue regions indicate negative values.

We marked the degenerate case $\beta_A = \beta_B$, $\eta_A = \eta_B$ with a green dot and plotted the coexistence curve with a solid line for comparison with theoretical predictions. The results are given in Figures 6.1 and 6.2 for the well-mixed and vertex models, respectively.

Computational materials and methods See Computational materials and methods in Section 2.3.1 (page 36) for the vertex model and in Section 4.4 (page 77) for the well-mixed model.

6.4.3 Results

Well-mixed model Figure 6.1 shows that the separation between $\overline{\Delta^{\neq}_{A|B}} > 0$ and $\overline{\Delta^{\neq}_{A|B}} < 0$ is generally consistent with the predicted coexistence curve. There are some deviations for $\eta_B = 0.05$ where the blue region is shaped more like a blob than a triangle. The overall structure of the parameter space remains intact, however.

Vertex model We report that 6 out of 4 320 Chaste simulations failed. Figure 6.2 shows that the observed parameter regions match the analytical predictions reasonably well for the vertex model. The largest deviation is found in the top right corner of the bottom triangle, where the red region transgresses the predicted coexistence curve. Moreover, comparing the range of the colour bar with Figure 6.1 reveals that the observed values for $\overline{\Delta^{\neq}_{A|B}}$ are less extreme than for the well-mixed model.

This effect is more pronounced for segregated initial conditions than for random initial conditions. In particular, the blue region in the bottom right corner is lighter. In addition, the red region under the coexistence curve is larger for segregated simulations. It thus seems that spatial segregation results in larger deviations from the well-mixed results, suggesting that spatial mixing may affect cell competition.

6.5 Homotypic survival difference

A consistent discrepancy in survival probabilities between cell types, i.e. $\Delta^{\neq}_{A|B} \neq 0$, indicates that they are participating in some form of competition. However, by itself it does not prove that a competitive interaction is taking place. After all, two non-interacting cell types with a different intrinsic survival probability have a nonzero heterotypic survival difference despite not interacting. Therefore, in this section we define a metric to test whether cells experience a change in their survival probability as a result of competitive interactions.

Firstly, the homotypic survival probabilities of cell types A and B are defined as

$$\lambda_A = \Psi_A\left(\frac{T_{\dagger,A}}{c_A(1-\beta_A)}\right), \tag{6.44}$$

$$\lambda_B = \Psi_B\left(\frac{T_{\dagger,B}}{c_B(1-\beta_B)}\right), \tag{6.45}$$

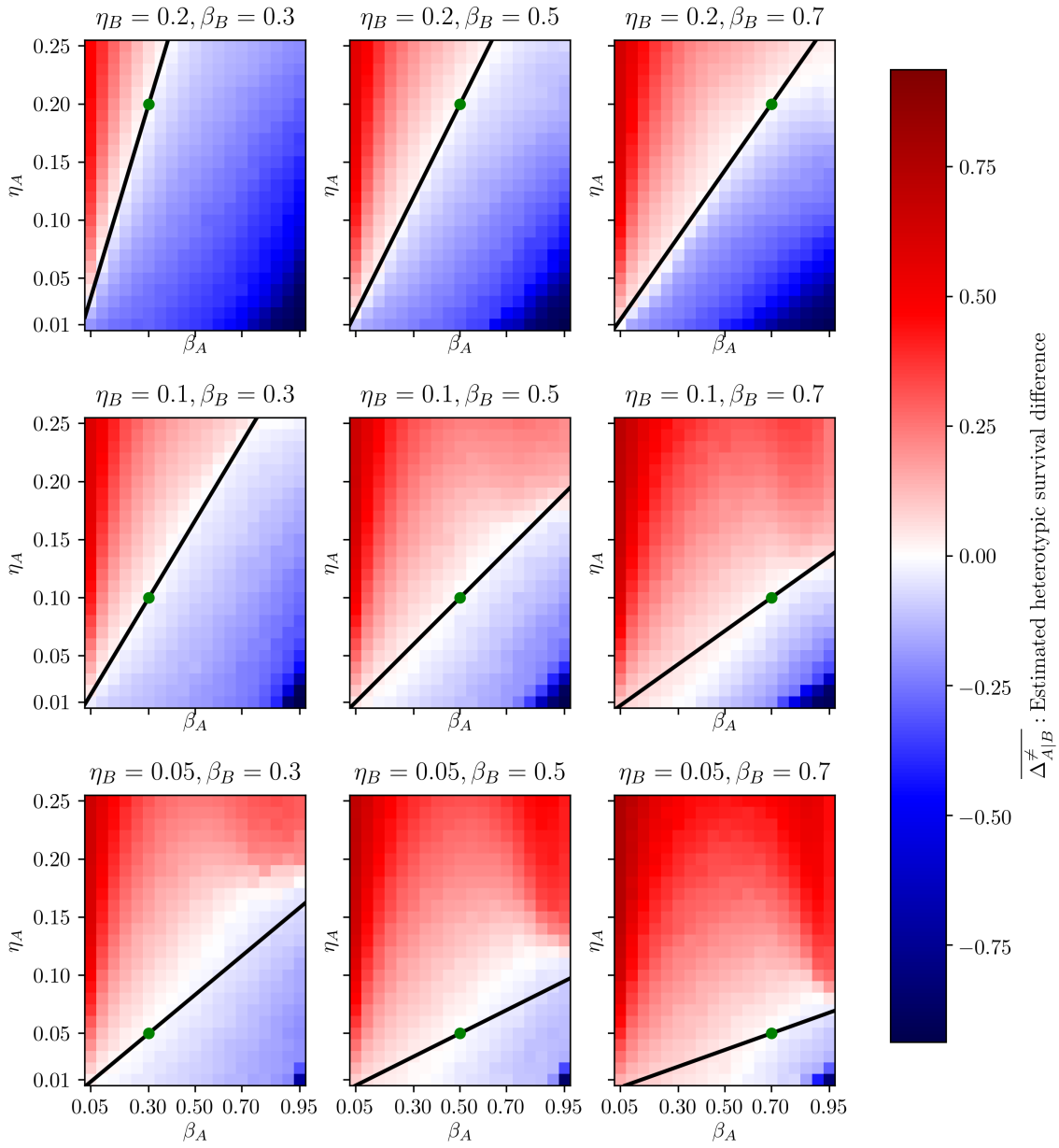

Figure 6.1: Estimated heterotypic survival difference, $\overline{\Delta^{\neq}_{A|B}}$, defined in Equation (6.43), for the well-mixed model. The solid line and the green dot correspond to the coexistence curve and the point ($\beta_A = \beta_B, \eta_A = \eta_B$), respectively.

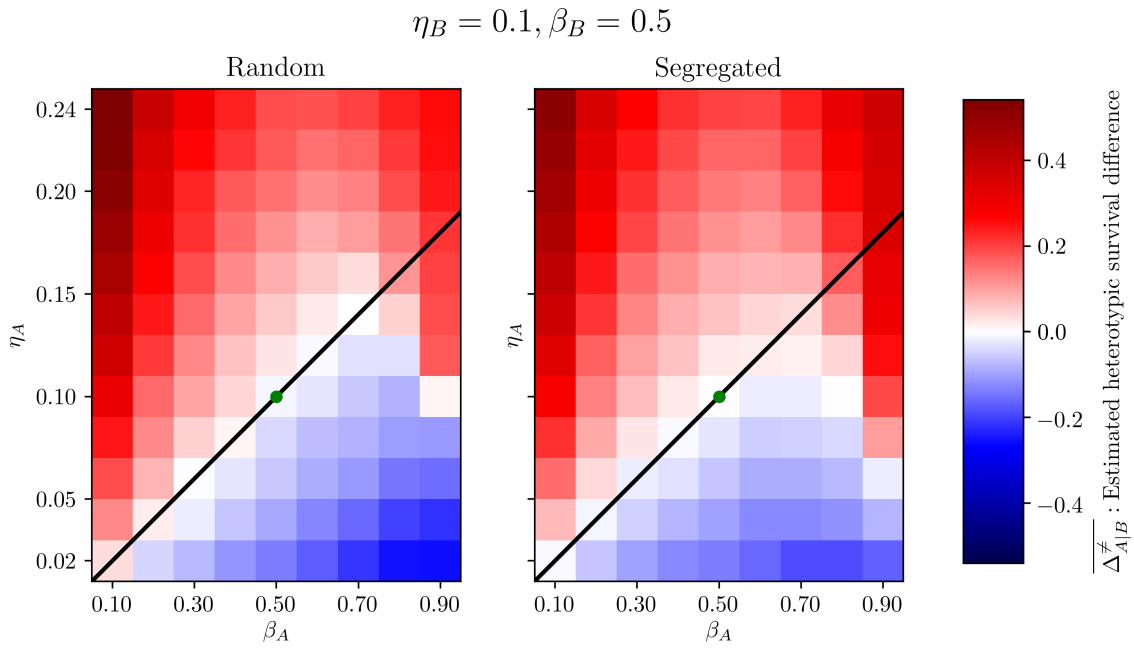

Figure 6.2: Estimated heterotypic survival difference, $\overline{\Delta^{\neq}_{A|B}}$, defined in Equation (6.43), for the vertex model. The solid line and the green dot correspond to the coexistence curve and the point $(\beta_A = \beta_B, \eta_A = \eta_B)$, respectively.

respectively. This is the survival probability that either cell type would experience in a homotypic environment, analogous to Equation (5.10) in Chapter 5. The **homotypic survival difference** is then defined as

$$\Delta^{=}_{A|B} \equiv \xi_{A|B} - \lambda_A \,, \tag{6.46}$$

$$\Delta^{=}_{B|A} \equiv \xi_{B|A} - \lambda_B \,, \tag{6.47}$$

for cell types A and B, respectively. The homotypic survival difference compares the fitness of a cell type in the heterotypic environment to its fitness in a homotypic environment. For a concrete interpretation, suppose that the population is initially homotypic for cell type A, i.e. only A-type cells populate the population. Then, a cell mutates and divides to give rise to a group of B-type cells, which renders the population heterotypic. The change in fitness for cell type A that is caused by the introduction of cell type B is the homotypic survival difference $\Delta^{=}_{A|B}$, and vice versa for $\Delta^{=}_{B|A}$.

The sign of the homotypic survival difference indicates whether a cell type becomes more or less fit as a result of the heterotypic interaction, compared to its homotypic fitness. If $\Delta^{=}_{A|B} > 0$, then we say that cell type A is more fit when competing with cell type B, and vice versa for $\Delta^{=}_{A|B} < 0$. Even though a positive sign for $\Delta^{=}_{A|B}$ indicates that cell type A benefits

from the interaction, it does not mean that the interaction is mutualistic, since that would require both B-type and A-type cells to benefit from the interaction simultaneously (i.e. $\Delta^{=}_{A|B}, \Delta^{=}_{B|A} > 0$). We show below that this situation is impossible when using an exponential cell cycle model under the ergodic approximation. Finally, if $\Delta^{=}_{A|B} = 0$, then we say that A-type cells are in **neutral competition** with B-type cells, since the interaction with B-type cells does not produce a net change in absolute fitness of A-type cells.

In the case of an exponential cell cycle model specifically, the homotypic survival probabilities are

$$\lambda_A = 1 - \exp\left(-\frac{\eta_A}{\beta_A(1-\beta_A)}\right), \tag{6.48}$$

$$\lambda_B = 1 - \exp\left(-\frac{\eta_B}{\beta_B(1-\beta_B)}\right). \tag{6.49}$$

Focusing our derivation of the homotypic survival difference on cell type A, we have

$$\begin{aligned}\Delta^{=}_{A|B} &= \left[1 - \exp\left(-\frac{\eta_A}{\beta_A\langle 1-\beta\rangle}\right)\right] - \left[1 - \exp\left(-\frac{\eta_A}{\beta_A(1-\beta_A)}\right)\right] \\ &= \exp\left(-\frac{\eta_A}{\beta_A(1-\beta_A)}\right) - \exp\left(-\frac{\eta_A}{\beta_A\langle 1-\beta\rangle}\right).\end{aligned} \tag{6.50}$$

Similarly to Section 6.4, we apply the sign function to arrive at

$$\begin{aligned}\operatorname{sgn}\left(\Delta^{=}_{A|B}\right) &= \operatorname{sgn}\left(\frac{\eta_A}{\beta_A\langle 1-\beta\rangle} - \frac{\eta_A}{\beta_A(1-\beta_A)}\right) \\ &= \operatorname{sgn}\left(\frac{\eta_A}{\beta_A}\right)\operatorname{sgn}\left(\frac{1}{\langle 1-\beta\rangle} - \frac{1}{1-\beta_A}\right) \\ &= \operatorname{sgn}\left(\frac{1}{\langle 1-\beta\rangle} - \frac{1}{1-\beta_A}\right).\end{aligned} \tag{6.51}$$

Expanding $\langle 1-\beta\rangle(t)$ (see definition in Equation (6.17)) in the argument of Equation (6.51), we have

$$\begin{aligned}\frac{1}{\langle 1-\beta\rangle(t)} - \frac{1}{1-\beta_A} &= \frac{n_A(t) + n_B(t)}{n_A(t)(1-\beta_A) + n_B(t)(1-\beta_B)} - \frac{1}{1-\beta_A} \\ &= \frac{n_B(t)(\beta_B - \beta_A)}{\left[n_A(t)(1-\beta_A) + n_B(t)(1-\beta_B)\right](1-\beta_B)},\end{aligned} \tag{6.52}$$

where we note that the denominator is always positive. Equation (6.52) shows that the sign of $\Delta^{=}_{A|B}$ is dependent on the system state. Specifically, in the degenerate case of $n_B(t) = 0$, we

are reduced to a homotypic population composed solely of A-type cells, and thus $\Delta^{=}_{A|B} = 0$. If we limit our scope to only the heterotypic case, i.e. $n_A(t), n_B(t) > 0$, we can rewrite Equation (6.51) using Equation (6.52) as

$$\operatorname{sgn}\left(\Delta^{=}_{A|B}\right) = \operatorname{sgn}(\beta_B - \beta_A)\,. \tag{6.53}$$

For cell type B, we derive an analogous expression of the form

$$\operatorname{sgn}\left(\Delta^{=}_{B|A}\right) = \operatorname{sgn}(\beta_A - \beta_B)\,. \tag{6.54}$$

Comparing Equations (6.53) and (6.54), we see that the following identity holds:

$$\operatorname{sgn}\left(\Delta^{=}_{A|B}\right) = -\operatorname{sgn}\left(\Delta^{=}_{B|A}\right)\,. \tag{6.55}$$

Equation (6.55) shows us that the homotypic survival differences for two competing cell types using the exponential cell cycle model must have opposite signs. Hence, one cell type's loss is another cell type's gain, and a mutualistic relationship is impossible.

In Section 6.4, we factored out the death signal $\langle 1 - \beta \rangle$ to find an expression for the sign of $\Delta^{\neq}_{A|B}$ and found that winner/loser status is determined by the difference in tolerance to death signals. Here, in contrast, we factored out the tolerance to death signals, η_A/β_A, in Equation (6.51) to find that the sign of the homotypic survival difference depends on the difference between the emission of death signal in homotypic conditions, $1 - \beta_A$, and in heterotypic conditions, $\langle 1 - \beta \rangle$. We can see this by rewriting Equation (6.51) as

$$\operatorname{sgn}\left(\Delta^{=}_{A|B}\right) = \operatorname{sgn}\Big((1 - \beta_A) - \langle 1 - \beta \rangle\Big)\,. \tag{6.56}$$

This suggests that we can rewrite Equation (6.53) as a difference in the homotypic emission of death signals, i.e.

$$\operatorname{sgn}\left(\Delta^{=}_{A|B}\right) = \operatorname{sgn}\Big((1 - \beta_A) - (1 - \beta_B)\Big)\,. \tag{6.57}$$

In words, Equation (6.57) states that cell type A fares better in heterotypic conditions if cell type B has lower intrinsic emission of death signal than cell type A and vice versa.

6.5.1 Neutral competition curve

Similarly to Section 6.4, Equations (6.53) and (6.54) reveal that the signs of $\Delta^{=}_{A|B}$ and $\Delta^{=}_{B|A}$ are independent of the system state, except in the degenerate homotypic cases $n_A(t) = 0$ or $n_B(t) = 0$. Likewise, this distinction lets us partition parameter space into two regions: one

where $\Delta^{=}_{A|B} > 0, \Delta^{=}_{B|A} < 0$, and one where $\Delta^{=}_{A|B} < 0, \Delta^{=}_{B|A} > 0$.

These regions are separated by the hypersurface $\Delta^{=}_{A|B} = 0$, or, equivalently, $\Delta^{=}_{B|A} = 0$. We define the **neutral competition curve** as the curve in (β_A, η_A)-space that satisfies $\Delta^{=}_{A|B} = 0$ for fixed values of β_B and η_B. Using Equation (6.53), we derive that the neutral competition curve is given by

$$\beta_B - \beta_A = 0\,. \tag{6.58}$$

We visualise the neutral competition curve in Figure 6.7 (page 152). We will use the neutral competition curve in Section 6.6 to classify competitive interactions.

6.5.2 Computational validation

Equations (6.53) and (6.54) predict the sign of the homotypic survival difference for cell types A and B, respectively. In this section, we verify whether this result is supported by simulations. The homotypic survival difference is calculated by subtracting the homotypic survival probability from the heterotypic survival probability. Therefore, in order to estimate the homotypic survival difference, we need to run both homotypic and heterotypic simulations to estimate the survival probability in each, and subtract the results. We have already run both types of simulations in Sections 5.4.1 and 6.4.2, however, so we combined the data from those sections to estimate the homotypic survival difference instead of running additional simulations.

Parameter choice We use the same parameters listed in Section 5.4.1 (Parameter choice, page 104) and Section 6.4.2 (Parameter choice, page 138) to estimate the homotypic and heterotypic survival probabilities, respectively. The parameters in those sections were chosen such that for every value for (β_A, η_A) in the heterotypic simulations, there is a corresponding parameter set in the homotypic simulations for which $(\beta = \beta_A, \eta = \eta_A)$, and likewise for cell type B.

Data processing We used the data from Section 6.4.2 for the heterotypic survival frequencies $\overline{\xi_{A|B}}$ and $\overline{\xi_{B|A}}$, and the data from Section 5.4.1 for the homotypic survival frequencies $\overline{\lambda_A}$ and $\overline{\lambda_B}$. We estimated the homotypic survival differences as follows:

$$\overline{\Delta^{=}_{A|B}} = \overline{\xi_{A|B}} - \overline{\lambda_A}\,, \tag{6.59}$$

$$\overline{\Delta^{=}_{B|A}} = \overline{\xi_{B|A}} - \overline{\lambda_B}\,, \tag{6.60}$$

where $\overline{\lambda_A}$ indicates that the homotypic survival frequency was computed from the homotypic simulations that match the parameters of cell type A, and similarly for $\overline{\lambda_B}$.

Data visualisation We plot $\overline{\Delta^{=}_{A|B}}$ and $\overline{\Delta^{=}_{B|A}}$ for fixed values of β_B and η_B using heat maps. We also split the results between segregated and random initial conditions for the vertex model. We denote the maximum absolute values of $\overline{\Delta^{=}_{A|B}}$ and $\overline{\Delta^{=}_{B|A}}$ across all parameter sets with $|\Delta^{\neq}_{A|B}|$ and $|\Delta^{\neq}_{B|A}|$, respectively. The colour map ranges are set to $[-|\Delta^{\neq}_{A|B}|, |\Delta^{\neq}_{A|B}|]$ and $[-|\Delta^{\neq}_{B|A}|, |\Delta^{\neq}_{B|A}|]$ with central values zero for the heat maps of $\overline{\Delta^{=}_{A|B}}$ and $\overline{\Delta^{=}_{B|A}}$, respectively. Hence, red regions correspond to positive values and blue regions correspond to negative values.

We mark the degenerate case $\beta_A = \beta_B$, $\eta_A = \eta_B$ with a green dot and plot the neutral competition curve with a dashed line for comparison with theoretical predictions. The well-mixed model results for $\overline{\Delta^{=}_{A|B}}$ and $\overline{\Delta^{=}_{B|A}}$ are given in Figures 6.3 and 6.4, respectively. The vertex model results for $\overline{\Delta^{=}_{A|B}}$ and $\overline{\Delta^{=}_{B|A}}$ are given in Figures 6.5 and 6.6, respectively.

6.5.3 Results

Well-mixed model Figures 6.3 and 6.4 show that the neutral curve is in good agreement with the border between red and blue regions. For $\overline{\Delta^{=}_{A|B}}$ in Figure 6.3, we see that the parameter space is mostly dominated by blue regions where cell type A has a lower survival probability than in the homotypic case, whereas the red regions are relatively faint. The relative faintness of the red regions compared to the blue regions suggests that the death signal mechanism generally does not confer a fitness advantage and that cell competition is more driven by the elimination of loser cells rather than the elevation of winner cells.

When we look at the parameter space to the right of the neutral curve, we can distinguish roughly three subregions. First, there is a dark blue region that runs along the right-hand edge of the parameter space with a triangular shape. Second, the dark blue fades into light blue and eventually white as it approaches the neutral curve on the bottom left. Third, the dark blue fades into lighter blue towards the top. For clarity, we outlined these subregions in the central plot of Figure 6.3

In the dark blue region, B-type cells significantly lower the viability of A-type cells compared to homotypic conditions. This region transitions into lighter blue towards the bottom left because there it overlaps with the homotypic Nonviable Regime, which was discussed in Section 5.4 and is sketched in Figure 5.7. In this subregion, A-type cells are homotypically nonviable to start with, hence there is less room for B-type cells to decrease the survival rate of A-type cells even further. On the other hand, as η_A is increased, A-type

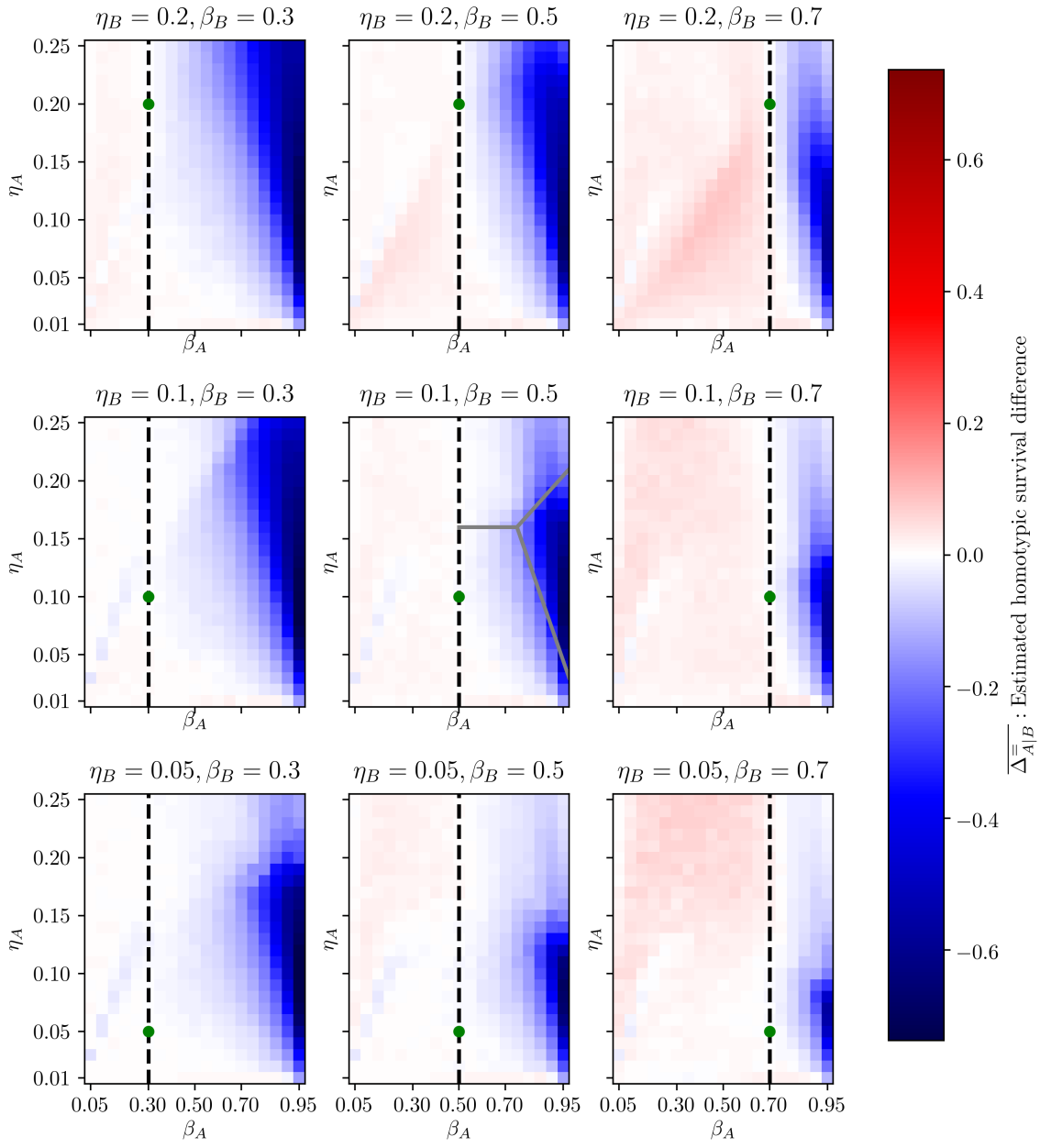

Figure 6.3: Estimated homotypic survival difference of cell type A, $\overline{\Delta^{=}_{A|B}}$, defined in Equation (6.59), for the well-mixed model. The dashed line and the green dot correspond to the neutral competition curve and the point ($\beta_A = \beta_B, \eta_A = \eta_B$), respectively.

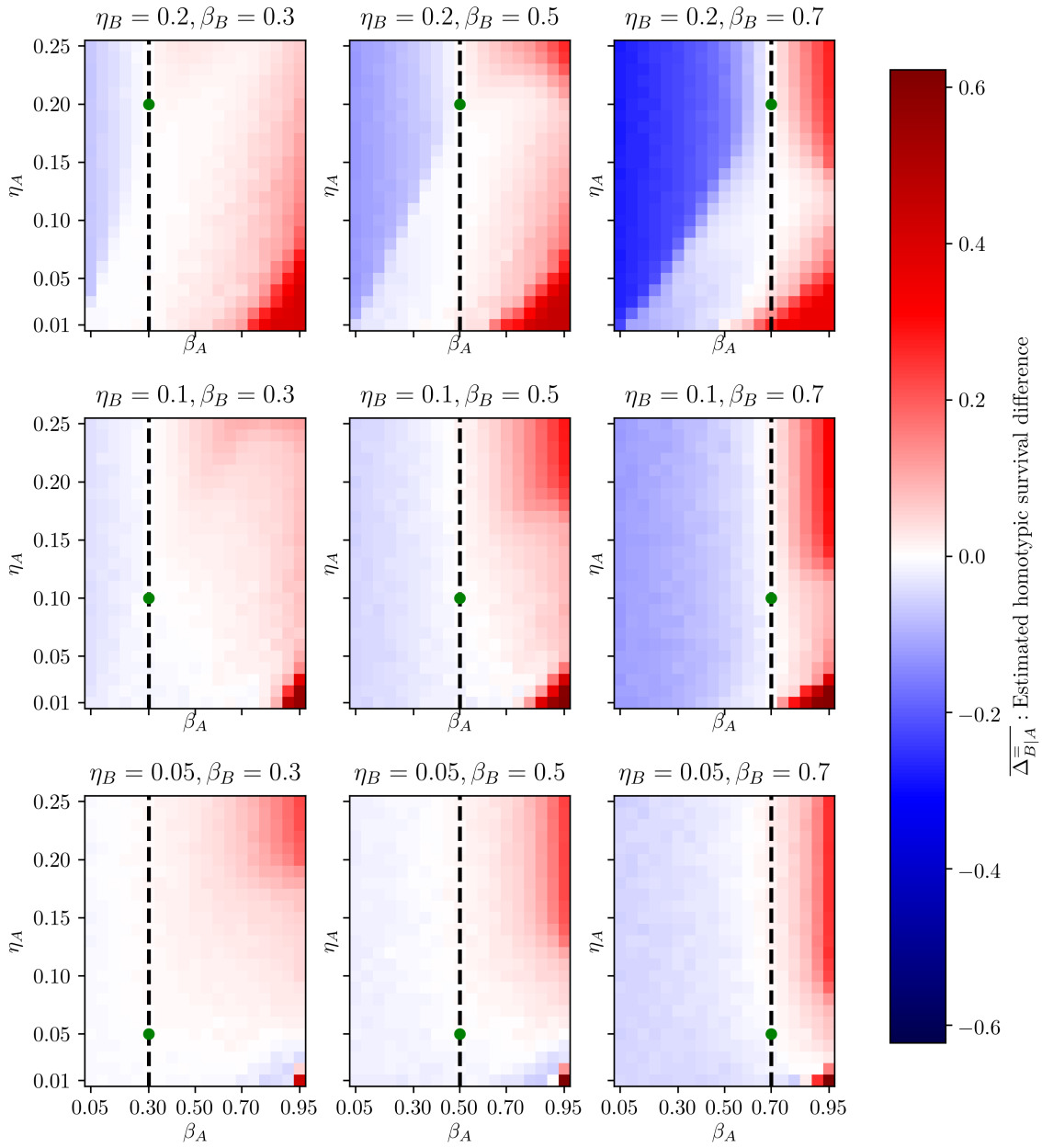

Figure 6.4: Estimated homotypic survival difference of cell type B, $\overline{\Delta^{=}_{B|A}}$, defined in Equation (6.60), for the well-mixed model. The dashed line and the green dot correspond to the neutral competition curve and the point ($\beta_A = \beta_B, \eta_A = \eta_B$), respectively.

cells are increasingly insensitive to death signals, hence the impact of a cell type with a higher emission of death signal is curtailed, so we see lighter shades of blue towards the top.

For $\overline{\Delta^{=}_{B|A}}$, we see in Figure 6.4 more pronounced red regions, particularly in the bottom right corner. This corresponds to the dark blue bottom right corner for $\overline{\Delta^{\neq}_{A|B}}$ in Figure 6.1. Inspection of individual heterotypic simulations (not shown) reveals that A-type cells in this region perish so quickly that B-type cells do not even have the opportunity to initiate apoptosis before the A-type cells go extinct and the simulation terminates. As a result, the heterotypic survival frequency of B-type cells is recorded as equal or close to unity, translating into a high homotypic survival difference. This behaviour is so robust that for $\eta_B = 0.2$, $\beta_B = 0.7$ the red region intersects the neutral curve. It should be noted, however, that this is more of a simulation artefact rather than a biologically meaningful phenomenon.

Above this dark red region and to the right of the neutral curve, we see a lighter middle region, and, following that, a darker upper region. In this lighter middle region, A-type cells are also going extinct, but not so quickly that the heterotypic survival frequency of B-type cells is recorded as unity. This results in a lower, but still positive, homotypic survival difference. In the upper region, A-type cells have a sufficiently high tolerance to persist alongside B-type cells. As a result, B-type cells consistently experience a lower average death signal compared to homotypic conditions, which is why we see an elevated homotypic survival difference.

For $\eta_B = 0.2$, $\beta_B = 0.7$ in Figure 6.4, we see a dark blue region in the top-left corner that is bounded to the right by the neutral curve and bounded below by the coexistence curve. Above the coexistence curve, A-type cells are more numerous than B-type cells, so they can exert their comparatively higher level of death signal to diminish the viability of B-type cells. Below the coexistence curve, however, B-type cells are more numerous than A-type cells, hence the impact of A-type cells is curbed because there are fewer of them.

Vertex model In Figures 6.5 and 6.6, we see that the predicted neutral curve agrees well with that established via simulation. In addition, the blue and red regions have the same overall shapes as for the well-mixed model. However, comparing the colour bars with those in Figures 6.3 and 6.4, we see that the observed homotypic survival difference is much smaller than in the well-mixed case. Moreover, the figures show that this effect is exacerbated for segregated initial conditions. Since $\overline{\Delta^{=}_{A|B}}$ and $\overline{\Delta^{=}_{B|A}}$ quantify the effects of competition, this suggests that spatial segregation acts as a buffer against competition.

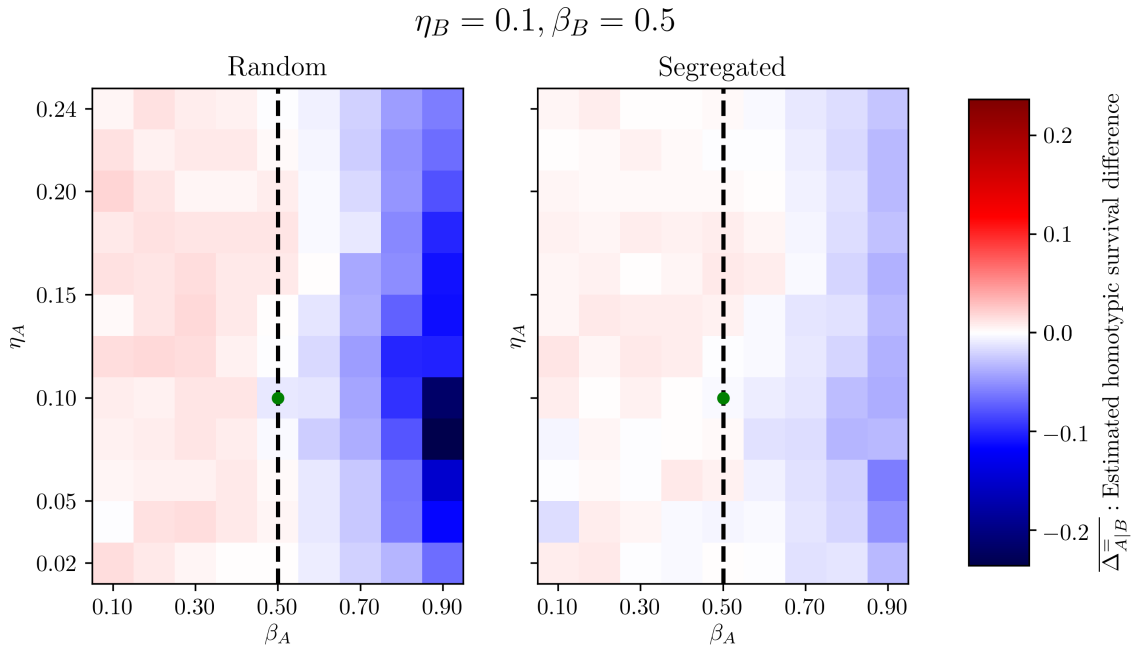

Figure 6.5: Estimated homotypic survival difference of cell type A, $\overline{\Delta^{=}_{A|B}}$, defined in Equation (6.59), for the vertex model. The dashed line and the green dot correspond to the neutral competition curve and the neutral coexistence point, respectively.

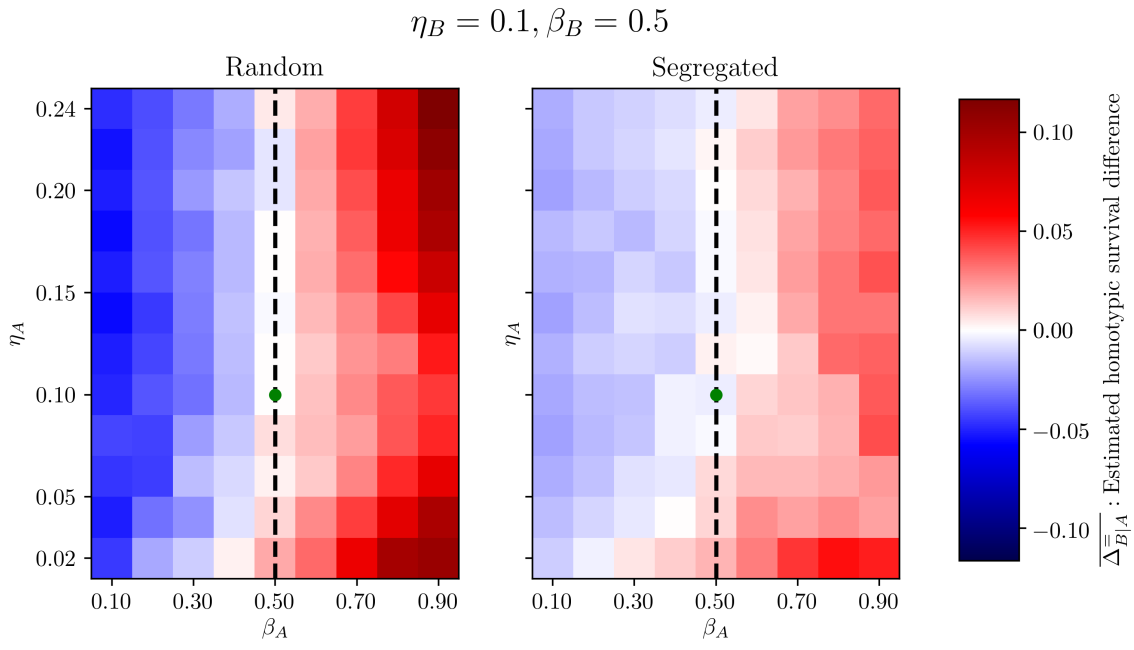

Figure 6.6: Estimated homotypic survival difference of cell type B, $\overline{\Delta^{=}_{B|A}}$, defined in Equation (6.60), for the vertex model. The dashed line and the green dot correspond to the neutral competition curve and the neutral coexistence point, respectively.

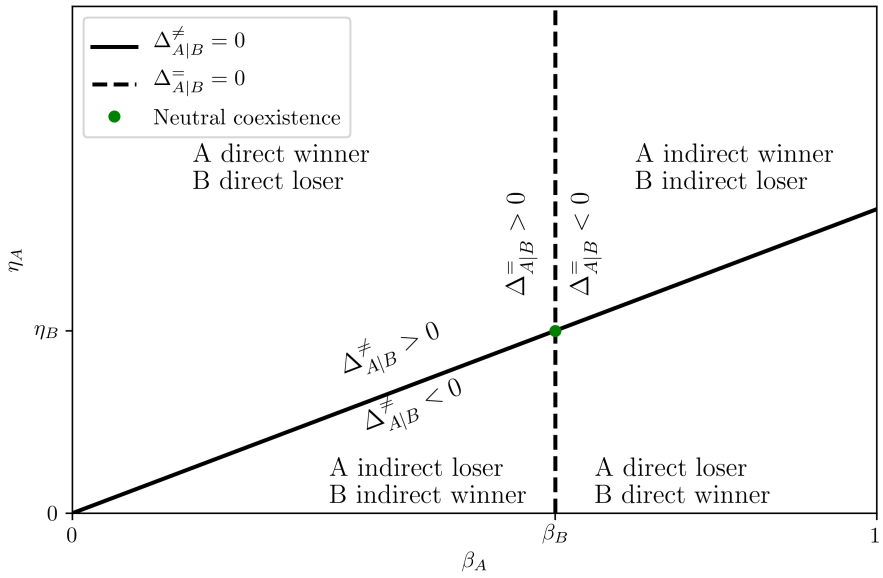

Figure 6.7: Diagram situating the different types of competitive interactions in (β_A, η_A)-space, given fixed values for β_B and η_B. The full and dashed lines correspond to the coexistence and neutral competition curves, respectively.

6.6 Classification of competitive interactions

Now that we have defined both the heterotypic and homotypic survival differences, we are able to refine notions of competition by combining them to tease out different types of competitive interactions. Based on the signs of $\Delta^{\neq}_{A|B}$ and $\Delta^{=}_{A|B}$, there are a total of nine categories, which are tabulated in Table 6.3. In this section, first we examine these categories one by one, and then situate the different types of competitive interactions in the parameter space (β_A, η_A), as plotted in Figure 6.7.

To recap, the heterotypic survival difference $\Delta^{\neq}_{A|B}$, defined in Section 6.4, indicates which cell type in the heterotypic population has a higher or lower survival probability compared to the competing cell type, thus letting us discriminate between winners and losers. On the other hand, the homotypic survival differences $\Delta^{=}_{A|B}$ and $\Delta^{=}_{B|A}$, defined in Section 6.5, indicate whether a cell type does better or worse as a result of the heterotypic interaction, when compared to a homotypic environment.

In general, there are three possible values (negative, zero, and positive) for the signs of of $\Delta^{\neq}_{A|B}$, $\Delta^{=}_{A|B}$, and $\Delta^{=}_{B|A}$, which add up to a total of $3^3 = 27$ different modes of competition. For the exponential cell cycle model specifically, however, we know by Equation (6.55) that $\operatorname{sgn}\left(\Delta^{=}_{A|B}\right)$ is uniquely determined by $\operatorname{sgn}\left(\Delta^{=}_{B|A}\right)$, leaving $3^2 = 9$ categories. We can narrow these categories even further by considering their symmetry with respect to cell

Table 6.3: Classification of competitive interactions.

$\Delta^{=}_{A\|B}$ \ $\Delta^{\neq}_{A\|B}$	$-$	0	$+$
$-$	A direct loser B direct winner	Coexistence	A indirect winner B indirect loser
0	A neutral loser B neutral winner	Neutral coexistence	A neutral winner B neutral loser
$+$	A indirect loser B indirect winner	Coexistence	A direct winner B direct loser

types. In particular, it follows from the definition of the heterotypic survival difference in Equation (6.34) that $\Delta^{\neq}_{A|B} = -\Delta^{\neq}_{B|A}$. Combined with Equation (6.55), which states that $\Delta^{=}_{A|B} = -\Delta^{=}_{B|A}$, we see that the signs of $\Delta^{\neq}_{A|B}$ and $\Delta^{=}_{A|B}$ are antisymmetric with respect to cell types. On the basis of this symmetry, we collapse the nine categories to five unique categories:

Neutral coexistence $\left\{\Delta^{\neq}_{A|B} = 0,\ \Delta^{=}_{A|B} = 0\right\}$**:** This is the degenerate case that corresponds to a homotypic population. The competition is neutral because there is no effect on either cell type's absolute fitness, and the cell types are said to be coexisting because there is no difference in their relative fitness.

Coexistence $\left\{\Delta^{\neq}_{A|B} = 0,\ \Delta^{=}_{A|B} \neq 0\right\}$**:** Despite a change in absolute fitness compared to the homotypic case, there is no survival advantage for either cell type. Hence neither cell type will dominate, but the fitness of the tissue as a whole may be affected.

Neutral competition $\left\{\Delta^{\neq}_{A|B} \neq 0,\ \Delta^{=}_{A|B} = 0\right\}$**:** Because of the nonzero heterotypic survival difference, there exists a difference in relative fitness. Thus, a winner and a loser cell type emerge, with the winner cell type likely becoming the dominant cell type in the population. However, the cells do not experience a difference in absolute fitness compared to the homotypic environment. Thus, cells that are viable in a homotypic environment remain viable when in neutral competition with another cell type, even as they become losers through the competition.

Indirect competition $\left\{\Delta^{\neq}_{A|B} \neq 0,\ \Delta^{=}_{A|B} \neq 0,\ \operatorname{sgn}\left(\Delta^{\neq}_{A|B}\right) \neq \operatorname{sgn}\left(\Delta^{=}_{A|B}\right)\right\}$**:** As in neutral competition, there are winners and losers in the case of indirect competition, and the population will become increasingly populated by winners. However, the sign of the

homotypic survival difference is opposite to the sign of the heterotypic survival difference. This creates the potential for two counter-intuitive scenarios. Firstly, not only are intrinsically viable loser cell types still viable when interacting with the winner cell type, some nonviable loser cell types may *become viable* as a direct result of the competition and are hence paradoxically "rescued" by the winner cell type. Secondly, the winner cell type could be made nonviable by the loser cells. However, as we will discuss in Section 6.7, if the winner cell type was viable prior to the interaction, they are likely to remain viable as the population will tend towards a majority of winner cells over time, and their homotypic fitness will be restored as the population becomes more homogeneous for winner cells.

Direct competition $\left\{\Delta^{\neq}_{A|B} \neq 0,\ \Delta^{=}_{A|B} \neq 0,\ \operatorname{sgn}\left(\Delta^{\neq}_{A|B}\right) = \operatorname{sgn}\left(\Delta^{=}_{A|B}\right)\right\}$**:** Like the other forms of competition, there is a distinction between winner and loser cells, and over time the population tends towards domination by the winner cells. Moreover, the homotypic survival difference has the same sign as the heterotypic survival difference, meaning that the winners fare better within a heterotypic environment than in a homotypic one, and the losers are worse off in a heterotypic environment as compared to a homotypic one. Importantly, direct competition allows for the distinct possibility of previously viable cell types becoming nonviable losers and being completely eliminated from the population as a direct result of the competitive interaction. This is the hallmark of cell competition, where only one from two initially viable cell populations survive the competition.

In order to locate the five types of competitive interactions in parameter space, we fix β_B and η_B, and divide the two-dimensional parameter space (β_A, η_A) by the coexistence curve and the neutral competition curve, defined in Sections 6.4.1 and 6.5.1, respectively.

As Figure 6.7 shows, the coexistence and neutral competition curves translate to straight lines in (β_A, η_A)-space, on which we can find coexistence and neutral competition, respectively. The coexistence curve has a slope η_B/β_B with respect to the β_A-axis, whereas the neutral competition curve is a vertical line that intersects the β_A-axis at $\beta_A = \beta_B$. They intersect at $\beta_A = \beta_B, \eta_A = \eta_B$, which is the only place where $\Delta^{\neq}_{A|B} = \Delta^{=}_{A|B} = 0$. Therefore, we call this the **neutral coexistence point**. Finally, these curves divide the parameter space into four sectors, with the top left and bottom right sectors corresponding to direct competition, and the top right and bottom left sectors corresponding to indirect competition.

6.7 Proliferation regimes

In the previous section, we discussed different types of competitive interactions. However, being a winner is not a sufficient condition for survival. Our definition of winning is based on the heterotypic survival difference $\Delta^{\neq}_{A|B}$, and it only means that the survival probability of the winner cell type is higher than that of the loser cell type. But even winners can be nonviable. In that case, both cell types are headed to extinction, with losers taking the lead. Conversely, losers are not necessarily nonviable and can continue proliferating in the tissue, even if they are less abundant than winners. In this section, we define the asymptotic survival probability and use it to investigate the proliferation regimes of heterotypic populations. We then plot the resulting proliferation regimes in parameter space (β_A, η_A) for particular values of β_B and η_B, and discuss the results. Finally, we validate the theoretical predictions with simulations.

6.7.1 Asymptotic survival probability

At the end of Section 6.3, we assumed that $t_{G,A} = t_{G,B}$, such that the differences in heterotypic survival probability alone determine proliferative success. Hence, regardless of the type of competitive interaction taking place between winners and losers, over time the population is more likely to be dominated by the winner cell type. Therefore, the proportion of winner cells tends towards unity. As a result, we expect the population-weighted average death signal $\langle 1-\beta \rangle(t)$ to approach the homotypic death signal of the winner cell. Assuming that cell type A is the winner, i.e. $\Delta^{\neq}_{A|B} > 0$, this means that

$$\langle 1-\beta \rangle(t) \to 1-\beta_A \quad \text{as} \quad t \to \infty \,. \tag{6.61}$$

We can therefore substitute $1-\beta_A$ for $\langle 1-\beta \rangle$ in the heterotypic survival probability of A-type and B-type cells in Equations (6.31) and (6.32) to obtain the **asymptotic survival probability**:

$$\xi_{A|B}(\beta_A, \eta_A, 1-\beta_A) = 1 - \exp\left(-\frac{\eta_A}{\beta_A(1-\beta_A)}\right), \tag{6.62}$$

$$\xi_{B|A}(\beta_B, \eta_B, 1-\beta_A) = 1 - \exp\left(-\frac{\eta_B}{\beta_B(1-\beta_A)}\right). \tag{6.63}$$

Comparing Equation (6.62) with Equation (6.48), we find that the asymptotic survival probability is the same as the homotypic survival probability:

$$\xi_{A|B}(\beta_A, \eta_A, 1-\beta_A) = \lambda_A \,. \tag{6.64}$$

In other words, the heterotypic survival probability of a winner cell type tends towards its homotypic survival probability as winners become more abundant than losers. We define

$$\xi^{\infty}_{B|A} \equiv 1 - \exp\left(-\frac{\eta_B}{\beta_B(1-\beta_A)}\right), \tag{6.65}$$

which is the heterotypic survival probability of cell type B in an environment dominated by cell type A. Combining Equations (6.63) and (6.65), we write

$$\xi_{B|A}(\beta_B, \eta_B, 1-\beta_A) = \xi^{\infty}_{B|A} \,. \tag{6.66}$$

Similarly, when cell type B is the winner, i.e. $\Delta^{\neq}_{A|B} < 0$, we define

$$\xi^{\infty}_{A|B} \equiv 1 - \exp\left(-\frac{\eta_A}{\beta_A(1-\beta_B)}\right), \tag{6.67}$$

and we can derive that

$$\xi_{A|B}(\beta_A, \eta_A, 1-\beta_B) = \xi^{\infty}_{A|B} \,, \tag{6.68}$$

$$\xi_{B|A}(\beta_B, \eta_B, 1-\beta_B) = \lambda_B \,. \tag{6.69}$$

6.7.2 Classification

In Section 4.2, we learned that a survival probability less than or equal to a half results in guaranteed extinction, whereas a survival probability greater than a half is likely to result in population explosion. Moreover, in Section 5.4, we characterised the proliferation regimes for a homotypic population based on the homotypic survival probability. Here, we use the asymptotic survival probability to characterise the proliferative behaviour of a heterotypic population.

Assuming that cell type A is the winner, i.e. $\Delta^{\neq}_{A|B} > 0$, we can distinguish between the following outcomes:

Case $\left\{\lambda_A \leq \frac{1}{2}\right\}$: If the winner cells are not viable, then the losers are not either, since they have, by definition, a lower survival probability than the winner. Hence, both winners and losers go extinct when winner cells are intrinsically nonviable.

Table 6.4: Classification of proliferation regimes.

Hypersurface	Parameter space			Proliferation regime
			$\xi^{\infty}_{B\|A} > \frac{1}{2}$	A viable winner B viable loser
B loser viability			$\xi^{\infty}_{B\|A} = \frac{1}{2}$	
		$\lambda_A > \frac{1}{2}$	$\xi^{\infty}_{B\|A} < \frac{1}{2}$	A viable winner B nonviable loser
A winner viability		$\lambda_A = \frac{1}{2}$		
	$\Delta^{\neq}_{A\|B} > 0$	$\lambda_A < \frac{1}{2}$		A nonviable winner B nonviable loser
Coexistence	$\Delta^{\neq}_{A\|B} = 0$			
	$\Delta^{\neq}_{A\|B} < 0$	$\lambda_B < \frac{1}{2}$		A nonviable loser B nonviable winner
B winner viability		$\lambda_B = \frac{1}{2}$		
		$\lambda_B > \frac{1}{2}$	$\xi^{\infty}_{A\|B} < \frac{1}{2}$	A nonviable loser B viable winner
A loser viability			$\xi^{\infty}_{A\|B} = \frac{1}{2}$	
			$\xi^{\infty}_{A\|B} > \frac{1}{2}$	A viable loser B viable winner

Case $\left\{\lambda_A > \frac{1}{2}\right\}$**:** In this case, winner cells are intrinsically viable and are likely to multiply and dominate the population. However, whether the loser population goes extinct or not depends on the value of $\xi^{\infty}_{B|A}$.

Subcase $\left\{\xi^{\infty}_{B|A} \leq \frac{1}{2}\right\}$**:** The loser cells are nonviable in the heterotypic population and are entirely eliminated from the tissue.

Subcase $\left\{\xi^{\infty}_{B|A} > \frac{1}{2}\right\}$**:** Despite a lower survival probability than the winners, the losers are still viable. In the developed tissue, we would expect to see mostly winner cells, but a small loser population may still be present.

We thus end up with three proliferation regimes for $\Delta^{\neq}_{A|B} > 0$. Three analogous proliferation regimes exist for $\Delta^{\neq}_{A|B} < 0$, for a total of six proliferation regimes. Visualising four-dimensional $(\beta_A, \eta_A, \beta_B, \eta_B)$-space is tricky, so we only provide a sketch of the parameter space here.

Firstly, the **coexistence hypersurface** $\Delta^{\neq}_{A|B} = 0$ divides the parameter space into two subspaces, $\Delta^{\neq}_{A|B} > 0$ and $\Delta^{\neq}_{A|B} < 0$, where the A and B cell types are winners, respectively. Secondly, within each of these sections there is a further subdivision into two regions where the winner is either viable or nonviable. For $\Delta^{\neq}_{A|B} > 0$ this corresponds to $\lambda_A > 1/2$ and $\lambda_A < 1/2$, respectively. The border is given by the **A winner viability hypersurface** $\lambda_A = 1/2$. Finally, the region in which the winner is viable, i.e. $\lambda_A > 1/2$, is split into two parts, based on whether the loser is viable ($\xi^{\infty}_{B|A} > 1/2$) or nonviable ($\xi^{\infty}_{B|A} < 1/2$). The boundary is given by the **B loser viability hypersurface** $\xi^{\infty}_{B|A} = 1/2$. Since either cell type can be a loser or winner, there are two winner viability hypersurfaces and two loser viability hypersurfaces, for a total of five hypersurfaces that partition the parameter space: $\Delta^{\neq}_{A|B} = 0$, $\lambda_A = 1/2$, $\xi^{\infty}_{B|A} = 1/2$, $\lambda_B = 1/2$, and $\xi^{\infty}_{A|B} = 1/2$. In Table 6.4, we provide a schematic diagram of the parameter space. Horizontal lines correspond to hypersurfaces, and the spaces between lines correspond to distinct proliferation regimes.

One caveat concerning this classification is that we assume that the initial loser and winner populations are large enough so that a cell type with a survival probability larger than a half is virtually guaranteed to survive. As we saw in Chapter 4, at low cell counts it is still possible for a population to go extinct with a survival probability over a half. This assumption is violated, for example, when we examine the dynamics of a mutation appearing in an initially homotypic population, since the mutation appears in a single cell only.

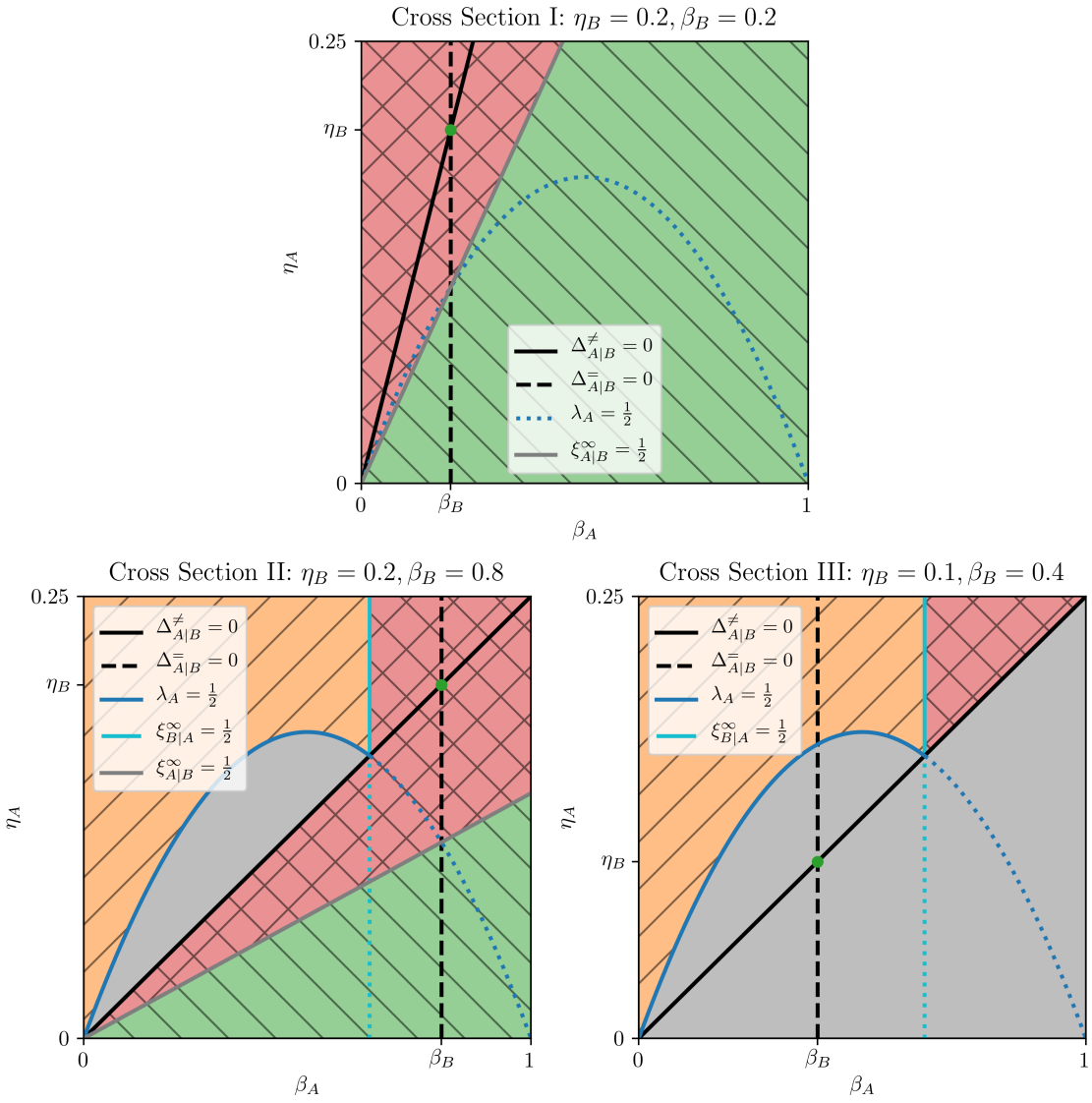

Figure 6.8: Diagrams for Cross Sections I, II, and III, situating the different proliferation regimes. The green dot corresponds to the point (β_B, η_B). : cell types A and B go extinct. : cell type A goes extinct, cell type B survives. : cell type A survives, cell type B goes extinct. : both cell types A and B survive.

6.7.3 Cross sections

Here, we treat β_B and η_B as fixed parameters and consider cross sections in (β_A, η_A)-space to visualise the proliferation regimes. For particular values of β_B and η_B, we have the following curves:

$$\text{Coexistence curve:} \qquad \Delta^{\neq}_{A|B} = 0 \qquad \Leftrightarrow \qquad \eta_A = \frac{\eta_B}{\beta_B}\beta_A\,, \tag{6.70}$$

$$\text{A winner viability curve:} \qquad \lambda_A = \frac{1}{2} \qquad \Leftrightarrow \qquad \eta_A = \ln(2)\beta_A(1-\beta_A)\,, \tag{6.71}$$

$$\text{B loser viability curve:} \qquad \xi^{\infty}_{B|A} = \frac{1}{2} \qquad \Leftrightarrow \qquad \beta_A = 1 - \frac{\eta_B}{\ln(2)\beta_B}\,, \tag{6.72}$$

$$\text{A loser viability curve:} \qquad \xi^{\infty}_{A|B} = \frac{1}{2} \qquad \Leftrightarrow \qquad \eta_A = \ln(2)(1-\beta_B)\beta_A\,. \tag{6.73}$$

We note that the A winner viability curve is analogous to the homotypic viability curve for the exponential cell cycle model, defined in Section 5.4. The B winner viability hypersurface does not map on to a curve in (β_A, η_A)-space because it only depends on β_B and η_B. Therefore, we consider the cases $\lambda_B < 1/2$ and $\lambda_B > 1/2$ in separate cross sections.

In addition, if $\eta_B/\beta_B > \ln(2)$, then Equation (6.72) does not have a solution for positive β_A. Hence, the B loser viability curve does not appear in cross sections for which this condition is satisfied. Since the inclusion of the B loser viability curve qualitatively changes the proliferation regimes found within a cross section, we treat this case in a separate cross section as well. It can be easily verified that $\eta_B/\beta_B > \ln(2)$ implies $\lambda_B > 1/2$, so we pick three cross sections in total:

Cross Section I $\{\beta_B = 0.2, \eta_B = 0.2\}$**:** This satisfies $\eta_B/\beta_B > \ln(2)$.

Cross Section II $\{\beta_B = 0.8, \eta_B = 0.2\}$**:** This satisfies $\eta_B/\beta_B < \ln(2)$ and $\lambda_B > 1/2$.

Cross Section III $\{\beta_B = 0.4, \eta_B = 0.1\}$**:** This satisfies $\eta_B/\beta_B < \ln(2)$ and $\lambda_B < 1/2$.

In Figure 6.8, we show a diagram of the proliferation regimes for each of these cross sections.

Cross Section I We observe three distinct regimes. For $\Delta^{\neq}_{A|B} > 0$, both cell types are proliferating, with A-type cells as the winner. On the other side of the coexistence curve, B-type cells are winning. Depending on the values of β_A and η_A, A-type cells can either proliferate alongside B-type cells or be eliminated from the population entirely. Interestingly, there are no regions in which B-type cells go extinct. In other words, faced with any competitor, the cell type B is never fully eliminated from the population.

Cross Section II We identify five distinct regimes. For $\Delta^{\neq}_{A|B} < 0$, the same two regimes appear as in Cross Section I. The wedge-shaped region where both winners and losers proliferate is surprising because a significant portion overlaps with the area under the homotypic viability curve. What this means is that the A-type cells in this region would go extinct in a homotypic environment, i.e. they are intrinsically nonviable, but they are "rescued" by the competitive interaction with B-type cells. This is also present in Cross Section I, but it is much more pronounced in Cross Section II, so we discuss it here. We note that this wedge is entirely contained within the indirect competition sector, showing that indirect competition is required for rescuing intrinsically nonviable loser cells.

We see three regimes on the other side of the coexistence curve. For $\lambda_A < 1/2$ the winner A-type cells are nonviable, rendering both cell types nonviable. This is despite the fact that the B-type cells in this regime would be viable in a homotypic environment. Hence, this is a case of pathological competition, where a winner cell type renders a tissue nonviable that was viable prior to the appearance of the winner. For $\lambda_A > 1/2$ the winner A-type cells are viable. In this space, whether the B-type cells survive depends on β_A. If it is less than $1 - \eta_B/(\ln(2)\beta_B)$, then the resulting death signal is too high for B-type cells to survive. However, on the other side of this critical value, the death signal is sufficiently low for B-type cells to survive. We can also see that the B loser viability curve is located at the intersection of the A winner viability curve and the coexistence curve. Indeed, the condition $\eta_B/\beta_B < \ln(2)$ is equivalent to the condition that the A winner viability curve intersects the coexistence curve for positive β_A.

Cross Section III In Cross Section III, the point (β_B, η_B) satisfies $\lambda_B < 1/2$. Hence, under the coexistence curve, where B-type cells are the winners, all cell types go extinct. Above the coexistence curve, we find the same regions as in Cross Section II. Because cell type B would be nonviable in a homotypic environment, we observe that the top right triangular area is another case of intrinsically nonviable loser cells being rescued by competition, and is analogous to the wedge-shaped areas in Cross Sections I and II. Again, this area is fully contained within the indirect competition sector.

Finally, we observe in every cross section that along the neutral competition curve A-type cells proliferate for the same parameter values that render them viable in a homotypic environment. This is because the cells experience the same death signal in neutral competition as they would in the homotypic case.

Table 6.5: Model and simulation parameter values used to estimate the proliferation regimes for the well-mixed model.

Parameter	Cell type A	Cell type B
t_G		100
c		1
η	$0.01, 0.02, \ldots, 0.25$	$\begin{bmatrix} 0.2 \\ 0.2 \end{bmatrix}, \begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix}, \begin{bmatrix} 0.1 \\ 0.4 \end{bmatrix}$
β	$0.05, 0.10, \ldots, 0.95$	
Initial cell count	50	50
Simulation end time		10 000
Minimum cell count		10
Maximum cell count		1 000
N_{sim}		50

6.7.4 Computational validation

Figure 6.8 shows the predicted proliferation regimes for three representative cross sections in (β_A, η_A)-space. Here, we validate these results by estimating the survival frequency in each of the cross sections using a Monte Carlo method and comparing them to predictions.

Parameter choice We followed the same approach as in Section 6.4.2 (Parameter choice, page 138), except that we chose parameters for cell type B corresponding to the cross sections defined in Section 6.7.3. The parameters are listed in Tables 6.5 and 6.6 for the well-mixed and vertex models, respectively.

For the well-mixed model, we ran 50 simulations for every unique parameter set, so the total number of simulations is $25 \times 19 \times 3 \times 50 = 71\,250$. For the vertex model, we additionally combined every parameter set with two different initial conditions, and ran 20 simulations per unique combination of parameter set and type of initial conditions. Hence, the total number of simulations is $12 \times 9 \times 3 \times 2 \times 20 = 12\,960$.

Initial conditions See Initial conditions in Section 6.4.2 (page 139).

Termination conditions We set a minimum cell count of 10, a maximum cell count of 1 000, and a simulation end time of 10 000. We do not terminate the simulation when the population becomes homotypic because we are interested in the long-term behaviour of the model.

Table 6.6: Model and simulation parameter values used to estimate the proliferation regimes for the vertex model.

Parameter	Cell type A	Cell type B
t_G		100
c		1
η	$0.02, 0.04, \ldots, 0.24$	$\begin{bmatrix} 0.2 \\ 0.2 \end{bmatrix}, \begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix}, \begin{bmatrix} 0.1 \\ 0.4 \end{bmatrix}$
β	$0.1, 0.2, \ldots, 0.9$	
Initial cell count	50	50
Pattern		random, segregated
Simulation end time		10 000
Minimum cell count		10
Maximum cell count		1 000
N_{sim}		20

Data processing We computed the estimated heterotypic survival frequencies $\overline{\xi_{A|B}}$ and $\overline{\xi_{B|A}}$ using Equations (2.13) and (2.14), respectively[3].

Data visualisation We visualised $\overline{\xi_{A|B}}$ and $\overline{\xi_{B|A}}$ for each cross section using heat maps. We are interested in the regions where the heterotypic survival frequencies are below or above $1/2$, so we set the colour map range to $[0, 1]$, with central value $1/2$. Hence, red and blue regions correspond to survival frequencies above and below $1/2$, respectively.

The results are given in Figure 6.9 for the well-mixed model. The results for random and segregated initial conditions in the vertex model are given in Figures 6.10 and 6.11, respectively.

Computational materials and methods See Computational materials and methods in Section 2.3.1 (page 36) for the vertex model and in Section 4.4 (page 77) for the well-mixed model.

6.7.5 Results

Well-mixed model The left and right columns in Figure 6.9 show the survival frequency for cell types A and B, respectively. Comparing this with Figure 6.8, we see an excellent agreement between simulations and predictions. This includes the wedge-shaped region

[3] See Section 6.4.2 (Data processing, page 140).

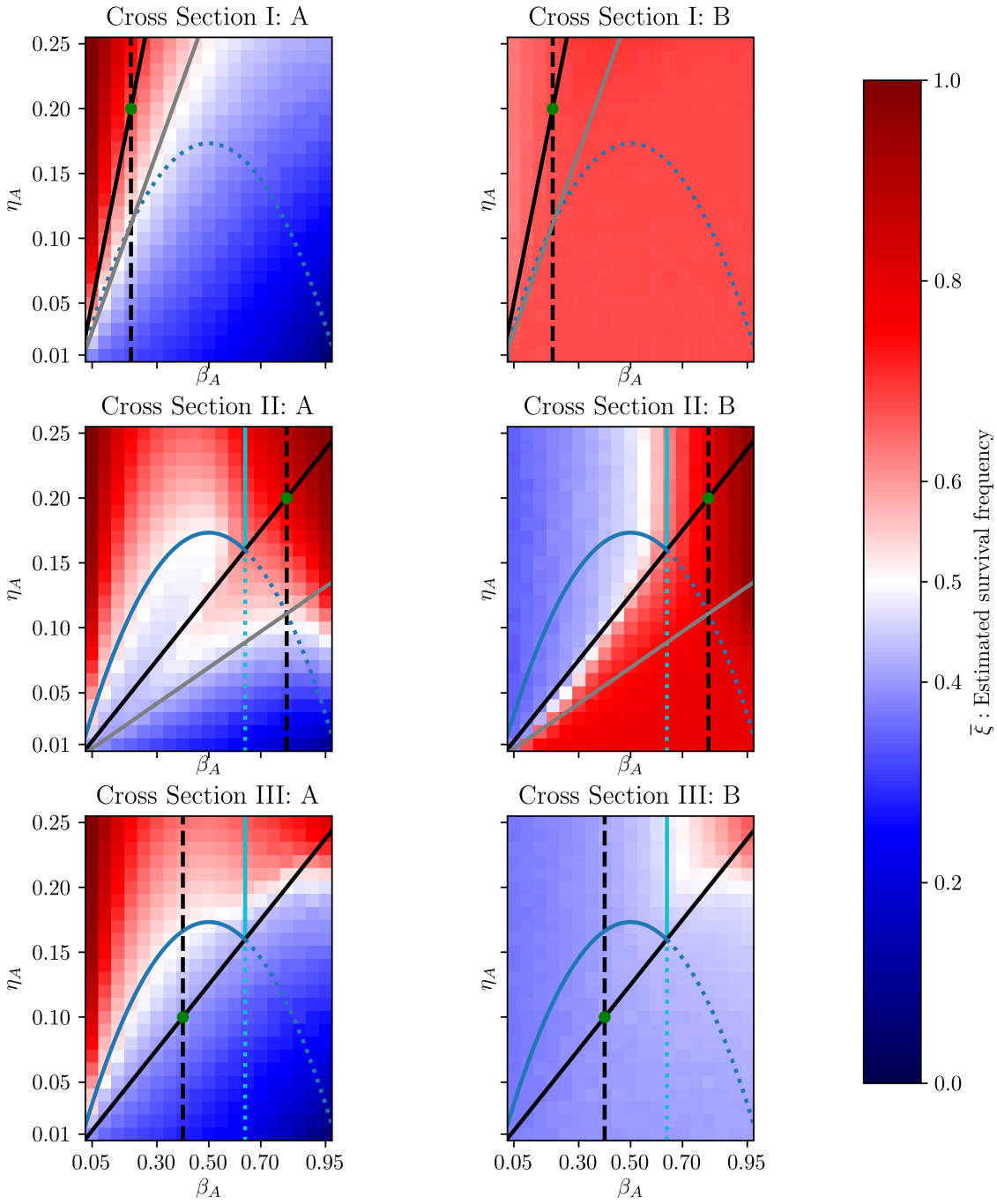

Figure 6.9: Estimated survival frequency of cell types A and B using the well-mixed model. The left column displays the estimated survival frequency of cell type A, $\overline{\xi_{A|B}}$, defined in Equation (6.41). The right column displays the estimated survival frequency of cell type B, $\overline{\xi_{B|A}}$, defined in Equation (6.42). All curves are the same as in Figure 6.8.

in Cross Section II between the coexistence curve and the A loser viability curve, where intrinsically nonviable A-type loser cells are rescued by competition, as well as the region between the A winner viability curve and the coexistence curve, where intrinsically viable B-type cells are rendered nonviable by nonviable A-type winner cells. We see some slight deviations for high β_A values in Cross Sections II and III, with red regions crossing over the A loser viability curve (Cross Section II) and the coexistence curve (Cross Section III). However, overall the observed proliferation regimes match predictions remarkably well.

Vertex model In Figure 6.10, where we plotted the results for random initial conditions, we can roughly make out the same proliferation regimes as in the well-mixed case. However, the observed contours are shifted with respect to predictions. Moreover, some distinctions are blurred, such as the area under the A winner viability curve in Cross Section II for cell type A. Here, we expected a blue region directly under the A winner viability curve, followed by a red region in the wedge-shaped area under the coexistence curve. However, we see a continuation of the red region above the A winner viability curve in the former, and a transition to the blue region in the bottom right corner in the latter. Particularly for high β_A, we see red regions for cell type A in all cross sections that extend below their predicted contours.

In Figure 6.11, where we plotted the results for segregated initial conditions, we see large deviations from the predicted proliferation regimes. In particular, there is no blue region in Cross Section II and no red region in Cross Section III for cell type B. Moreover, the boundary between red and blue for cell type A roughly follows the A winner viability curve. When we compare the plots for cell type A with our estimation of the homotypic proliferation regimes in Figure 5.8, we see that A-type cells behave almost as if they were in a homotypic environment. This is also the case for cell type B, since the homotypic survival probability is greater than $1/2$ for Cross Sections I and II, and less than $1/2$ for Cross Section III, and, accordingly, Cross Sections I and II are coloured red and Cross Section III is coloured blue, regardless of the parameters of cell type A. This shows that spatial segregation results in cells largely behaving as if they were in a homotypic population.

6.8 Classification of competition regimes

So far, we have characterised the proliferation regimes for homotypic populations (Section 5.4), classified different types of competitive interactions (Section 6.6), and derived the proliferation regimes for heterotypic populations (Section 6.7). In this section, we combine these classifications to identify biologically relevant competition regimes for the G2 death

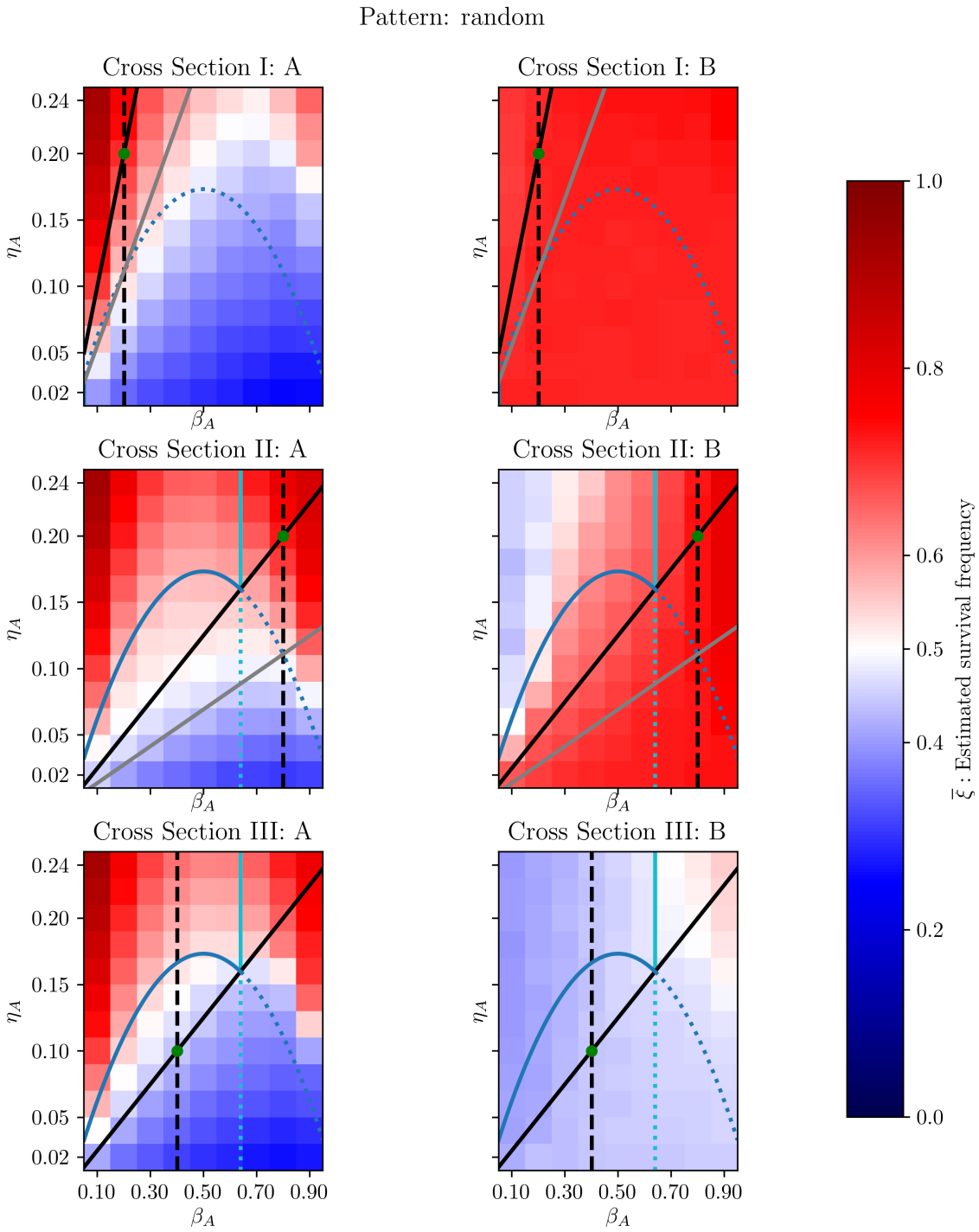

Figure 6.10: Estimated survival frequency of cell types A and B using the vertex model and the random initial pattern. The left column displays the estimated survival frequency of cell type A, $\overline{\xi_{A|B}}$, defined in Equation (6.41). The right column displays the estimated survival frequency of cell type B, $\overline{\xi_{B|A}}$, defined in Equation (6.42). All curves are the same as in Figure 6.8.

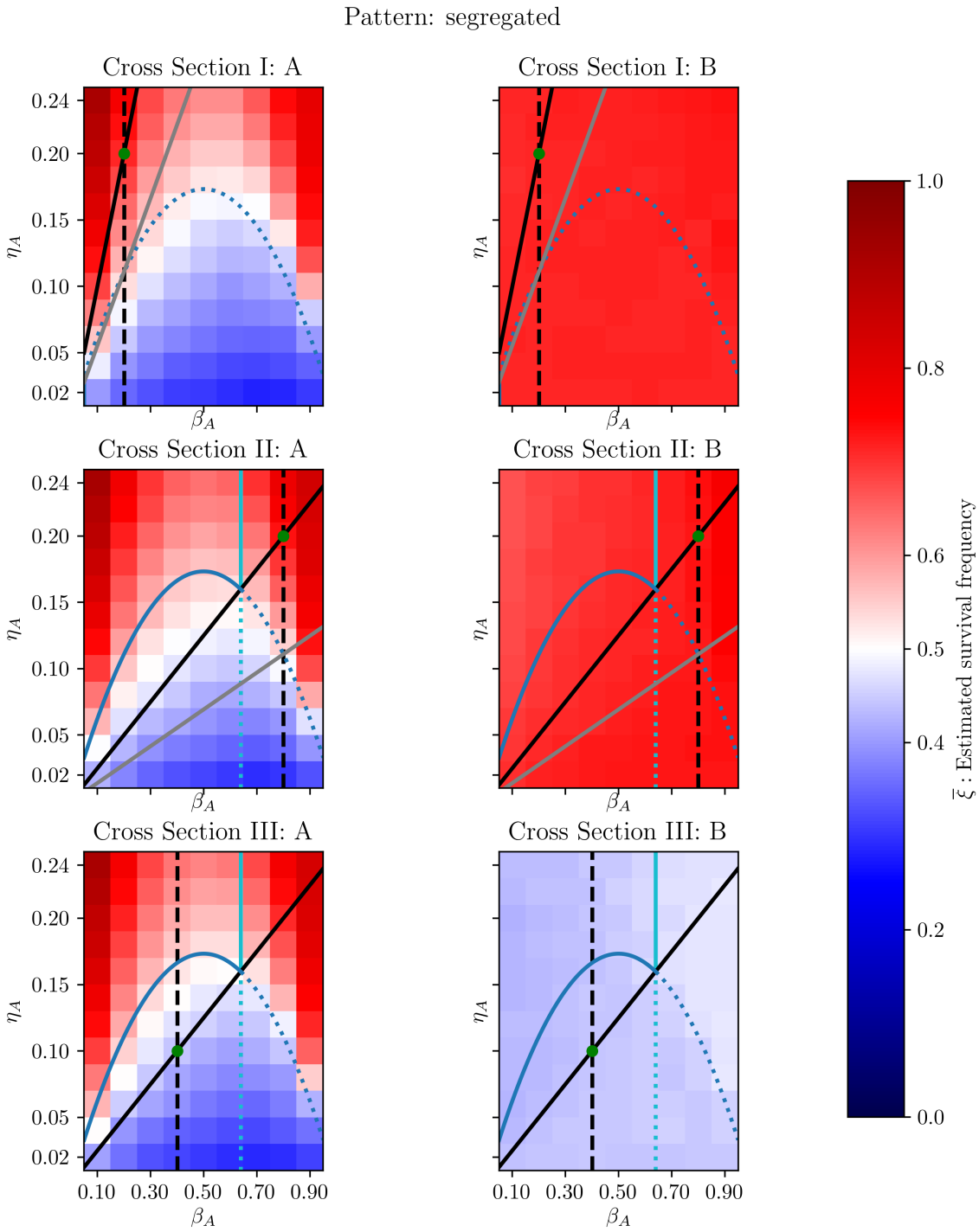

Figure 6.11: Estimated survival frequency of cell types A and B using the vertex model and the segregated initial pattern. The left column displays the estimated survival frequency of cell type A, $\overline{\xi_{A|B}}$, defined in Equation (6.41). The right column displays the estimated survival frequency of cell type B, $\overline{\xi_{B|A}}$, defined in Equation (6.42). All curves are the same as in Figure 6.8.

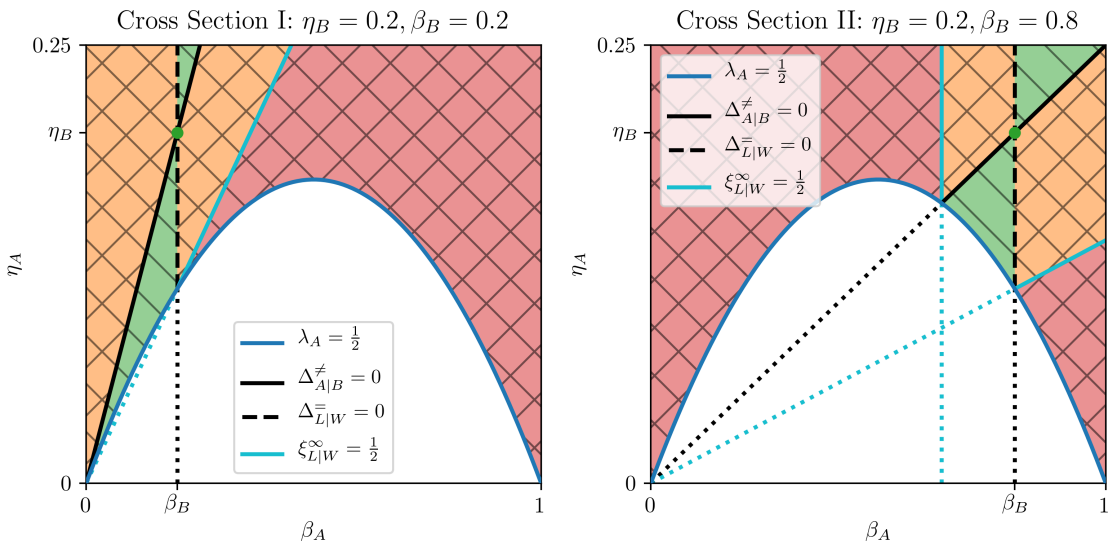

Figure 6.12: Diagrams of competition regimes for Cross Sections I and II. The green dot corresponds to the point (β_B, η_B). The labels W and L are used to refer to the winner and loser cell types, respectively, i.e. W = A, L = B for $\Delta^{\neq}_{A|B} > 0$ and W = B, L = A for $\Delta^{\neq}_{A|B} < 0$. ⟍: intrinsic viability. ⨉: cell competition. ■: complete cell competition. ■: incomplete cell competition. ■: indirect cell competition. See also Table 6.7 for the legend.

Table 6.7: Classification of competition regimes. The competition regime (bolded) can be subdivided in two ways: loser elimination and loser survival regimes (top section), or cell competition, neutral competition, and indirect competition regimes (bottom section). The underlined conditions are implied by the other conditions on the same row. The legend column maps the regimes to areas and curves plotted in Figure 6.12.

Regime	λ_W, λ_L	$\Delta^{\neq}_{W\|L}$	$\xi^{\infty}_{L\|W}$	$\Delta^{=}_{L\|W}$	Legend
Intrinsic viability	$> 1/2$	-	-	-	⟍⟍⟍
Coexistence	$> 1/2$	$= 0$	-	-	▬
Competition	$> 1/2$	> 0	-	-	-
Loser elimination	$> 1/2$	> 0	$\leq 1/2$	-	-
Loser survival	$> 1/2$	> 0	$> 1/2$	-	-
Cell competition	$> 1/2$	> 0	-	< 0	⨉⨉⨉
Complete cell competition	$> 1/2$	> 0	$< 1/2$	$\underline{< 0}$	■
Critical cell competition	$> 1/2$	> 0	$= 1/2$	$\underline{< 0}$	■
Incomplete cell competition	$> 1/2$	> 0	$> 1/2$	< 0	■
Neutral competition	$> 1/2$	> 0	$\underline{> 1/2}$	$= 0$	▬ ▬
Indirect competition	$> 1/2$	> 0	$\underline{> 1/2}$	> 0	■

signal model.

The first condition for cell competition is that both competing cell types are intrinsically viable, i.e. $\lambda_A, \lambda_B > 1/2$. In order to satisfy $\lambda_A > 1/2$, we only consider the parameter space above the homotypic viability curve, as shown in Figure 6.12. To satisfy the viability of cell type B, we only consider cross sections that satisfy $\lambda_B > 1/2$. In particular, this means that Cross Section III, defined in Section 6.7.3, is excluded, so we plot only Cross Sections I and II in Figure 6.12. We can verify the condition $\lambda_B > 1/2$ visually by checking whether the neutral coexistence point (where $\beta_A = \beta_B$ and $\eta_A = \eta_B$) is above the homotypic viability curve. We define this regime as the **intrinsic viability regime**:

$$\lambda_A, \lambda_B > \frac{1}{2}\,. \tag{6.74}$$

The second condition for cell competition is that there are winners and losers. In Section 6.4, we defined this as a nonzero heterotypic survival difference, i.e. $\Delta^{\neq}_{A|B} \neq 0$. This distinction splits the intrinsic viability regime into the **coexistence regime**:

$$\lambda_A, \lambda_B > \frac{1}{2} \quad \wedge \quad \Delta^{\neq}_{A|B} = 0\,, \tag{6.75}$$

and the **competition regime**:

$$\lambda_A, \lambda_B > \frac{1}{2} \quad \wedge \quad \Delta^{\neq}_{A|B} \neq 0\,. \tag{6.76}$$

The competition regime can further be subdivided according to which cell type is the winner. The parameter space is symmetric with respect to swapping cell type labels, so we could choose cell type A as the winner, $\Delta^{\neq}_{A|B} > 0$, or cell type B as the winner, $\Delta^{\neq}_{A|B} < 0$, to continue the discussion and we would obtain the same parameter regimes in either case. What matters most here is that there *are* winner and loser cell types, rather than *which* cell type is the winner or loser. To emphasise this fact in the rest of this section, we relabel the winner cell type as W and the loser cell type as L. Concretely, this means that W = A, L = B for $\Delta^{\neq}_{A|B} > 0$ and W = B, L = A for $\Delta^{\neq}_{A|B} < 0$, such that $\Delta^{\neq}_{W|L} > 0$ is true by construction.

As we saw in Section 6.7, when the winner cell type is intrinsically viable, there are two possible outcomes for the loser cell type. The losers are either eliminated if $\xi^{\infty}_{L|W} \leq 1/2$ or they survive if $\xi^{\infty}_{L|W} > 1/2$. Hence, we define the **loser elimination regime** as

$$\lambda_W, \lambda_L > \frac{1}{2} \quad \wedge \quad \Delta^{\neq}_{W|L} > 0 \quad \wedge \quad \xi^{\infty}_{L|W} \leq \frac{1}{2}\,, \tag{6.77}$$

and the **loser survival regime** as

$$\lambda_W, \lambda_L > \frac{1}{2} \quad \wedge \quad \Delta^{\neq}_{W|L} > 0 \quad \wedge \quad \xi^{\infty}_{L|W} > \frac{1}{2}\,. \tag{6.78}$$

The third condition for cell competition is that losers are eliminated. Hence, we could simply equate the loser elimination regime with cell competition and stop there. However, by also considering the type of competitive interaction, we can further refine our notion of cell competition. In particular, we see in Figure 6.12 that the neutral competition curve, defined by $\Delta^{=}_{L|W} = 0$, runs through the loser survival regime. Hence, we define the **neutral competition regime** as

$$\lambda_W, \lambda_L > \frac{1}{2} \quad \wedge \quad \Delta^{\neq}_{W|L} > 0 \quad \wedge \quad \Delta^{=}_{L|W} = 0\,. \tag{6.79}$$

In this regime, both cell types experience neither harm nor benefit from the presence of the other cell type. The neutral competition curve separates the loser survival regime into two subregimes where $\Delta^{=}_{L|W} < 0$ and $\Delta^{=}_{L|W} > 0$, respectively. In the case of $\Delta^{=}_{L|W} < 0$, the fitness of losers is reduced by the winners, but not enough to cause loser elimination. Therefore, we define this as the **incomplete cell competition regime**[4]:

$$\lambda_W, \lambda_L > \frac{1}{2} \quad \wedge \quad \Delta^{\neq}_{W|L} > 0 \quad \wedge \quad \xi^{\infty}_{L|W} > \frac{1}{2} \quad \wedge \quad \Delta^{=}_{L|W} < 0\,. \tag{6.80}$$

This suggests a partition of the loser elimination regime into the **complete cell competition regime**:

$$\lambda_W, \lambda_L > \frac{1}{2} \quad \wedge \quad \Delta^{\neq}_{W|L} > 0 \quad \wedge \quad \xi^{\infty}_{L|W} < \frac{1}{2}\,, \tag{6.81}$$

where loser elimination is complete, and the **critical cell competition regime**:

$$\lambda_W, \lambda_L > \frac{1}{2} \quad \wedge \quad \Delta^{\neq}_{W|L} > 0 \quad \wedge \quad \xi^{\infty}_{L|W} = \frac{1}{2}\,, \tag{6.82}$$

which is the threshold regime between complete and incomplete cell competition. This regime maps on to the loser viability curve defined in Section 6.7. The shared property of complete, critical, and incomplete cell competition is that the winners have a negative impact on the losers. We group these regimes under the **cell competition regime**:

$$\lambda_W, \lambda_L > \frac{1}{2} \quad \wedge \quad \Delta^{\neq}_{W|L} > 0 \quad \wedge \quad \Delta^{=}_{L|W} < 0\,. \tag{6.83}$$

[4] We colour the names of competition regimes with the same colours used to draw them in Figure 6.12.

Lastly, on the other side of the neutral competition curve we have $\Delta^{=}_{L|W} > 0$. This means that loser cells have a higher fitness compared to the homotypic case. The competitive outcome is therefore not caused by upregulating loser cell death, but rather by winners being less sensitive to death signals. This corresponds with an indirect competitive interaction, so we define this as the **indirect competition regime**:

$$\lambda_W, \lambda_L > \frac{1}{2} \quad \wedge \quad \Delta^{\neq}_{W|L} > 0 \quad \wedge \quad \Delta^{=}_{L|W} > 0\,. \tag{6.84}$$

Figure 6.12 shows the competition regimes for Cross Sections I and II, and Table 6.7 summarises the competition regimes.

The competition regimes also let us specify different types of winners and losers. In particular, we define *complete winners*, *critical winners*, *incomplete winners*, *neutral winners*, and *indirect winners* as the winner cell types in the respective competition regimes, and *direct winners* as the winner cell types in the cell competition regime. Furthermore, we define analogous terms for the loser cell types. We note that only critical and complete losers are eliminated by the competitive interaction, while the other types of losers remain viable.

6.9 Discussion

In this chapter, we applied the G2 death signal model to heterotypic populations with the constraint that the cell types can only vary in their death clock parameters. Considering only the exponential cell cycle model, we extended the theoretical framework of the previous chapter. Concrete predictions include the signs of the heterotypic and homotypic survival differences, which we used to classify different types of competitive interactions, and the asymptotic survival probability, which enabled us to characterise proliferation regimes. These predictions were validated using simulations, with great success for the well-mixed model and mixed results for the vertex model. Finally, we used the developed framework to identify and classify the biologically relevant competition regimes that emerge from the heterotypic G2 death signal model.

The main achievement of this chapter is that we were able to reproduce the features of cell competition without assigning *a priori* winner/loser identities. To interpret model predictions we rigorously defined criteria for what it means to be a winner or a loser (heterotypic survival difference), for distinguishing direct from indirect competition (homotypic survival difference), and for distinguishing loser elimination from loser survival (asymptotic survival probability). Not only did we succeed in reproducing the defining characteristics of cell

competition by varying model parameters, but also we discovered a range of competitive outcomes in the process. This enabled us to refine our notion of cell competition into complete, critical, and incomplete cell competition.

Winner/loser status is determined by the relative tolerance to death signals Assuming that the total cell cycle time is the same for both competing cell types, i.e. $t_{G,A} = t_{G,B}$, relative proliferative success is determined solely by the heterotypic survival difference, $\Delta^{\neq}_{A|B}$. We say that the cell type with the higher survival probability is the winner cell type, because those cells become more abundant than the other cell type, which we call the loser cell type.

In Section 6.4, we showed that i) the heterotypic survival difference is invariant with respect to the system state and only depends on model parameters, and ii) winner/loser status is predicted by differences in tolerance to the death signal. In particular, the cell type that has a higher tolerance to death signals, i.e. a higher η/β, becomes the winner.

This result is based on the assumption that cell cycle phases are ergodic, i.e. the spatial distribution of cell cycle phases mirrors the temporal distribution of cell cycle phases, and that the system is well-mixed, i.e. the local distribution of cell cycle phases mirrors the global distribution of cell cycle phases. Together, these assumptions mean that both cell types are experiencing the same death signal, such that the heterotypic survival difference is determined by sensitivity to the common death signal.

Loser elimination is determined by the relative emission rate of death signals The winner cell type, by definition, becomes more abundant. However, this is not sufficient for cell competition, which requires that losers are eliminated from the population. Consider, for instance, neutral competition, where two cell types have the same values for β and vary only in their η parameters. In Section 6.8, we showed that neutral competition implies loser survival, i.e. $\xi^{\infty}_{L|W} > 1/2$, as indicated in Table 6.7. This is because neutral winners and losers experience the same level of death signal as in their respective homotypic environments. Since the losers are assumed to be intrinsically viable, they continue to be viable when entering into a neutral competition environment.

As winners outnumber losers over time, the death signal experienced throughout the population eventually becomes dominated by the homotypic death signal of the winner cell type. In Section 6.8, we observed that loser elimination, $\xi^{\infty}_{L|W} < 1/2$, implies direct competition, i.e. $\Delta^{=}_{L|W} < 0$. Moreover, in Section 6.5, we showed that direct competition is based on differences in the emission rate of the death signal, $1 - \beta$. Therefore, the winner must emit death signals at a higher rate than the loser in order to eliminate the loser.

Therefore, the model predicts that two conditions must be satisfied in order for cell

competition between two homotypically viable cell types to occur. Firstly, one cell type must have a lower sensitivity to the death signal than the other cell type to become the winner. Secondly, the winners must emit death signals at a higher rate than the losers in order to decrease loser viability. When these two conditions are satisfied, we say that cell competition is taking place. However, these conditions are necessary but not sufficient for loser elimination, the defining feature of cell competition. The final condition is that the rate of emission of the death signal by the winner is sufficiently high to render the losers nonviable. In this case, we say that complete cell competition has occurred. This regime in particular is designated in the literature as cell competition.

Tolerance and emission Having identified the tolerance to and emission rate of death signals as the fundamental cell properties driving cell competition, here we present an alternative formulation of the model parameters that more directly expresses these properties. In addition, we reinterpret the competition regimes, defined in Section 6.8, in terms of these transformed parameters.

We define the **tolerance** of cell type X as

$$\tilde{\eta}_X \equiv \frac{\eta_X}{\ln(2)\beta_X} = \frac{T_{\dagger,X}}{\ln(2)c_X t_{\mathrm{G1},X}}\,, \tag{6.85}$$

where we use the factor ln(2) to simplify the expressions below. In Equation (6.85), we expressed the tolerance in terms of the dimensional parameters $T_{\dagger,X}$, c_X, and $t_{\mathrm{G1},X}$ by substituting Equations (6.3) and (6.29). As we saw in Section 4.3, the ratio $T_\dagger/c$ is the death time for a cell experiencing a constant death signal c (see Equation (4.8)). Hence, $T_{\dagger,X}/c_X$ can be interpreted as the death time of a cell whose neighbours are all in G2 phase, such that the proportion of neighbours in G2 phase is unity, $g(t) = 1$, and the G2 death signal is equal to $f_X(t) = c_X g(t) = c_X$.

Therefore, the tolerance, $\tilde{\eta}_X$, can be interpreted as the ratio of this death time, $T_{\dagger,X}/c_X$, over the average G1 duration, $t_{\mathrm{G1},X}$. For a fixed $t_{\mathrm{G1},X}$, increasing the death time makes it less probable for a cell to hit the death time in G1 phase, thus increasing its tolerance to death signals. For a fixed $T_{\dagger,X}/c_X$, increasing $t_{\mathrm{G1},X}$ extends the average duration of G1 phase, making it more likely for a cell to reach the death time while in G1 phase, and hence decreasing the cell's tolerance to death signals.

We define the **emission** of cell type X as

$$d_X \equiv 1 - \beta_X = \frac{t_{\mathrm{G2},X}}{t_{\mathrm{G},X}}\,, \tag{6.86}$$

where we used Equation (6.3) to express the emission in terms of $t_{G2,X}$ and $t_{G,X}$. The emission is thus interpreted as the average fraction of time spent in G2 phase, which is the same as the ergodic approximation to $g(t)$ for homotypic populations, as expressed in Equation (5.7). This expresses that the more time that cells spend in G2 phase on average, the more death signals that they emit.

The transformed parameters $\tilde{\eta}_X$ and d_X allow us to characterise the competition regimes in terms of model parameters using simple and intuitive expressions. We formulate the **intrinsic viability** condition, $1/2 < \lambda_X$, using $\tilde{\eta}_X$ and d_X by substituting Equation (5.14) and rearranging to give

$$\frac{1}{2} < \lambda_X \quad \Leftrightarrow \quad 1 - \beta_X < \frac{\eta_X}{\ln(2)\beta_X} \quad \Leftrightarrow \quad d_X < \tilde{\eta}_X \,, \tag{6.87}$$

where we substituted the definitions of tolerance and emission in the last step. Note that the factor ln(2) is absorbed by $\tilde{\eta}_X$. Equation (6.87) expresses that cells must have a higher tolerance than emission in order to be intrinsically viable.

For heterotypic populations with cell types A and B, winners and losers emerge if $\Delta^{\neq}_{A|B} \neq 0$. Using the convention of Section 6.8 to denote the winner and loser cell types using W and L, respectively, we express the condition for **competition**, $0 < \Delta^{\neq}_{W|L}$, by substituting Equation (6.37) to give

$$0 < \Delta^{\neq}_{W|L} \quad \Leftrightarrow \quad 0 < \frac{\eta_W}{\beta_W} - \frac{\eta_L}{\beta_L} \quad \Leftrightarrow \quad \tilde{\eta}_L < \tilde{\eta}_W \,. \tag{6.88}$$

This expresses that winners have a higher tolerance to death signals than losers. Similarly, we express the condition for **direct competition**, $\Delta^{=}_{L|W} < 0$, by substituting Equation (6.57) to give

$$\Delta^{=}_{L|W} < 0 \quad \Leftrightarrow \quad (1 - \beta_L) - (1 - \beta_W) < 0 \quad \Leftrightarrow \quad d_L < d_W \,. \tag{6.89}$$

This expresses that winners must have a higher emission than losers for the competition to be considered a form of direct competition. We can also express the **loser elimination** condition, $\xi^{\infty}_{L|W} < 1/2$, by substituting Equation (6.65) and rearranging to give

$$\xi^{\infty}_{L|W} < \frac{1}{2} \quad \Leftrightarrow \quad \frac{\eta_L}{\ln(2)\beta_L} < 1 - \beta_W \quad \Leftrightarrow \quad \tilde{\eta}_L < d_W \,. \tag{6.90}$$

Equation (6.90) expresses that loser elimination occurs when the winner emission is higher than the loser tolerance.

The defining features of cell competition are that i) winners and losers are intrinsically viable and ii) losers are eliminated from the population. We can express the first property using

Equation (6.87) as $d_W < \tilde{\eta}_W$ and $d_L < \tilde{\eta}_L$, and the second property using Equation (6.90) as $\tilde{\eta}_L < d_W$. Combining these expressions, we summarise the **hallmarks of cell competition** in a single statement:

$$d_L < \tilde{\eta}_L < d_W < \tilde{\eta}_W \,, \tag{6.91}$$

which can be read as

$$\text{loser emission} < \text{loser tolerance} < \text{winner emission} < \text{winner tolerance} \,. \tag{6.92}$$

This is the regime that we classified in Section 6.8 as **complete cell competition**. Equation (6.91) expresses the intrinsic viability of winners and losers, as well as loser elimination. In addition, it also shows the properties implied by complete cell competition, such as direct competition, $d_L < d_W$, and that winners have a higher tolerance than losers $\tilde{\eta}_L < \tilde{\eta}_W$.

Similarly, we can derive that critical cell competition requires that $\tilde{\eta}_L = d_W$, i.e. winner emission is equal to loser tolerance, and that incomplete cell competition requires $\tilde{\eta}_L > d_W$. To summarise, we express the competition regimes in terms of tolerance and emission as

Cell competition: $d_L < \tilde{\eta}_L < \tilde{\eta}_W \wedge d_L < d_W$.

Complete cell competition: $d_L < \tilde{\eta}_L < d_W < \tilde{\eta}_W$.

Critical cell competition: $d_L < \tilde{\eta}_L = d_W < \tilde{\eta}_W$.

Incomplete cell competition: $d_L < d_W < \tilde{\eta}_L < \tilde{\eta}_W$.

Neutral competition: $d_L = d_W < \tilde{\eta}_L < \tilde{\eta}_W$.

Indirect competition: $d_W < d_L < \tilde{\eta}_L < \tilde{\eta}_W$.

These relationships can also be verified visually in Figure 6.13, where the competition regimes are plotted in the transformed parameter space $(d_A, \tilde{\eta}_A)$-space.

Partial inhibition of apoptosis can prevent elimination We learned in Section 6.7 that a cell type for which $\eta/\beta > \ln(2)$ cannot be eliminated by any competing cell type (see Cross Section I in Figure 6.8). This is also reflected in the fact that the complete and critical cell competition regimes are absent above the coexistence line in Cross Section I (see Figure 6.12). Since β is constrained to the open interval $(0, 1)$, a sufficient condition for being immune to elimination is $\eta \geq \ln(2)$.

As discussed above, we can interpret η/β as the cell's tolerance to death signals. We can express the condition for immunity to elimination using the transformed parameter $\tilde{\eta}$ as $\tilde{\eta} > 1$. Hence, the condition $\tilde{\eta} > 1$ expresses a lower bound for $\tilde{\eta}$ such that the cell

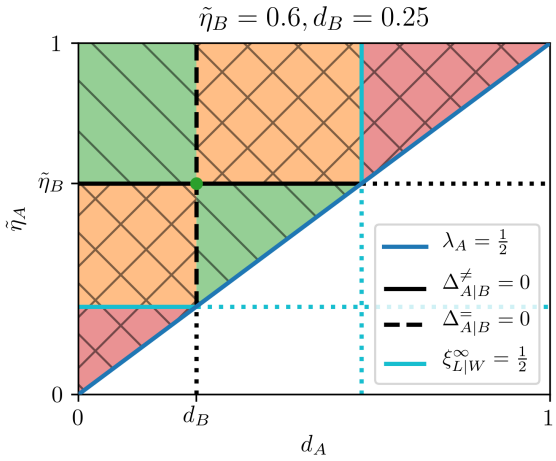

Figure 6.13: Diagram of competition regimes using the transformed parameters $\tilde{\eta}_X$, defined in Equation (6.85), and d_X, defined in Equation (6.86). The green dot corresponds to the point $(d_B, \tilde{\eta}_B)$. See Table 6.7 for the legend. The symbol $\xi^{\infty}_{L|W}$ is used to refer to the asymptotic survival probability of the loser cell type, i.e. cell type B for $\Delta^{\neq}_{A|B} > 0$ and cell type A for $\Delta^{\neq}_{A|B} < 0$.

type cannot be eliminated through competition. The underlying reason, as expressed by the tolerance and emission, is β is constrained as $0 < \beta < 1$, means that the emission is constrained as $0 < d < 1$. The upper bound on emission of death signal translates to a lower bound on the tolerance such that elimination is impossible. Hence, our model shows that partial inhibition of apoptosis can be sufficient to preclude elimination through cell competition. Since evasion of apoptosis is one of the hallmarks of cancer [16], this observation may have implications for super-competition and tumour formation.

Spatial mixing is required for cell competition In Section 6.4, we remarked that we observed smaller heterotypic survival differences in the vertex model than in the well-mixed model. Similarly, in Section 6.5, we observed smaller homotypic survival differences for the vertex model than for the well-mixed model. In addition, this effect was stronger for segregated initial conditions than random initial conditions. Finally, in Section 6.7 we saw that the loser elimination regime was smaller than predicted for random initial conditions, and completely absent for segregated initial conditions. In fact, comparing the heterotypic survival frequencies in Section 6.7 with the homotypic survival frequencies in Section 5.4, we see that the segregated cell populations behave essentially the same as in homotypic conditions.

The reason for this discrepancy between the predicted proliferative behaviour and the results of the vertex model is that the assumption of a well-mixed population, which we used

to derive the ergodic approximation to the heterotypic survival probability, is not respected in the vertex model. In particular, when cell types are spatially segregated, cells interact more often with cells of the same type rather than with cells of the other type. Therefore, the assumption that the local cell neighbourhood reflects the whole population does not hold. In the well-mixed model, on the other hand, this assumption is satisfied by construction, which explains why the well-mixed results agree better with predictions.

The proliferation regimes derived in Section 6.7 therefore do not describe the global behaviour of a tissue with competing cell populations, but rather the local behaviour at clone boundaries where multiple cell types meet and exchange death signals such that the system can *locally* be considered well-mixed. The local behaviour within clones, on the other hand, can be described by the proliferation regimes of homotypic populations because cells only interact with cells of the same type. Therefore, the degree of competition is dependent on the degree of heterotypic contact between cell types, i.e. spatial mixing is required for cell competition. This observation is also seen *in vivo* [206] and has been replicated in other cell-based models of cell competition [78].

Chapter 7

Discussion

7.1 Summary

In Chapter 1, we reviewed cell competition and identified a lack of models that investigate the principles determining winner/loser status. In particular, currently published models assert *a priori* winner/loser identities and therefore cannot explain how winner/loser status is determined. This motivated the research aim of this thesis: to study the emergence of winners and losers using mathematical modelling. To achieve this aim, we stated the following research objectives in Section 1.1:

Research Objective 1: construct a model of cell competition that produces winners and losers as an emergent phenomenon.

Research Objective 2: use this model to identify and study the key factors driving winner/loser identification.

Research Objective 3: use predictions from the model to propose experiments for validating the model.

In this section, we summarise the thesis and the progress made on these research objectives.

Chapter 2 describes our first attempt at achieving Research Objective 1. We investigated the hypothesis that differences in the mechanical parameters of a heterotypic population in a vertex model are sufficient to reproduce the dynamics of cell competition. In particular, we ran a large parameter sweep varying mechanical parameters of one cell type while keeping the parameters of the other cell type constant. Although we observed winners and losers in the sense that some cell types were more successful at proliferating, homotypic simulations showed that these fitness differences were due to intrinsic viability, and not competitive interactions.

Thus, we were unable to reproduce cell competition by varying mechanical parameters alone. This result agrees with reports in the literature that cell competition requires an active mechanism of cell death, such as apoptosis. Therefore, in the following chapters, we investigate a different hypothesis to achieve Research Objective 1: cell competition is mediated through the intercellular exchange of death signals.

In Chapter 3, we constructed the death clock framework to explore this hypothesis. The main assumptions underlying this framework are that i) cells perceive death signals from their surroundings; ii) cells trigger apoptosis if the accumulated death signal reaches a threshold; and iii) apoptosis can only be initiated in the stochastic G1 phase. An important distinction with current models of cell competition that are based on active cell death is that the death clock mechanism is active in all cells, and not only losers.

We also discussed the cell-based models that are used to implement the death clock framework. Specifically, we implemented the death clock mechanism in the vertex model using Chaste's object-oriented interface, and constructed a well-mixed model *de novo*. The vertex model lets us study the spatial dimensions of cell competition at the cost of expensive simulations. The well-mixed model, on the other hand, is much cheaper to run and has better analytical tractability, but cannot be used to study the spatial aspects of cell competition.

In Chapter 4, we analysed the simplest possible death clock model: the constant death signal model. We related the model's behaviour to that of a discrete-time birth–death Markov chain to establish different proliferation regimes, and found that population extinction and explosion are the only possible outcomes. In addition, we used cell-based simulations to verify these predictions. Although the constant death signal model cannot reproduce cell competition, the chapter provides an important theoretical foundation for the research carried out in the next two chapters.

In order to achieve Research Objective 1 with the death clock framework, we hypothesise a death signal that is emitted by cells in G2 phase as a mechanism for cell competition. In Chapter 5, we characterised the behaviour of the G2 death signal model for homotypic populations. In particular, we derived the parameter regimes for homotypic viability, and validated the predictions with cell-based simulations. This is significant because the precondition for cell competition is that the competing cell types are viable in homotypic conditions.

In Chapter 6, we studied the G2 death signal model in heterotypic populations. Through theoretical predictions and cell-based simulations, we found that the G2 death signal model is indeed able to reproduce the hallmarks of cell competition, thus achieving Research Objective 1. In particular, we discovered parameter regimes in which two intrinsically viable cell types compete such that only one cell type remains. Moreover, we also describe

closely related parameter regimes representing other forms of competition. Finally, we identified tolerance to death signals and emission of death signals as the main determinants of winner/loser status, thus achieving Research Objective 2.

In Section 7.2, we discuss the implications of the model from a biological perspective and make suggestions for how it can be experimentally validated, thus achieving Research Objective 3.

7.2 Discussion

The tolerance–emission model of cell competition Our work makes the following prediction: cell competition requires that winners have a higher tolerance to death signals *and* a higher rate of death signal emission than losers. As we showed in Section 6.9, this follows from the intrinsic viability of the winner, i.e. $d_W < \tilde{\eta}_W$, the intrinsic viability of the loser, i.e. $d_L < \tilde{\eta}_L$, and the loser elimination condition, i.e. $\tilde{\eta}_L < d_W$. The intuitive interpretation is that winners emit death signals that are tolerated by the winners, but not by the losers, hence losers must have a lower tolerance to death signals. Moreover, losers are viable when exposed to their intrinsic emission of death signals, but are eliminated by the emission of winners, hence the winners must have a higher emission of death signals than losers.

Importantly, our prediction suggests that mutations which affect only the tolerance, only the emission, or neither do *not* create a cell type that outcompetes or is outcompeted by wild-type cells. This explains why some mutations result in cell competition and others do not [42]. Apoptosis inhibition can be regarded as a mutation which affects tolerance but not emission. Consistent with our model-based predictions, it has been shown in experiments that inhibiting apoptosis prevents endogenous cell competition from taking place [42, 53]. In addition, inhibiting apoptosis in a complete loser[1] prevents loser elimination [53].

A question that arises from this prediction is whether there exist any mutations that increase emission but not tolerance? If there are, they would be difficult to identify, because our theory suggests that such mutants would be intrinsically nonviable. However, our model suggests that such a mutation may be successful when paired with apoptosis inhibition. This suggests that we may look for emission-enhancing mutations in apoptosis-inhibited cells. Our framework predicts that the combination of these two mutations would create a super-competitor.

Experimental support for the tolerance–emission hypothesis Experimental evidence in *Myc*-based cell competition supports the tolerance–emission hypothesis. In [207], the

[1] See end of Section 6.8 (page 171) for the definition of a "complete loser".

authors demonstrate that the ligand Spätzle is necessary for the elimination of loser cells in the *Drosophila* wing disc when outcompeted by *Myc* super-competitors. They thus establish that Spätzle is, in their words, a "killing signal". However, they also observe that Spätzle is produced continuously in wild-type conditions, but in quantities that are too small to induce cell death in wild-type cells. The production of Spätzle is upregulated in *Myc* super-competitors, indicating a higher emission of death signals. Furthermore, enhanced Spätzle signalling is responsible for inducing cell death in loser cells, but does not induce cell death in super-competitors, which indicates a higher tolerance in winners.

The death signal in this case is mediated through a death ligand. We did not model diffusible death ligands in this thesis, but we hypothesise that the tolerance–emission model applies to death ligand-mediated cell competition as well. The winners and losers should be able to tolerate their own rate of death ligand secretion for intrinsic viability, and the loser cell type should not tolerate the winner cell type's intrinsic death ligand level, such that the premises of the tolerance–emission model remain unchanged.

An important difference with contact-based death signals, however, is that death ligands can diffuse away from the heterotypic boundary. As discussed in Section 6.9, we assume that the populations are well-mixed in the derivation of the competition regimes, which limits the scope of our analysis to interactions at heterotypic boundaries, where multiple cell types mix and exchange death signals. A diffusible death ligand would increase the area where the death signal can be considered well-mixed to bands centred on the heterotypic boundaries. This tolerance–emission model may explain therefore why we observe loser cell death at a distance from the winner clone in *Myc* cell competition and not in *Minutes* cell competition. Since *Myc* super-competitors secrete death ligands at a higher rate than wild-type cells, the death ligand potentially travels further away from the winner clone.

The influence of proliferation rates A potential limitation of our model is that we did not consider differential proliferation rates, i.e. $\kappa_{A/B} \neq 1$. Differences in proliferation rates are often associated with cell competition, with *Myc*-transformed cells having a higher proliferation rate and *Minutes*-transformed cells having a lower proliferation rate. Indeed, differential proliferation rates could alter the relative abundance of cell types. For example, a cell type with a lower survival probability can become more abundant than another cell type if it also has a much shorter cell cycle time. Hence, this could affect which cell type becomes more abundant. However, without increasing emission the faster-growing cell type cannot create the conditions for loser elimination, no matter how fast they grow.

Furthermore, not all mutations that increase proliferation rates result in cell competition [42]. Our hypothesis suggests that such mutations likely do not result in cell competition

because they do not affect the tolerance and emission of death signals. Mutations that only increase proliferation rates can be considered a form of neutral competition, where cells become more abundant without affecting death signal emission and thus without increasing cell death in wild-type cells.

The function of cell competition The question remains, however, why are mutations associated with cell competition then so closely linked with proliferation rates? This is not a question that we can answer within the scope of this thesis, but we speculate that it is related to the function of cell competition. It has been proposed that cell competition is a mechanism for tissue homeostasis, eliminating unfit cells to improve overall tissue health. There is a cost associated with cell competition, however, both in terms of the death signal machinery, and in terms of eliminating cells that i) are viable and ii) the body has already expended resources on.

Hence, the gain from cell competition should compensate for this cost. In this respect, it is interesting to note that cell competition takes place during development, where it is necessary to create a large number of cells in a short amount of time. For example, the *Drosophila* wing disc expands from 50 to 50 000 cells in the span of four days [208]. During mouse embryogenesis, the epidermis, where cell competition has also been observed [59], expands its surface area 30-fold.

Since our hypothesis suggests that differential proliferation rates are not the mechanism of cell competition, we conjecture that they are the *target* of cell competition, in the sense that cell competition aims to eliminate cells that are worse than their peers at dividing quickly. From an evolutionary perspective, cell competition may only evolve in systems where the ability to contribute usefully to the tissue can be linked to a death signal. In this respect, it is interesting to note that the death signal is linked to protein synthesis in *Minutes*-based competition, since protein synthesis is an important process in rapid cell division and growth.

Experimental validation of the tolerance–emission hypothesis In our model, inhibiting apoptosis corresponds to taking the limit $\tilde{\eta} \to \infty$, such that the survival probability approaches unity. Our model correctly predicts that apoptosis inhibition in the loser cell type prevents loser elimination, which agrees with experiments [42, 53]. More specifically, taking $\tilde{\eta}_L \to \infty$ moves the system from complete cell competition into the indirect competition regime with complete winners becoming indirect losers and complete losers becoming indirect winners. In addition, our model correctly predicts that apoptosis inhibition in the winner cell type does not alter the outcome of cell competition, which has also been verified

experimentally [42]. However, since these experiments were available during the construction of the model, these are only validations in the weak sense that the model reproduces known data. The true test of the model will be to extrapolate to experimental conditions that so far have not been tested.

If we relax the condition of loser viability, there are more proliferation regimes available to us, as described in Section 6.7. A particularly interesting regime is nonviable loser rescue. This is an example of indirect competition where an intrinsically viable winner, say cell type A, with a lower emission than the loser creates the conditions of viability for a loser cell type, say cell type B, that is intrinsically nonviable. In a tissue, we would observe this as small clones of cell type B surviving in a background of cell type A, with cell type B being nonviable when the whole tissue is composed of them.

However, conducting such an experiment would require looking for mutations that have not yet been identified. Moreover, there is the added difficulty that cell type B is not viable on its own. We therefore suggest an alternative experiment that simulates the intrinsic nonviability conditions by running the experiment in a high emission background. Consider a triple co-culture where cell type C outcompetes cell type B and cell type B outcompetes cell type A. Both cell types A and B are eliminated in a C-type background, simulating intrinsic nonviability. If we inhibit apoptosis in B, we expect B to survive and A to be eliminated. This is not different to the situation described above where apoptosis inhibition rescues a viable loser cell type from elimination.

However, when we inhibit apoptosis in cell type A (and assuming that inhibiting apoptosis does not affect emission), then we expect the apoptosis-inhibited A-type cells to survive and create a low emission environment in which B-type cells can also survive. We expect this to be the case because B-type cells outcompete normal A-type cells, which, according to our hypothesis, means that A-type cells have a lower emission than B-type cells and that B-type cells can tolerate the A-type emission of death signals. Furthermore, we expect that B-type cells can *only* survive in the apoptosis-inhibited A-type clones, and not outside of them.

7.3 Future work

Exploring the role of relative proliferation rates In this chapter, we assumed that $t_{\mathrm{G},A} = t_{\mathrm{G},B}$ for the sake of simplicity. This means that the autonomous proliferation rates are identical, which allowed us to characterise the proliferative behaviour of heterotypic populations purely in terms of survival probabilities. If we relax this assumption and let the ratio $\kappa_{A/B} = t_{\mathrm{G},A}/t_{\mathrm{G},B}$ be less or greater than unity, we would expect that the relative

proliferation rate affects the abundance of cell types, and hence the winner/loser status. Indeed, preliminary results (not shown) suggest that a shorter total cell cycle time can turn a loser cell type into a winner cell type, everything else being equal.

Because the proliferation rate does not affect the emission or tolerance, however, there are some constraints on the types of winners that losers can become and vice versa by changing $\kappa_{A/B}$. For instance, suppose that cell type A is a complete loser, i.e. $d_A < \tilde{\eta}_A < d_B < \tilde{\eta}_B$, when the cell cycle times are equal, i.e. $\kappa_{A/B} = 1$. If we change $t_{\mathrm{G},A}$ such that $\kappa_{A/B} \ll 1$ while keeping everything else equal, and assume that cell type A becomes the winner instead, then cell type A must be an indirect winner because it has a lower intrinsic death signal than cell type B. Future work is needed to characterise the conditions for which changing $\kappa_{A/B}$ can switch winner/loser status and to verify this with simulations.

Extending the death clock framework The death signal model is very flexible and can be implemented in any continuous-time cell-based model that has rules for division and apoptosis. In addition, any stochastic distribution with non-negative support can be used for the stochastic G1 phase. Most importantly, different death signals can be used to represent different hypothesised mechanisms of cell competition, such as diffusible death ligands and mechanical compression. We discuss three such death signals in the following paragraphs.

Constant emission death signal The transformed parameters defined in Section 6.9 suggest a simpler model where every cell α emits death signals to its neighbours at a constant rate d_α, instead of only in the G2 phase. The death signal received by a particular cell is then obtained by averaging the emitted death signals of its neighbours. We call this the constant emission death signal model, or the constant emission model for brevity.

Preliminary theoretical analysis (not shown) demonstrates that the homotypic survival probability and proliferation regimes in homotypic and heterotypic conditions for the constant emission model can be derived in an analogous manner to the G2 death signal model (see Figure 7.1). Importantly, the classification of competitive interactions and competition regimes apply virtually identically to the constant emission model. In particular, compare the bottom right plot in Figure 7.1 to Figure 6.13.

An important difference, however, is that there is no need to assert ergodicity. In effect, the purpose of the ergodic approximation is to reduce the time-varying G2 death signal to a constant rate of emission. The constant emission model obviates the need for such an approximation, so it is potentially more generalisable. Nevertheless, one should always pick the most appropriate death signal model based on the particular application and context. As we saw in Chapter 5, the property of ergodicity in the G2 death signal model, and

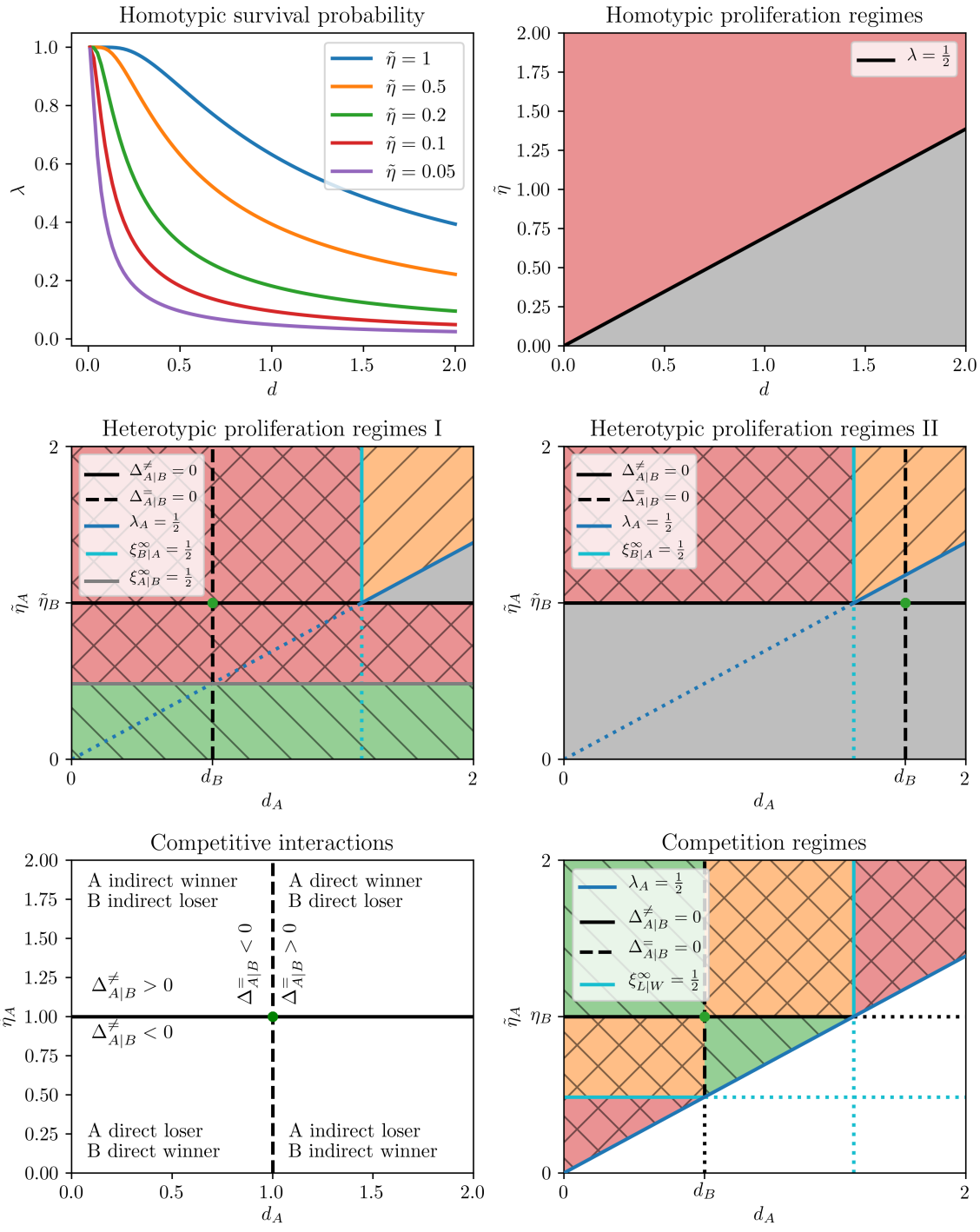

Figure 7.1: Diagrams for the constant emission death signal model with the exponential cell cycle model. **Top left**: homotypic survival probability, analogous to Figure 5.1. **Top right**: homotypic proliferation regimes, analogous to Figure 5.7. **Middle left and right**: cross sections showing heterotypic proliferation regimes, analogous to Figure 6.8. **Bottom left**: classification of competitive interactions, analogous to Figure 6.7. **Bottom right**: competition regimes, analogous to Figures 6.12 and 6.13.

the theoretical analysis based on it, is contingent on the cell cycle model. We therefore hypothesise that the constant emission model is more robust with respect to the cell cycle model in terms of conforming to the theoretical framework we developed in this thesis. Future work is needed to validate the theoretical analysis in Figure 7.1 with cell-based simulations.

Diffusible death ligand Experiments suggest that cell competition in the *Drosophila* wing disc is mediated by diffusible death ligands [54, 207, 209]. This suggests implementing a death clock model where cells produce death ligands (emission) that are perceived by the cells as a death signal. The well-mixed model could be extended with an ODE model that models the production and uptake of this ligand by competing cells. In the vertex model, the production, diffusion, and uptake of the death ligand could be modelled using a reaction–diffusion model [71, 210]. Future work is needed to implement, analyse, and simulate a diffusible death ligand model.

Mechanical cell competition The motivation for developing the death clock framework was the observation that varying mechanical parameters is not enough to reproduce cell competition in a heterotypic vertex model. However, mechanical parameters may be important for mechanical cell competition, in which cells are outcompeted through mechanical compression [64]. In particular, losers remove themselves from the tissue by apoptosis in response to mechanical compression. This suggests that mechanical compression could be modelled as a death signal for cell competition. Tolerance would then correspond to a cell's ability to resist compression, while emission would correspond to a cell's ability to proliferate and take up space. Future work is needed to implement, analyse, and simulate a mechanical death signal model.

Emergence of winners and losers Finally, our work shows the value of constructing models where the winner/loser status is emergent, rather than asserted. In particular, the methodology used in this thesis can be applied more generally to cell competition models. The first step is to construct a heterotypic model with two cell types that vary only in their parameters. The second step is to study the conditions for homotypic viability. The third step is to study the conditions for heterotypic viability. The fourth step is to combine the results of the second and third step to characterise the competition regimes. This can be done either analytically by deriving a survival probability, as we did for the death clock model, or computationally by running a parameter sweep and classifying simulations based on viability criteria, as we did in Chapter 2. Future work is needed to apply this general

framework to study the emergence of winners and losers in other types of models.

Appendix A

Chaste contributions

As explained in Section 3.3.2, we had to identify and fix certain issues in Chaste in order to reduce the failure rate of simulations to an acceptable level. The Chaste project uses a ticketing system to keep track of, and organise, outstanding and resolved issues. We opened a total of four tickets (3062, 3064, 3065, and 3067), and closed three of them (3062, 3064, and 3065). In addition, we made a contribution to a pre-existing ticket (2401).

Ticket 3062: Boundary node intersection bug

URL `https://chaste.cs.ox.ac.uk/trac/ticket/3062`

Background It was already known that in certain situations cells can become concave and self-intersecting. The intersection swap was implemented to deal with self-intersections by rearranging the connectivity between the nodes of the self-intersecting cell and its neighbours [157]. See Figure A.1 for an illustration.

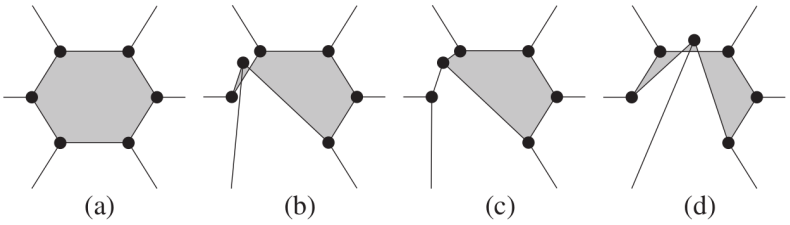

Figure A.1: Illustration of the intersection swap. (a) A convex cell. (b) A concave, self-intersecting cell. (c) The intersection swap modifies the topology of the intersecting node to resolve the self-intersection. (d) The intersection swap cannot resolve this type of self-intersection, as it would result in splitting the cell into two. Figure reproduced from [157] with permission from Elsevier.

Problem In principle, intersection swaps are agnostic with respect to whether the intersecting node is an interior node or a boundary node. However, the implementation of the function `PerformIntersectionSwap()` was written under the assumption of interior nodes. In the case of interior nodes, there are four cells that contain the intersecting node, but for boundary nodes one or more of these cells may be missing. Therefore, for boundary nodes, `PerformIntersectionSwap()` would attempt to operate on cells that did not exist, resulting in a fatal crash.

Solution We modified `PerformIntersectionSwap()` so that it checks for the existence of a neighbouring cell before operating on it. In this way, both interior and boundary nodes can be handled by the same function.

Outcome Ticket closed and fix included in the Chaste 2021.1 release[1].

Ticket 3064: Bug in T2 swap on element without neighbours

URL `https://chaste.cs.ox.ac.uk/trac/ticket/3064`

Background As discussed extensively in Section 4.2, the entire cell population can go extinct in the constant death signal model. When a cell initiates apoptosis in the vertex model, its area shrinks until it reaches a threshold and is removed through a T2 swap, replacing the cell with a node (see Figure 2.2).

Problem The node replacing the extruded cell in a T2 swap is normally associated to the neighbours of the extruded cell. However, when the entire tissue goes extinct, there are no neighbours to associate the node to. Instead, the last cell is replaced by an isolated node, which is an invalid mesh.

Solution We added a check to the beginning of `PerformT2Swap()` to see if the cell being extruded has any neighbours. If not, then we simply remove the cell without creating a replacing node and exit the function.

Outcome Ticket closed and fix included in the Chaste 2021.1 release.

[1] `https://github.com/Chaste/Chaste/releases/tag/release_2021.1`

Ticket 3065: T3 swaps versus intersection swaps

URL `https://chaste.cs.ox.ac.uk/trac/ticket/3065`

Background The function `CheckForIntersections()` is responsible for performing T3 swaps and intersection swaps. By default, this function only performs T3 swaps and no intersection swaps.

Problem The flag `mCheckForIntersections` enables intersection swaps while simultaneously disabling T3 swaps. Hence, T3 swaps and intersection swaps were programmatically mutually exclusive, even though the conditions for both types of mesh rearrangements could occur in the same simulation.

Solution We implemented the flag `mCheckForT3Swaps`, which enables T3 swaps independently of intersection swaps. Hence, now T3 swaps and intersection swaps can both be enabled within the same simulation.

Outcome Ticket closed and fix merged into the Chaste development branch for inclusion in the next Chaste release.

Ticket 3067: T1 swap can result in concave triangular element

URL `https://chaste.cs.ox.ac.uk/trac/ticket/3067`

Background A triangular cell should never become concave because its area would hit zero and be extruded by a T2 swap before that can happen.

Problem For very elongated quadrilateral cells, a T1 swap occurring on a short edge can result in a concave triangular cell because of the sudden topological change, as illustrated in Figure A.2.

Outcome Ticket opened. We did not implement a fix because the issue only manifests in less than 1% of simulations and a fix would require significant structural changes in how Chaste handles mesh rearrangements.

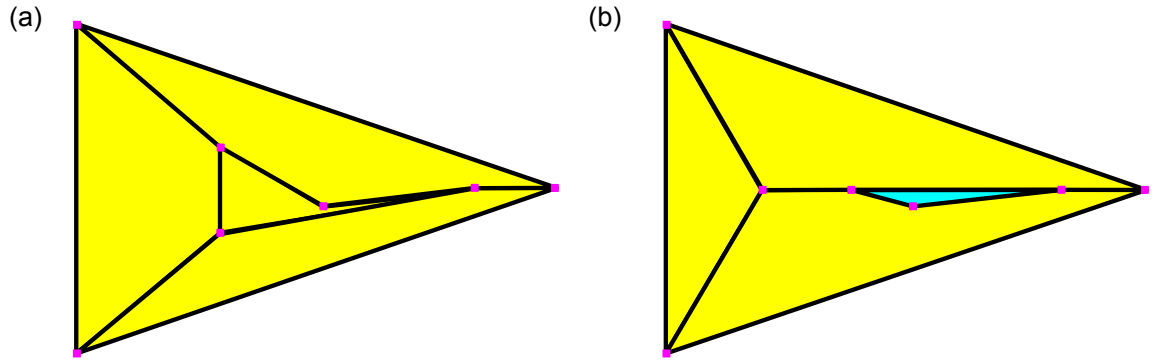

Figure A.2: Illustration of a T1 swap resulting in a concave triangular element. (a) **Before T1 swap.** The quadrilateral cell in the centre is elongated horizontally and has a short edge on the left-hand side with an edge length below the T1 threshold. (b) **After T1 swap.** The T1 swap on the short edge creates a concave triangular element, indicated in cyan.

Ticket 2401: Rationalise and look for efficiencies in vertex remeshing code

URL `https://chaste.cs.ox.ac.uk/trac/ticket/2401`

Background Previously, the main loop in `CheckForIntersections()` tried every possible pair of node and cell to check whether they intersect before proceeding with an intersection swap.

Problem The number of checks grows on the order of the number of nodes times the number of cells. This is prohibitively time-consuming for large meshes.

Solution We optimised the loop by only considering the cells that are neighbours of the cells containing the node in question, instead of all cells. Assuming a roughly constant number of neighbours per cell, the run time of the loop was reduced from $O(N_{\text{node}} \cdot N_{\text{cell}})$ to $O(N_{\text{node}})$.

Outcome Optimisation included in the Chaste 2021.1 release.

Bibliography

[1] OED Online, "competition, n." *Oxford University Press*, https://www.oed.com/view/Entry/37578, 2021. [Online; accessed 5 September 2021].

[2] T. R. Malthus, *An Essay on the Principle of Population*. London: J. Johnson, 1798.

[3] C. Darwin, *On the Origin of Species*. London: John Murray, 1859.

[4] F. A. Huntingford and A. K. Turner, *Animal Conflict*. London: Chapman and Hall, 1987.

[5] J. M. Craine and R. Dybzinski, "Mechanisms of plant competition for nutrients, water and light," *Functional Ecology*, vol. 27, no. 4, pp. 833–840, 2013.

[6] M. E. Hibbing, C. Fuqua, M. R. Parsek, and S. B. Peterson, "Bacterial competition: surviving and thriving in the microbial jungle," *Nature Reviews Microbiology*, vol. 8, pp. 15–25, 2010.

[7] E. T. Granato, T. A. Meiller-Legrand, and K. R. Foster, "The evolution and ecology of bacterial warfare," *Current Biology*, vol. 29, pp. R521–R537, 2019.

[8] E. T. Granato and K. R. Foster, "The evolution of mass cell suicide in bacterial warfare," *Current Biology*, vol. 30, pp. 2836–2843, 2020.

[9] A. J. Desmond and J. R. Moore, *Darwin*. London: Penguin Books Ltd, 1991.

[10] S. E. Smith and D. J. Read, *Mycorrhizal Symbiosis*. London: Academic Press, 3rd ed., 2010.

[11] E. M. Quigley, "Gut bacteria in health and disease," *Gastroenterology and Hepatology*, vol. 9, no. 9, pp. 560–569, 2013.

[12] A. M. Kramer, B. Dennis, A. M. Liebhold, and J. M. Drake, "The evidence for Allee effects," *Population Ecology*, vol. 51, pp. 341–354, 2009.

[13] R. E. Ellis, J. Yuan, and H. R. Horvitz, "Mechanisms and functions of cell death," *Annual Review of Cell Biology*, vol. 7, pp. 663–698, 1991.

[14] M. C. Raff, "Social controls on cell survival and cell death," *Nature*, vol. 356, pp. 397–400, 1992.

[15] J. C. Ameisen, "On the origin, evolution, and nature of programmed cell death: a timeline of four billion years," *Cell Death and Differentiation*, vol. 9, pp. 367–393, 2002.

[16] D. Hanahan and R. A. Weinberg, "The hallmarks of cancer," *Cell*, vol. 100, no. 1, pp. 57–70, 2000.

[17] L. A. Johnston, "Competitive interactions between cells: death, growth, and geography," *Science*, vol. 324, no. 5935, pp. 1679–1682, 2009.

[18] S. Vivarelli, L. Wagstaff, and E. Piddini, "Cell wars: regulation of cell survival and proliferation by cell competition," *Essays In Biochemistry*, vol. 53, pp. 69–82, 2012.

[19] R. Levayer and E. Moreno, "Mechanisms of cell competition: themes and variations," *Journal of Cell Biology*, vol. 200, no. 6, pp. 689–698, 2013.

[20] J. P. Vincent, A. G. Fletcher, and L. A. Baena-Lopez, "Mechanisms and mechanics of cell competition in epithelia," *Nature Reviews Molecular Cell Biology*, vol. 14, no. 9, pp. 581–591, 2013.

[21] M. Amoyel and E. A. Bach, "Cell competition: how to eliminate your neighbours," *Development*, vol. 141, no. 5, pp. 988–1000, 2014.

[22] S. Bowling, K. Lawlor, and T. A. Rodríguez, "Cell competition: the winners and losers of fitness selection," *Development*, vol. 146, p. dev167486, 2019.

[23] P. J. Morin, *Community Ecology*. John Wiley & Sons, 2nd ed., 2011.

[24] V. Smil, "Detonator of the population explosion," *Nature*, vol. 400, p. 415, 1999.

[25] R. E. Evenson and D. Gollin, "Assessing the impact of the Green Revolution, 1960 to 2000," *Science*, vol. 300, no. 5620, pp. 758–762, 2003.

[26] D. Schluter, "Ecological causes of adaptive radiation," *The American Naturalist*, vol. 148, pp. S40–S64, 1996.

[27] J. H. Connell, "The influence of interspecific competition and other factors on the distribution of the barnacle Chthamalus Stellatus," *Ecology*, vol. 42, no. 4, pp. 710–723, 1961.

[28] N. A. Campbell and J. B. Reece, *Biology*. San Francisco: Pearson Education publishing as Benjamin Cummings, 7th ed., 2005.

[29] R. H. Masure and W. C. Allee, "The social order in flocks of the common chicken and the pigeon," *The Auk*, vol. 51, no. 3, pp. 306–327, 1934.

[30] G. K. Noble, "The rôle of dominance in the social life of birds," *The Auk*, vol. 56, no. 3, pp. 263–273, 1939.

[31] S. A. Queiroz and V. U. Cromberg, "Aggressive behavior in the genus Gallus sp," *Brazilian Journal of Poultry Science*, vol. 8, no. 1, pp. 1–14, 2006.

[32] A. M. Guhl and W. C. Allee, "Some measurable effects of social organization in flocks of hens," *Physiological Zoology*, vol. 17, no. 3, pp. 320–347, 1944.

[33] S. Tonegawa, "Somatic generation of antibody diversity," *Nature*, vol. 302, pp. 575–581, 1983.

[34] F. W. Alt, E. M. Oltz, F. Young, J. Gorman, G. Taccioli, and J. Chen, "VDJ recombination," *Immunology Today*, vol. 13, no. 8, pp. 306–314, 1992.

[35] L. Klein, B. Kyewski, P. M. Allen, and K. A. Hogquist, "Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see)," *Nature Reviews Immunology*, vol. 14, pp. 377–391, 2014.

[36] T. K. Starr, S. C. Jameson, and K. A. Hogquist, "Positive and negative selection of T cells," *Annual Review of Immunology*, vol. 21, pp. 139–176, 2003.

[37] G. Morata and P. Ripoll, "Minutes: mutants of Drosophila autonomously affecting cell division rate," *Developmental Biology*, vol. 42, no. 2, pp. 211–221, 1975.

[38] S. J. Marygold, J. Roote, G. Reuter, A. Lambertsson, M. Ashburner, G. H. Millburn, P. M. Harrison, Z. Yu, N. Kenmochi, T. C. Kaufman, S. J. Leevers, and K. R. Cook, "The ribosomal protein genes and Minute loci of Drosophila melanogaster," *Genome Biology*, vol. 8, no. 10, 2007.

[39] P. Simpson and G. Morata, "Differential mitotic rates and patterns of growth in compartments in the Drosophila wing," *Developmental Biology*, vol. 85, no. 2, pp. 299–308, 1981.

[40] P. Simpson, "Parameters of cell competition in the compartments of the wing disc of Drosophila," *Developmental Biology*, vol. 69, no. 1, pp. 182–193, 1979.

[41] L. A. Johnston, D. A. Prober, B. A. Edgar, R. N. Eisenman, and P. Gallant, "Drosophila myc regulates cellular growth during development," *Cell*, vol. 98, no. 6, pp. 779–790, 1999.

[42] C. De La Cova, M. Abril, P. Bellosta, P. Gallant, and L. A. Johnston, "Drosophila myc regulates organ size by inducing cell competition," *Cell*, vol. 117, no. 1, pp. 107–116, 2004.

[43] E. Moreno and K. Basler, "dMyc transforms cells into super-competitors," *Cell*, vol. 117, no. 1, pp. 117–129, 2004.

[44] M. Vita and M. Henriksson, "The Myc oncoprotein as a therapeutic target for human cancer," *Seminars in Cancer Biology*, vol. 16, no. 4, pp. 318–330, 2006.

[45] E. Moreno, "Is cell competition relevant to cancer?," *Nature Reviews Cancer*, vol. 8, no. 2, pp. 141–147, 2008.

[46] G. D. Dakubo, J. P. Jakupciak, M. A. Birch-Machin, and R. L. Parr, "Clinical implications and utility of field cancerization," *Cancer Cell International*, vol. 7, pp. 1–12, 2007.

[47] D. Bilder, "Epithelial polarity and proliferation control: links from the Drosophila neoplastic tumor suppressors," *Genes and Development*, vol. 18, no. 16, pp. 1909–1925, 2004.

[48] J. P. Campanale, T. Y. Sun, and D. J. Montell, "Development and dynamics of cell polarity at a glance," *Journal of Cell Science*, vol. 130, no. 7, pp. 1201–1207, 2017.

[49] N. A. Grzeschik, L. M. Parsons, M. L. Allott, K. F. Harvey, and H. E. Richardson, "Lgl, aPKC, and Crumbs regulate the Salvador/Warts/Hippo pathway through two distinct mechanisms," *Current Biology*, vol. 20, no. 7, pp. 573–581, 2010.

[50] Y. Tamori, C. U. Bialucha, A. G. Tian, M. Kajita, Y. C. Huang, M. Norman, N. Harrison, J. Poulton, K. Ivanovitch, L. Disch, T. Liu, W. M. Deng, and Y. Fujita, "Involvement of Lgl and mahjong/VprBP in cell competition," *PLoS Biology*, vol. 8, no. 7, 2010.

[51] J. Menendez, A. Perez-Garijo, M. Calleja, and G. Morata, "A tumor-suppressing mechanism in Drosophila involving cell competition and the Hippo pathway," *Proceedings of the National Academy of Sciences*, vol. 107, no. 33, pp. 14651–14656, 2010.

[52] C.-L. Chen, M. C. Schroeder, M. Kango-Singh, C. Tao, and G. Halder, "Tumor suppression by cell competition through regulation of the Hippo pathway," *Proceedings of the National Academy of Sciences*, vol. 109, no. 2, pp. 484–489, 2012.

[53] E. Moreno, K. Basler, and G. Morata, "Cells compete for decapentaplegic survival factor to prevent apoptosis in Drosophila wing development," *Nature*, vol. 416, no. 6882, pp. 755–759, 2002.

[54] S. N. Meyer, M. Amoyel, C. Bergantiños, C. De La Cova, C. Schertel, K. Basler, and L. A. Johnston, "An ancient defense system eliminates unfit cells from developing tissues during cell competition," *Science*, vol. 346, no. 6214, 2014.

[55] C. Rhiner, J. M. López-Gay, D. Soldini, S. Casas-Tinto, F. A. Martín, L. Lombardía, and E. Moreno, "Flower forms an extracellular code that reveals the fitness of a cell to its neighbors in Drosophila," *Developmental Cell*, vol. 18, no. 6, pp. 985–998, 2010.

[56] E. R. Oliver, "Ribosomal protein L24 defect in Belly spot and tail (Bst), a mouse Minute," *Development*, vol. 131, no. 16, pp. 3907–3920, 2004.

[57] C. Clavería, G. Giovinazzo, R. Sierra, and M. Torres, "Myc-driven endogenous cell competition in the early mammalian embryo," *Nature*, vol. 500, no. 7460, pp. 39–44, 2013.

[58] M. Sancho, A. Di-Gregorio, N. George, S. Pozzi, J. M. Sánchez, B. Pernaute, and T. A. Rodríguez, "Competitive interactions eliminate unfit embryonic stem cells at the onset of differentiation," *Developmental Cell*, vol. 26, no. 1, pp. 19–30, 2013.

[59] S. J. Ellis, N. C. Gomez, J. Levorse, A. F. Mertz, Y. Ge, and E. Fuchs, "Distinct modes of cell competition shape mammalian tissue morphogenesis," *Nature*, vol. 3, 2019.

[60] B. I. Shraiman, "Mechanical feedback as a possible regulator of tissue growth," *Proceedings of the National Academy of Sciences*, vol. 102, no. 9, pp. 3318–3323, 2005.

[61] E. Marinari, A. Mehonic, S. Curran, J. Gale, T. Duke, and B. Baum, "Live-cell delamination counterbalances epithelial growth to limit tissue overcrowding," *Nature*, vol. 484, no. 7395, pp. 542–545, 2012.

[62] G. T. Eisenhoffer, P. D. Loftus, M. Yoshigi, H. Otsuna, C. B. Chien, P. A. Morcos, and J. Rosenblatt, "Crowding induces live cell extrusion to maintain homeostatic cell numbers in epithelia," *Nature*, vol. 484, no. 7395, pp. 546–549, 2012.

[63] R. Levayer, C. Dupont, and E. Moreno, "Tissue crowding induces caspase-dependent competition for space," *Current Biology*, vol. 26, no. 5, pp. 670–677, 2016.

[64] L. Wagstaff, M. Goschorska, K. Kozyrska, G. Duclos, I. Kucinski, A. Chessel, L. Hampton-O'Neil, C. R. Bradshaw, G. E. Allen, E. L. Rawlins, P. Silberzan, R. E. Salas, and E. Piddini, "Mechanical cell competition kills cells via induction of lethal p53 levels," *Nature Communications*, vol. 7, 2016.

[65] S. A. Gudipaty, J. Lindblom, P. D. Loftus, M. J. Redd, K. Edes, C. F. Davey, V. Krishnegowda, and J. Rosenblatt, "Mechanical stretch triggers rapid epithelial cell division through Piezo1," *Nature*, vol. 543, no. 7643, pp. 118–121, 2017.

[66] C. Brás-Pereira and E. Moreno, "Mechanical cell competition," *Current Opinion in Cell Biology*, vol. 51, pp. 15–21, 2018.

[67] A. Szabó and R. M. Merks, "Cellular Potts modeling of tumor growth, tumor invasion, and tumor evolution," *Frontiers in Oncology*, vol. 3, pp. 1–12, 2013.

[68] G. B. Ermentrout and L. Edelstein-Keshet, "Cellular automata approaches to biological modeling," *Journal of Theoretical Biology*, vol. 160, no. 1, pp. 97–133, 1993.

[69] Y. Lee, S. Kouvroukoglou, L. V. McIntire, and K. Zygourakis, "A cellular automaton model for the proliferation of migrating contact-inhibited cells," *Biophysical journal*, vol. 69, no. 4, pp. 1284–1298, 1995.

[70] A. Deutsch and S. Dormann, *Cellular Automaton Modeling of Biological Pattern Formation: Characterization, Applications, and Analysis*. Birkhäuser Basel, 2005.

[71] J. M. Osborne, A. G. Fletcher, J. M. Pitt-Francis, P. K. Maini, and D. J. Gavaghan, "Comparing individual-based approaches to modelling the self-organization of multicellular tissues," *PLoS Computational Biology*, vol. 13, no. 2, p. e1005387, 2017.

[72] F. Graner and J. A. Glazier, "Simulation of biological cell sorting using a two-dimensional extended Potts model.," *Physical Review Letters*, vol. 69, no. 13, pp. 2013–2016, 1992.

[73] J. A. Izaguirre, R. Chaturvedi, C. Huang, T. Cickovski, J. Coffland, G. Thomas, G. Forgacs, M. Alber, G. Hentschel, S. A. Newman, and J. A. Glazier, "CompuCell, a multi-model framework for simulation of morphogenesis," *Bioinformatics*, vol. 20, no. 7, pp. 1129–1137, 2004.

[74] A. Shirinifard, J. S. Gens, B. L. Zaitlen, N. J. Popławski, M. Swat, and J. A. Glazier, "3D multi-cell simulation of tumor growth and angiogenesis," *PLoS ONE*, vol. 4, no. 10, 2009.

[75] J. Silvertown, S. Holtier, J. Johnson, and P. Dale, "Cellular automaton models of interspecific competition for space–the effect of pattern on process," *Journal of Ecology*, vol. 80, no. 3, pp. 527–533, 1992.

[76] R. Durrett and S. Levin, "Spatial aspects of interspecific competition," *Theoretical Population Biology*, vol. 53, no. 1, pp. 30–43, 1998.

[77] K. S. Korolev, J. B. Xavier, D. R. Nelson, and K. R. Foster, "A quantitative test of population genetics using spatiogenetic patterns in bacterial colonies," *American Naturalist*, vol. 178, no. 4, pp. 538–552, 2011.

[78] D. Gradeci, A. Bove, G. Vallardi, A. R. Lowe, S. Banerjee, and G. Charras, "Cell-scale biophysical determinants of cell competition in epithelia," *eLife*, vol. 10, p. e61011, 2021.

[79] E. M. Purcell, "Life at low Reynolds number," *American Journal of Physics*, vol. 45, no. 1, pp. 3–11, 1977.

[80] D. Drasdo and S. Höhme, "A single-cell-based model of tumor growth in vitro: monolayers and spheroids," *Physical Biology*, vol. 2, no. 3, pp. 133–147, 2005.

[81] Y. S. Chu, S. Dufour, J. P. Thiery, E. Ferez, and F. Pincet, "Johnson-Kendall-Roberts theory applied to living cells," *2005 NSTI Nanotechnology Conference and Trade*

Show - NSTI Nanotech 2005 Technical Proceedings, vol. 028102, no. January, pp. 560–562, 2005.

[82] F. A. Meineke, C. S. Potten, and M. Loeffler, "Cell migration and organization in the intestinal crypt using a lattice-free model," *Cell Proliferation*, vol. 34, no. 4, pp. 253–266, 2001.

[83] T. Nagai and H. Honda, "A dynamic cell model for the formation of epithelial tissues," *Philosophical Magazine B*, vol. 81, no. 7, pp. 699–719, 2001.

[84] R. Farhadifar, J. C. Röper, B. Aigouy, S. Eaton, and F. Jülicher, "The influence of cell mechanics, cell-cell interactions, and proliferation on epithelial packing," *Current Biology*, vol. 17, no. 24, pp. 2095–2104, 2007.

[85] A. G. Fletcher, M. Osterfield, R. E. Baker, and S. Y. Shvartsman, "Vertex models of epithelial morphogenesis," *Biophysical Journal*, vol. 106, no. 11, pp. 2291–2304, 2014.

[86] H. Honda and T. Nagai, "Cell models lead to understanding of multi-cellular morphogenesis consisting of successive self-construction of cells," *Journal of Biochemistry*, vol. 157, no. 3, pp. 129–136, 2015.

[87] C. Tamulonis, M. Postma, H. Q. Marlow, C. R. Magie, J. de Jong, and J. Kaandorp, "A cell-based model of Nematostella vectensis gastrulation including bottle cell formation, invagination and zippering," *Developmental Biology*, vol. 351, no. 1, pp. 217–228, 2011.

[88] T. J. Newman, "Modeling multicellular systems using subcellular elements," *Mathematical Biosciences and Engineering*, vol. 2, no. 3, pp. 613–624, 2005.

[89] J. D. Murray, *Mathematical Biology*. Berlin: Springer-Verlag, 2nd, corr. ed., 1993.

[90] F. Brauer and C. Castillo-Chávez, *Mathematical Models in Population Biology and Epidemiology*. New York: Springer-Verlag, 2001.

[91] G. M. Knight, N. J. Dharan, G. J. Fox, N. Stennis, A. Zwerling, R. Khurana, and D. W. Dowdy, "Bridging the gap between evidence and policy for infectious diseases: how models can aid public health decision-making," *International Journal of Infectious Diseases*, vol. 42, pp. 17–23, 2016.

[92] D. Adam, "Modeling the pandemic: the simulations driving the world's response to COVID-19," *Nature*, vol. 580, pp. 316–318, 2020.

[93] P. F. Verhulst, "Notice sur la loi que la population suit dans son accroissement," *Correspondance Mathématique et Physique*, vol. 10, pp. 113–121, 1838.

[94] P. F. Verhulst, "Recherches mathématiques sur la loi d'accroissement de la population," *Nouveaux Mémoires de l'Académie Royale des Sciences et Belles-Lettres de Bruxelles*, vol. 18, pp. 1–38, 1845.

[95] M. Paczkowski, W. W. Kretzschmar, B. Markelc, S. K. Liu, L. A. Kunz-Schughart, A. L. Harris, M. Partridge, H. M. Byrne, and P. Kannan, "Reciprocal interactions between tumour cell populations enhance growth and reduce radiation sensitivity in prostate cancer," *Communications Biology*, vol. 4, no. 6, pp. 1–13, 2021.

[96] A. Okubo, P. K. Maini, M. H. Williamson, and J. D. Murray, "On the spatial spread of the grey squirrel in Britain," *Proceedings of the Royal Society of London. Series B, Biological Sciences*, vol. 238, no. 1291, pp. 113–125, 1989.

[97] X. Yu and M. Gen, *Introduction to Evolutionary Algorithms*. London: Springer-Verlag, 2010.

[98] J. Maynard Smith, *Evolution and the Theory of Games*. Cambridge: Cambridge University Press, 1982.

[99] J. Nash, "Non-cooperative games," *Annals of Mathematics*, vol. 54, no. 2, pp. 286–295, 1951.

[100] J. Maynard Smith and G. R. Price, "The logic of animal conflict," *Nature*, vol. 246, pp. 15–18, 1973.

[101] J. Maynard Smith, "A new theory of sexual investment," *Behavioral Ecology and Sociobiology*, vol. 7, pp. 247–251, 1980.

[102] R. Axelrod and W. D. Hamilton, "The evolution of cooperation," *Science*, vol. 211, no. 4489, pp. 1390–1396, 1981.

[103] S. Hummert, K. Bohl, D. Basanta, A. Deutsch, S. Werner, G. Theißen, A. Schroeter, and S. Schuster, "Evolutionary game theory: cells as players," *Molecular BioSystems*, vol. 10, pp. 3044–3065, 2014.

[104] K. Bohl, S. Hummert, S. Werner, D. Basanta, A. Deutsch, S. Schuster, G. Theißen, and A. Schroeter, "Evolutionary game theory: molecules as players," *Molecular BioSystems*, vol. 10, pp. 3066–3074, 2014.

[105] M. A. Nowak and R. M. May, "Evolutionary games and spatial chaos," *Nature*, vol. 359, pp. 826–829, 1992.

[106] E. Lieberman, C. Hauert, and M. A. Nowak, "Evolutionary dynamics on graphs," *Nature*, vol. 433, pp. 9–11, 2005.

[107] L. Sagan, "On the origin of mitosing cells," *Journal of Theoretical Biology*, vol. 14, pp. 255–274, 1967.

[108] S. Widder, R. J. Allen, T. Pfeiffer, T. P. Curtis, C. Wiuf, W. T. Sloan, O. X. Cordero, S. P. Brown, B. Momeni, W. Shou, H. Kettle, H. J. Flint, A. F. Haas, B. Laroche, J. U. Kreft, P. B. Rainey, S. Freilich, S. Schuster, K. Milferstedt, J. R. Van Der Meer, T. Grobkopf, J. Huisman, A. Free, C. Picioreanu, C. Quince, I. Klapper, S. Labarthe, B. F. Smets, H. Wang, O. S. Soyer, S. D. Allison, J. Chong, M. C. Lagomarsino, O. A. Croze, J. Hamelin, J. Harmand, R. Hoyle, T. T. Hwa, Q. Jin, D. R. Johnson, V. de Lorenzo, M. Mobilia, B. Murphy, F. Peaudecerf, J. I. Prosser, R. A. Quinn, M. Ralser, A. G. Smith, J. P. Steyer, N. Swainston, C. E. Tarnita, E. Trably, P. B. Warren, and P. Wilmes, "Challenges in microbial ecology: building predictive understanding of community function and dynamics," *The ISME Journal*, vol. 10, pp. 2557–2568, 2016.

[109] K. R. Foster and T. Bell, "Competition, not cooperation, dominates interactions among culturable microbial species," *Current Biology*, vol. 22, no. 19, pp. 1845–1850, 2012.

[110] C. D. Nadell, K. Drescher, and K. R. Foster, "Spatial structure, cooperation and competition in biofilms," *Nature Reviews Microbiology*, vol. 14, pp. 589–600, 2016.

[111] M. Ghoul and S. Mitri, "The ecology and evolution of microbial competition," *Trends in Microbiology*, vol. 24, no. 10, pp. 833–845, 2016.

[112] D. Tilman, "Resource competition between plankton algae: an experimental and theoretical approach," *Ecology*, vol. 58, no. 2, pp. 338–348, 1977.

[113] J. Monod, "The growth of bacterial cultures," *Annual Review of Microbiology*, vol. 3, pp. 371–394, 1949.

[114] M. A. Riley and M. A. Chavan, *Bacteriocins*. Berlin: Springer, 2007.

[115] M. A. Kohanski, D. J. Dwyer, and J. J. Collins, "How antibiotics kill bacteria: from targets to networks," *Nature Reviews Microbiology*, vol. 8, pp. 423–435, 2010.

[116] Y. Michel-Briand and C. Baysse, "The pyocins of Pseudomonas aeruginosa," *Biochimie*, vol. 84, pp. 499–510, 2002.

[117] M. Basler, "Type VI secretion system: secretion by a contractile nanomachine," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 370, p. 20150021, 2015.

[118] X.-Y. Li, T. Lachnit, S. Fraune, T. C. G. Bosch, A. Traulsen, and M. Sieber, "Temperate phages as self-replicating weapons in bacterial competition," *Journal of The Royal Society Interface*, vol. 14, p. 20170563, 2017.

[119] S. D. Allison, "Cheaters, diffusion and nutrients constrain decomposition by microbial enzymes in spatially structured environments," *Ecology Letters*, vol. 8, pp. 626–635, 2005.

[120] S. Schuster, J.-U. Kreft, N. Brenner, F. Wessely, G. Theißen, E. Ruppin, and A. Schroeter, "Cooperation and cheating in microbial exoenzyme production–theoretical analysis for biotechnological applications," *Biotechnology journal*, vol. 5, pp. 751–758, 2010.

[121] H. Bachmann, D. Molenaar, M. Kleerebezem, and J. E. van Hylckama Vlieg, "High local substrate availability stabilizes a cooperative trait," *The ISME Journal*, vol. 5, pp. 929–932, 2011.

[122] R. R. Stein, V. Bucci, N. C. Toussaint, C. G. Buffie, G. Rätsch, E. G. Pamer, C. Sander, and J. B. Xavier, "Ecological modeling from time-series inference: insight into dynamics and stability of intestinal microbiota," *PLoS Computational Biology*, vol. 9, no. 12, p. e1003388, 2013.

[123] K. Z. Coyte, J. Schluter, and K. R. Foster, "The ecology of the microbiome: networks, competition, and stability," *Science*, vol. 350, no. 6261, pp. 663–666, 2015.

[124] J. B. Xavier, C. Picioreanu, and M. C. M. van Loosdrecht, "A framework for multidimensional modelling of activity and structure of multispecies biofilms," *Environmental Microbiology*, vol. 7, no. 8, pp. 1085–1103, 2005.

[125] C. D. Nadell, K. R. Foster, and J. B. Xavier, "Emergence of spatial structure in cell groups and the evolution of cooperation," *PLoS Computational Biology*, vol. 6, no. 3, p. e1000716, 2010.

[126] J. B. Xavier and K. R. Foster, "Cooperation and conflict in microbial biofilms," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 3, pp. 876–881, 2007.

[127] H. M. Byrne, "Dissecting cancer through mathematics: from the cell to the animal model," *Nature Reviews Cancer*, vol. 10, pp. 221–230, 2010.

[128] P. M. Altrock, L. L. Liu, and F. Michor, "The mathematics of cancer: integrating quantitative models," *Nature Reviews Cancer*, vol. 15, pp. 730–745, 2015.

[129] A. Masoudi-Nejad, G. Bidkhori, S. Hosseini Ashtiani, A. Najafi, J. H. Bozorgmehr, and E. Wang, "Cancer systems biology and modeling: microscopic scale and multi-scale approaches," *Seminars in Cancer Biology*, vol. 30, pp. 60–69, 2015.

[130] M. W. Anderson, S. H. Reynolds, M. You, and R. M. Maronpot, "Role of proto-oncogene activation in carcinogenesis," *Environmental Health Perspectives*, vol. 98, pp. 13–24, 1992.

[131] L.-H. Wang, C.-F. Wu, N. Rajasekaran, and Y. K. Shin, "Loss of tumor suppressor gene function in human cancer: an overview," *Cellular Physiology and Biochemistry*, vol. 51, pp. 2647–2693, 2018.

[132] J. Bachmann, A. Raue, M. Schilling, V. Becker, J. Timmer, and U. Klingmüller, "Predictive mathematical models of cancer signalling pathways," *Journal of Internal Medicine*, vol. 271, pp. 155–165, 2011.

[133] G. B. Leenders and J. A. Tuszynski, "Stochastic and deterministic models of cellular p53 regulation," *Frontiers in Oncology*, vol. 3, pp. 8–23, 2013.

[134] G. Simoni, F. Reali, C. Priami, and L. Marchetti, "Stochastic simulation algorithms for computational systems biology: exact, approximate, and hybrid methods," *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, vol. 11, p. e1459, 2019.

[135] R. J. Orton, O. E. Sturm, V. Vyshemirsky, M. Calder, D. R. Gilbert, and W. Kolch, "Computational modelling of the receptor-tyrosine-kinase-activated MAPK pathway," *Biochemical Journal*, vol. 392, pp. 249–261, 2005.

[136] R. J. Orton, M. E. Adriaens, A. Gormand, O. E. Sturm, W. Kolch, and D. R. Gilbert, "Computational modelling of cancerous mutations in the EGFR/ERK signalling pathway," *BMC Systems Biology*, vol. 3, no. 100, 2009.

[137] O. Warburg, F. Wind, and E. Negelein, "The metabolism of tumors in the body," *The Journal of General Physiology*, vol. 8, no. 6, pp. 519–530, 1927.

[138] M. V. Liberti and J. W. Locasale, "The Warburg effect: how does it benefit cancer cells?," *Trends in Biochemical Sciences*, vol. 41, no. 3, pp. 211–218, 2016.

[139] R. A. Gatenby and E. T. Gawlinski, "A reaction-diffusion model of cancer invasion," *Cancer Research*, vol. 56, no. 24, pp. 5745–5753, 1996.

[140] A. A. Patel, E. T. Gawlinski, S. K. Lemieux, and R. A. Gatenby, "A cellular automaton model of early tumor growth and invasion: the effects of native tissue vascularity and increased anaerobic tumor metabolism," *Journal of Theoretical Biology*, vol. 213, pp. 315–331, 2001.

[141] A. R. A. Anderson, "A hybrid mathematical model of solid tumour invasion: the importance of cell adhesion," *Mathematical Medicine and Biology: A Journal of the IMA*, vol. 22, pp. 163–186, 2005.

[142] P. Gerlee and A. R. A. Anderson, "A hybrid cellular automaton model of clonal evolution in cancer: the emergence of the glycolytic phenotype," *Journal of Theoretical Biology*, vol. 250, pp. 705–722, 2008.

[143] R. A. Gatenby, K. Smallbone, P. K. Maini, F. Rose, J. Averill, R. B. Nagle, L. Worrall, and R. J. Gillies, "Cellular adaptations to hypoxia and acidosis during somatic evolution of breast cancer," *British Journal of Cancer*, vol. 97, pp. 646–653, 2007.

[144] J. Metzcar, Y. Wang, R. Heiland, and P. Macklin, "A review of cell-based computational modeling in cancer biology," *JCO Clinical Cancer Informatics*, pp. 1–13, 2019.

[145] P. C. Nowell, "The clonal evolution of tumor cell populations," *Science*, vol. 194, pp. 23–28, 1976.

[146] R. A. Gatenby, A. S. Silva, R. J. Gillies, and B. R. Frieden, "Adaptive therapy," *Cancer Research*, vol. 69, no. 11, pp. 4894–4903, 2009.

[147] A. S. Silva, Y. Kam, Z. P. Khin, S. E. Minton, R. J. Gillies, and R. A. Gatenby, "Evolutionary approaches to prolong progression-free survival in breast cancer," *Cancer Research*, vol. 72, no. 24, pp. 6362–6370, 2012.

[148] K. Bacevic, R. Noble, A. Soffar, O. Wael Ammar, B. Boszonyik, S. Prieto, C. Vincent, M. E. Hochberg, L. Krasinska, and D. Fisher, "Spatial competition constrains resistance to targeted cancer therapy," *Nature Communications*, vol. 8, pp. 1–15, 2017.

[149] J. A. Gallaher, P. M. Enriquez-Navas, K. A. Luddy, R. A. Gatenby, and A. R. Anderson, "Spatial heterogeneity and evolutionary dynamics modulate time to recurrence in continuous and adaptive cancer therapies," *Cancer Research*, vol. 78, no. 8, pp. 2127–2139, 2018.

[150] M. A. R. Strobl, J. West, Y. Viossat, M. Damaghi, M. Robertson-Tessi, J. S. Brown, R. A. Gatenby, P. K. Maini, and A. R. A. Anderson, "Turnover modulates the need for a cost of resistance in adaptive therapy," *Cancer Research*, vol. 81, pp. 1135–1147, 2021.

[151] S. Nishikawa, A. Takamatsu, S. Ohsawa, and T. Igaki, "Mathematical model for cell competition: predator-prey interactions at the interface between two groups of cells in monolayer tissue," *Journal of Theoretical Biology*, vol. 404, 2016.

[152] S. Nishikawa and A. Takamatsu, "Effects of cell death-induced proliferation on a cell competition system," *Mathematical Biosciences*, vol. 316, no. April, p. 108241, 2019.

[153] A. Tsuboi, S. Ohsawa, D. Umetsu, Y. Sando, E. Kuranaga, T. Igaki, and K. Fujimoto, "Competition for space is controlled by apoptosis-induced change of local epithelial topology," *Current Biology*, vol. 28, no. 13, 2018.

[154] S. W. Lee and Y. Morishita, "Possible roles of mechanical cell elimination intrinsic to growing tissues from the perspective of tissue growth efficiency and homeostasis," *PLoS Computational Biology*, vol. 13, no. 7, 2017.

[155] D. Gradeci, A. Bove, G. Charras, A. R. Lowe, and S. Banerjee, "Single-cell approaches to cell competition: high-throughput imaging, machine learning and simulations," *Seminars in Cancer Biology*, vol. 63, pp. 60–68, 2020.

[156] A. Bove, D. Gradeci, Y. Fujita, S. Banerjee, G. Charras, and A. R. Lowe, "Local cellular neighborhood controls proliferation in cell competition," *Molecular Biology of the Cell*, vol. 28, no. 23, 2017.

[157] A. G. Fletcher, J. M. Osborne, P. K. Maini, and D. J. Gavaghan, "Implementing vertex dynamics models of cell populations in biology within a consistent computational

framework," *Progress in Biophysics and Molecular Biology*, vol. 113, no. 2, pp. 299–326, 2013.

[158] J. Kursawe, *Quantitative approaches to investigating epithelial morphogenesis*. PhD thesis, University of Oxford, 2017.

[159] J. A. Smith and L. Martin, "Do cells cycle?," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 70, no. 4, pp. 1263–1267, 1973.

[160] G. R. Mirams, C. J. Arthurs, M. O. Bernabeu, R. Bordas, J. Cooper, A. Corrias, Y. Davit, S.-J. Dunn, A. G. Fletcher, D. G. Harvey, M. E. Marsh, J. M. Osborne, P. Pathmanathan, J. Pitt-Francis, J. Southern, N. Zemzemi, and D. J. Gavaghan, "Chaste: an open source C++ library for computational physiology and biology," *PLoS Computational Biology*, vol. 9, no. 3, p. e1002970, 2013.

[161] J. Pitt-Francis, M. O. Bernabeu, J. Cooper, A. Garny, L. Momtahan, J. Osborne, P. Pathmanathan, B. Rodriguez, J. P. Whiteley, and D. J. Gavaghan, "Chaste: using agile programming techniques to develop computational biology software," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 366, no. 1878, pp. 3111–3136, 2008.

[162] J. D. Hunter, "Matplotlib: a 2D graphics environment," *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90–95, 2007.

[163] M. L. Waskom, "seaborn: statistical data visualization," *Journal of Open Source Software*, vol. 6, no. 60, p. 3021, 2021.

[164] M. Norman, K. A. Wisniewska, K. Lawrenson, P. Garcia-Miranda, M. Tada, M. Kajita, H. Mano, S. Ishikawa, M. Ikegawa, T. Shimada, and Y. Fujita, "Loss of Scribble causes cell competition in mammalian cells," *Journal of Cell Science*, vol. 125, no. 1, pp. 59–66, 2012.

[165] S. Elmore, "Apoptosis: a review of programmed cell death," *Toxicologic Pathology*, vol. 35, no. 4, pp. 495–516, 2007.

[166] H. Steller, "Regulation of apoptosis in Drosophila," *Cell Death and Differentiation*, vol. 15, no. 7, pp. 1132–1138, 2008.

[167] M. S. D'Arcy, "Cell death: a review of the major forms of apoptosis, necrosis and autophagy," *Cell Biology International*, vol. 43, no. 6, pp. 582–592, 2019.

[168] Y. Fuchs and H. Steller, "Programmed cell death in animal development and disease," *Cell*, vol. 147, no. 4, pp. 742–758, 2011.

[169] S. L. Spencer and P. K. Sorger, "Measuring and modeling apoptosis in single cells," *Cell*, vol. 144, no. 6, pp. 926–939, 2011.

[170] K. Schleich and I. N. Lavrik, "Mathematical modeling of apoptosis," *Cell Communication and Signaling*, vol. 11, no. 1, p. 44, 2013.

[171] S. Legewie, N. Blüthgen, and H. Herzel, "Mathematical modeling identifies inhibitors of apoptosis as mediators of positive feedback and bistability," *PLoS Computational Biology*, vol. 2, no. 9, p. e120, 2006.

[172] M. Rehm, H. J. Huber, H. Dussmann, and J. H. M. Prehn, "Systems analysis of effector caspase activation and its control by X-linked inhibitor of apoptosis protein," *The EMBO Journal*, vol. 25, no. 18, pp. 4338–4349, 2006.

[173] M. Bentele, I. Lavrik, M. Ulrich, S. Stößer, D. W. Heermann, H. Kalthoff, P. H. Krammer, and R. Eils, "Mathematical modeling reveals threshold mechanism in CD95-induced apoptosis," *Journal of Cell Biology*, vol. 166, no. 6, pp. 839–851, 2004.

[174] I. N. Lavrik, A. Golks, D. Riess, M. Bentele, R. Eils, and P. H. Krammer, "Analysis of CD95 threshold signaling: triggering of CD95 (FAS/APO-1) at low concentrations primarily results in survival signaling," *Journal of Biological Chemistry*, vol. 282, no. 18, pp. 13664–13671, 2007.

[175] K. L. King and J. A. Cidlowski, "Cell cycle and apoptosis: common pathways to life and death," *Journal of Cellular Biochemistry*, vol. 58, no. 2, pp. 175–180, 1995.

[176] K. L. King and J. A. Cidlowski, "Cell cycle regulation and apoptosis," *Annual Review of Physiology*, vol. 60, no. 1, pp. 601–617, 1998.

[177] B. Pucci, M. Kasten, and A. Giordano, "Cell cycle and apoptosis," *Neoplasia*, vol. 2, no. 4, pp. 291–299, 2000.

[178] K. Vermeulen, Z. N. Berneman, and D. R. Van Bockstaele, "Cell cycle and apoptosis," *Cell Proliferation*, vol. 36, no. 3, pp. 165–175, 2003.

[179] M. M. Kasten and A. Giordano, "pRb and the Cdks in apoptosis and the cell cycle," *Cell Death and Differentiation*, vol. 5, no. 2, pp. 132–140, 1998.

[180] J. A. Pietenpol and Z. A. Stewart, "Cell cycle checkpoint signaling: cell cycle arrest versus apoptosis," *Toxicology*, vol. 181-182, pp. 475–481, 2002.

[181] L. M. Facchini and L. Z. Penn, "The molecular role of Myc in growth and transformation: recent discoveries lead to new insights," *The FASEB Journal*, vol. 12, no. 9, pp. 633–651, 1998.

[182] S. B. McMahon, "MYC and the control of apoptosis," *Cold Spring Harbor Perspectives in Medicine*, vol. 4, no. 7, p. a014407, 2014.

[183] G. Bretones, M. D. Delgado, and J. León, "Myc and cell cycle control," *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, vol. 1849, no. 5, pp. 506–516, 2015.

[184] N. E. Baker, "Emerging mechanisms of cell competition," *Nature Reviews Genetics*, vol. 21, no. 11, pp. 683–697, 2020.

[185] M. Eroglu and W. B. Derry, "Your neighbours matter - non-autonomous control of apoptosis in development and disease," *Cell Death and Differentiation*, vol. 23, no. 7, pp. 1110–1118, 2016.

[186] N. Ninov, D. A. Chiarelli, and E. Martín-Blanco, "Extrinsic and intrinsic mechanisms directing epithelial cell sheet replacement during Drosophila metamorphosis," *Development*, vol. 134, no. 2, pp. 367–379, 2007.

[187] Y. Nakajima, E. Kuranaga, K. Sugimura, A. Miyawaki, and M. Miura, "Nonautonomous apoptosis is triggered by local cell cycle progression during epithelial replacement in Drosophila," *Molecular and Cellular Biology*, vol. 31, no. 12, pp. 2499–2512, 2011.

[188] A. Pérez-Garijo, Y. Fuchs, and H. Steller, "Apoptotic cells can induce non-autonomous apoptosis through the TNF pathway," *eLife*, vol. 2, p. e01004, 2013.

[189] J. R. Huh, M. Guo, and B. A. Hay, "Compensatory proliferation induced by cell death in the Drosophila wing disc requires activity of the apical cell death caspase Dronc in a nonapoptotic role," *Current Biology*, vol. 14, no. 14, pp. 1262–1266, 2004.

[190] A. Pérez-Garijo, F. A. Martín, and G. Morata, "Caspase inhibition during apoptosis causes abnormal signalling and developmental aberrations in Drosophila," *Development*, vol. 131, no. 22, pp. 5591–5598, 2004.

[191] H. D. Ryoo, T. Gorenc, and H. Steller, "Apoptotic cells can induce compensatory cell proliferation through the JNK and the wingless signaling pathways," *Developmental Cell*, vol. 7, no. 4, pp. 491–501, 2004.

[192] L. Ballesteros-Arias, V. Saavedra, and G. Morata, "Cell competition may function either as tumour-suppressing or as tumour-stimulating factor in Drosophila," *Oncogene*, vol. 33, pp. 4377–4384, 2014.

[193] A. Pérez-Garijo and H. Steller, "Spreading the word: non-autonomous effects of apoptosis during development, regeneration and disease," *Development*, vol. 142, no. 19, pp. 3253–3262, 2015.

[194] Y. Kawamoto, Y. Nakajima, and E. Kuranaga, "Apoptosis in cellular society: communication between apoptotic cells and their neighbors," *International Journal of Molecular Sciences*, vol. 17, no. 12, p. 2144, 2016.

[195] 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*, vol. 10, no. 7, p. 1003616, 2014.

[196] R. Shields, "Further evidence for a random transition in the cell cycle," *Nature*, vol. 273, no. 5665, pp. 755–758, 1978.

[197] 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*, vol. 79, no. 12, pp. 2905–2928, 2017.

[198] O. Kempthorne and E. Pollak, "Concepts of fitness in Mendelian populations," *Genetics*, vol. 64, pp. 125–145, 1970.

[199] T. E. Oliphant, "SciPy: open source scientific tools for Python," *Computing in Science and Engineering*, vol. 9, pp. 10–20, 2007.

[200] P. A. Gagniuc, *Markov Chains: From Theory to Implementation and Experimentation*. USA, New Jersey: John Wiley & Sons, Inc., 2017.

[201] K. Siegrist, "Random." `https://www.randomservices.org/random/`, 2019. [Online; accessed 10 May 2020].

[202] O. Tange, "GNU Parallel: the command-line power tool," *;login: The USENIX Magazine*, vol. 36, no. 1, pp. 42–47, 2011.

[203] W. Li and N. E. Baker, "Engulfment is required for cell competition," *Cell*, vol. 129, no. 6, pp. 1215–1225, 2007.

[204] The Matplotlib development team, "Choosing colormaps in Matplotlib." `https://matplotlib.org/stable/tutorials/colors/colormaps.html`, 2021. [Online; accessed 5 October 2021].

[205] B. Hay, T. Wolff, and G. Rubin, "Expression of baculovirus P35 prevents cell death in Drosophila," *Development*, vol. 120, pp. 2121–2129, 1994.

[206] R. Levayer, B. Hauert, and E. Moreno, "Cell mixing induced by myc is required for competitive tissue invasion and destruction," *Nature*, vol. 524, no. 7566, pp. 476–480, 2015.

[207] L. Alpar, C. Bergantiños, and L. A. Johnston, "Spatially restricted regulation of Spätzle/Toll signaling during cell competition," *Developmental Cell*, vol. 46, no. 6, pp. 1–14, 2018.

[208] R. M. Neto-Silva, B. S. Wells, and L. A. Johnston, "Mechanisms of growth and homeostasis in the Drosophila wing," *Annual Review of Cell and Developmental Biology*, vol. 25, no. June, pp. 197–220, 2009.

[209] N. Senoo-Matsuda and L. A. Johnston, "Soluble factors mediate competitive and cooperative interactions between cells expressing different levels of Drosophila Myc," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 47, pp. 18543–18548, 2007.

[210] A. M. Smith, R. E. Baker, D. Kay, and P. K. Maini, "Incorporating chemical signalling factors into cell-based models of growing epithelial tissues," *Journal of Mathematical Biology*, vol. 65, no. 3, pp. 441–463, 2012.