

UNIVERSITY COLLEGE LONDON

FINAL YEAR PROJECT

**Modelling in vitro cell
competition: p53-/- and WT
cells in early gastrulation**

Claudia Alicia De Sousa Miranda Pérez

supervised by
Dr Rubén PÉREZ CARRASCO

12th March, 2020

Acknowledgements

First of all, I would like to thank Dr Pérez Carrasco, for being always such an accessible, well-prepared, and demanding supervisor. Not only he gave me the academic support I needed, but also believed in me and my potential coding skills from the very beginning, even when I did not do so myself. Thanks to my friends too, who served as an inspiration for my project and helped me all throughout. And finally, thanks to my family, but especially to my grandma, who does not understand a single word of English, and still asked me about the MSci project every single day.

Contents

1 Abstract	6
2 Introduction	6
3 One species system	12
3.1 Constant intrinsic growth	12
3.1.1 Continuous description	12
3.1.2 Agent based model	13
3.2 Competition	15
3.2.1 Continuous description (mean-field)	15
3.2.2 Agent based model (short-range interactions)	19
4 Two species system	26
4.1 Continuous description (mean-field)	26
4.2 Agent based model (short-range interactions)	28
5 Conclusions and discussion	36
6 Appendix	39

1 Abstract

Cells are constantly evaluating the fitness of their neighbours in order to prevent DNA-damaged cells from proliferating and propagating through the organism. They do so via cell competition. Given the important role that cell fitness plays in ageing and cancer, it is no wonder why cell competition is a topic of increasing interest in the biomedical community. The aim of this project is to understand cell competition in the context of early in vitro gastrulation. To do so, we construct two models. On the one hand, a deterministic continuous model, and on the other, a stochastic agent-based model. The agent based models allowed us to study the role of diffusion and interaction length unravelling the different dynamics of cell-cell competition.

2 Introduction

Cell competition

Developmental biology studies the growth and development of tissues and organs, including that characteristic of embryonic development (embryogenesis).

Embryogenesis requires a level of coordination between cells that enables proliferation, migration, and the adoption of different identities [1]. This process is carefully regulated by gene expression and will culminate in the formation of a multicellular living system.

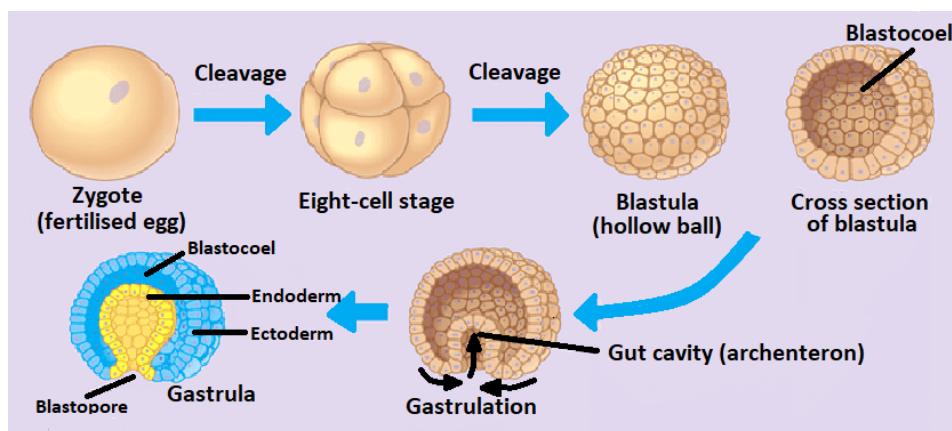


Figure 1: Early stages of embryogenesis. Picture adapted from [2].

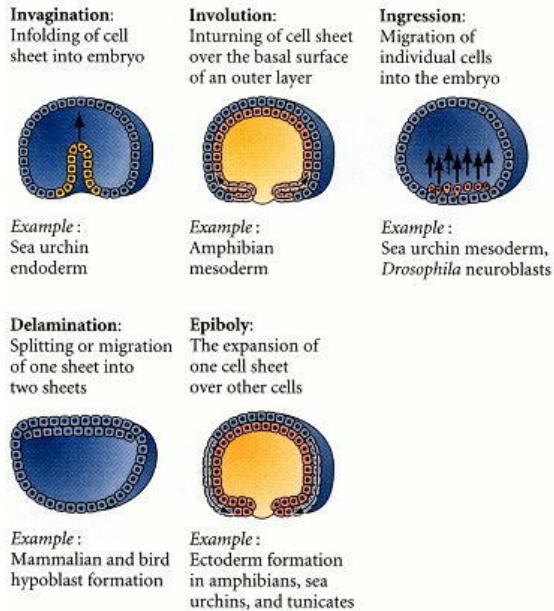


Figure 2: Types of cell movements during gastrulation. [1]

The starting point of embryogenesis is fecundation — the fusion of the egg and the sperm cell. Subsequent to fertilisation, the zygote is formed. This is subjected to several mitotic divisions [2] or cleavages (see Figure 1), which will eventually form the blastula, a yolk-filled cavity enclosed by a cell layer. Following the constitution of the single-layered blastula, the cells, called blastomeres, are reorganised into a multilayered configuration: the gastrula. By the end of gastrulation, cell differentiation and symmetry breaking take place, leading to organogenesis.

Gastrulation demands full organisation between ingression (cell migrations), and other movements that involve the entire embryo (see Figure 2 for full list). This means there is a strong spatial component inherent to the process of gastrulation. Furthermore, we know cell competition [3, 4] is one of the main mechanisms that drive morphogenesis [5], and hence, embryogenesis [6]. Cells are able to ‘sense’ fitness and eliminate those ones in their vicinity which are less fit or damaged [7] via cell competition.

Due to the complexity of the spatial interactions involved, the role that cell competition plays in the formation of the gastrula remains a conundrum. In order to comprehend how different competition mechanisms control gastrulation, a quantitative approach describing the cell population dynamics

is imperative. With this aim in mind, in this project, we will model and explore different aspects of cellular competition by looking at the cell population dynamics of a two-dimensional embryonic cell culture (see Figure 3A).

Recent research has revealed different genes essential in determining the fitness of neighbouring cells during cell competition. Particularly, the gene p53, which stops abnormal cells from proliferating. Gene p53 triggers apoptosis (programmed cell death) or produces cell-cycle arrest when a cell's DNA has been damaged, impeding cell proliferation.

Defective p53 could result in cancer [8] [9], and an excess of its protein may lead to earlier aging [10]. This is why it is of growing interest within the biomedical community to see how p53-gene mutants, in particular p53^{-/-} cells (cells lacking both p53 alleles), compete with its counterpart, the wild-type (WT) species (p53^{+/+}).

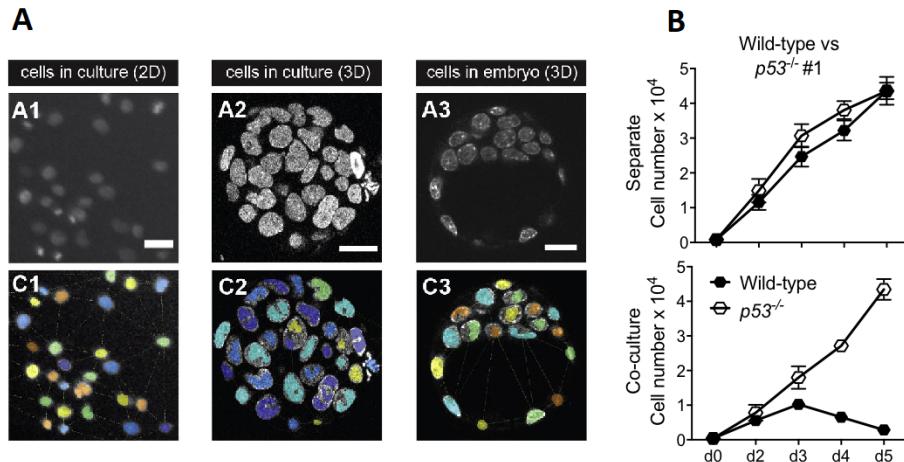


Figure 3: A) Figure is adapted from [11]. Images A1-A3 show fluorescent images of a 2D cell culture, a 3D cell culture, and an in vivo mouse embryo respectively. C1-C3 are graph coloured versions of A1-A3. . B) Experimental data courtesy of Tristan Rodriguez. The first graph shows the population growth of two different cultures of WT and p53^{-/-} cells, and the second, a co-culture of the two species.

For what has been stated in the previous paragraphs, one expects that in a co-culture of WT and p53^{-/-} cells, the higher fitness of p53^{-/-} cells will result in the eventual extinction of the WT species. This has been shown

to happen both experimentally (see Figure 3B) and theoretically [12]. However, in order to understand the relevance of this competition during early gastrulation, we need to understand the time scale of this process as well as other effects such as demographic stochasticity, one of the main sources of variability in ecological systems [13].

Stochasticity arises from the uncertainty and randomness associated to the interaction and cell fate decisions of individual cells [14]. There are two main sources of stochasticity: intrinsic noise (e.g. gene regulation, intra-cell variation of protein production, the diffusive nature of the intra-cell medium); and extrinsic noise (e.g. temperature variations, cell dynamics, environmental fluctuations). Mathematically, we can associate this stochasticity by assigning probabilities to cell dynamics, proliferation and apoptosis events. This allows us to build mathematical models that incorporate the demographical stochasticity observed during early gastrulation.

The aim of this project is to understand how competition during early gastrulation controls the time scale of the population growth. To do so, we will explore two different population models. On the one hand, a mean-field model describing the population dynamics as a coupled system of equations and, on the other hand, an agent-based model that incorporates spatial cell-to-cell interactions and cell motility.

Agent-based models

Agent-based modelling (ABM) is a computational modelling methodology used to describe complex systems as a collection of self-sufficient entities called agents. These entities interact with each other following a predetermined set of rules [15]. Examples of agents range from cells of the immune system to stock master investors [16].

ABM is a very useful tool to design, test, and analyse *in silico* experimental arrangements in a rule-based, discrete-event, and discrete-time fashion [17]. In addition, it is a great option when it comes to capturing the nonlinear and complex behaviour of biological systems like the one at hand [17].

Most mathematical and computation biomedical models are inductive models [17], meaning that, given a pattern of data, they infer the mechanisms via which the data was created. Nevertheless, ABMs aim to do the opposite: they start off with certain mechanisms or rules of behaviour and

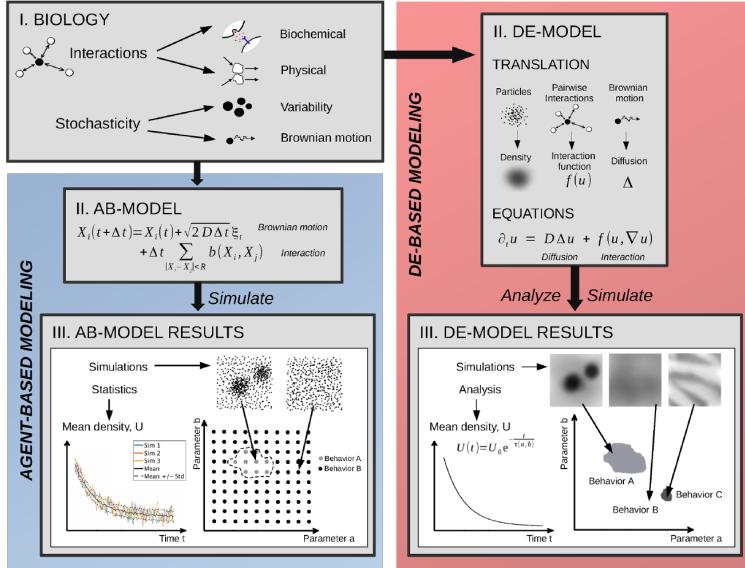


Figure 4: Reprinted from [18]. This diagram illustrates the main differences between agent-based (AB) modelling and differential equations (DE) modelling. It is a typical case scenario when we have interacting diffusive particles (e.g. cells, bacteria).

try to reconstruct the observed data patterns. Some of the advantages of implementing an agent-based model include:

1. ABMs easily integrate space. The agents' behaviours can be dependent on its immediate environment, which resembles the mechanism that drives cell competition. The more cells within a neighbourhood, the more strongly they will compete, and the more likely they are to die. This feature is very beneficial when looking at any population's dynamics given the spatial inherence of mathematical biology [19].
2. ABMs cover aggregated dynamics. If local conditions are changed, so will be the fates of each cell affected by them. This will lead of a higher-level system behaviour arising from the parallelism of the different behavioural trajectories of each agent. An example of this is collective cell migration [20]. A general leader in order is not needed in order to create distinct trajectory patterns.
3. ABMs can be designed to incorporate stochasticity. For example, we can estimate the proliferation rate for the cell population as a whole,

and then implement its respective probability function for each cell as an agent behavioral rule. By doing this, we can infer and test if our system follows deterministic rules, that could not be easily detected only with observations and experimental data.

4. Information can be added to ABMs effortlessly. Given their modular structure, we can conveniently introduce new cell types or change the behaviour rules of the existing ones without having to rebuild the whole model [17]. This allows us to easily change the course of our research to test different hypotheses if needed.
5. ABMs are useful to track emergent behaviours. ABMs portray accurately the nonintuitive nature of the entire system, which cannot normally be extracted from the rules of the agents alone. The different figures produced by birds in a flock are a good example of emergent behaviour [21].
6. ABMs can be built given incomplete knowledge. Testing very simple behavioural rules, e.g. each cell's death rate grows with number of neighbouring cells, can already give us a reasonable amount of information about the system at hand. Although a more detailed model will probably provide a better correlation with reality, given how little is our actual biological knowledge, we will never be able to predict the very exact system's behaviour.

For this project, given the computational resources available, we chose to use a hybrid modelling framework: on the one hand, ordinary differential equations (ODEs) are used in those cases where mean-field approximations are suitable to describe the system; and ABM, with the cells being the agents, when spatial effects are not negligible (e.g. when we take into account short cell-to-cell signalling) [17]. Modelling the concentration of the different cell species with partial differential equations (PDEs) would be a natural next step to bridge both limits (see Figure 5). However, this is beyond the scope of this project (see conclusions for further discussion).

3 One species system

3.1 Constant intrinsic growth

3.1.1 Continuous description

Mathematical models have played an essential role in building our understanding of population dynamics in biology [19]. One of the simplest models is the exponential growth model, sometimes also referred to as the Malthusian growth model [22]:

Let $N(t)$ be the number of cells at time t . Then, $\frac{dN}{dt}$ is the rate of change, and $\frac{1}{N} \frac{dN}{dt}$ is the per capita rate of change of this population [23]. In an isolated system with no migration, $N(t)$ can only vary as a result of proliferation and cell deaths. Assuming the per capita birth or proliferation rate b and per capita death or apoptosis rate d are positive constants, we can express this law of conservation of number of individuals as follows:

$$\frac{1}{N} \frac{dN}{dt} = b - d. \quad (1)$$

Equation (1) is normally rewritten as

$$\frac{dN}{dt} = (b - d) N \equiv rN, \quad (2)$$

where r is the difference between the per capita birth and death rates and is known as intrinsic growth rate.

Given $N(0) = N_0$ ($N_0 \in \mathbb{N}$) as the initial number of cells, we can easily solve Equation (2), which leads to

$$N(t) = N_0 e^{rt}.$$

This solution describes a non-stable population. If $r > 0$, the population will explode, meaning $\lim_{t \rightarrow \infty} N(t) = \infty$. For $r < 0$, the population will become extinct at large times, i.e., $\lim_{t \rightarrow \infty} N(t) = 0$. If $N_0 = 0$, $N(t)$ will remain at 0 for all times regardless of the sign of r , since $N^* = 0$ is an equilibrium point.

Exponential growth is an unrealistic model with several limitations:

1. It does not account for environmental resources availability or intraspecific competition or cooperation, which can have an impact on the

variation of number of individuals. Consequently, in this model, the population growth is unhindered, which is unrealistic. [23]

2. The proliferation and death rates, and hence the intrinsic growth rate, are constant in time. Therefore, cells do not age or differentiate during our analysis.
3. The number of individuals $N(t)$ ought to be discrete. Nevertheless, in order to be able to differentiate this quantity, we consider it to be a continuous variable.
4. It ignores stochastic effects.

3.1.2 Agent based model

Method

The cells in our agent-based model are represented as point-like particles in a two-dimensional continuous region. The position of each particle can be described as $\vec{x} = (x, y) \in \mathbb{R}^2$ where they can move, proliferate and die. At each discrete time increment Δt the particles can do any of the following actions:

1. Proliferate at a constant per capita birth rate b . The new cell is located at a distance $2R$ from the mother cell, where $R = 0.1$ is an arbitrary radius for the cells, at an angle $\theta = 2\pi \mathcal{U}(0, 1)$, where $\mathcal{U}(0, 1)$ is the uniform probability distribution. Therefore:

$$x_{\text{daughter}}(t_{\text{proliferation}}) = x_{\text{mother}}(t_{\text{proliferation}}) + 2R \cos \theta$$

$$y_{\text{daughter}}(t_{\text{proliferation}}) = y_{\text{mother}}(t_{\text{proliferation}}) + 2R \sin \theta$$

2. Die at a constant per capita apoptosis rate d .
3. Perform a two-dimensional Brownian motion (see Figure 6B) following the Stochastic differential equation

$$d\vec{x} = \sqrt{2D\Delta t} d\vec{W}$$

where D is the diffusion coefficient of the cell in the medium and \vec{W} is a Wiener Process. This is simulated using the Euler-Maruyama method [24],

$$x_i(t + \Delta t) = x_i(t) + \sqrt{2D\Delta t} \mathcal{N}(0, 1)$$

$$y_i(t + \Delta t) = y_i(t) + \sqrt{2D\Delta t} \mathcal{N}(0, 1)$$

where i is an index representing cell i and $\mathcal{N}(0, 1)$ is a random Gaussian number of zero mean and standard deviation equal to 1.

In order to simulate a typical *in vitro* experiment with a number of cells that can be greater than 10^5 , we are going to study a smaller square region of the dish such that $\vec{x} = \{(x, y) \mid x \in [0, L], y \in [0, L]\}$, i.e. $\vec{x} \in [0, L] \times [0, L]$, with periodic toroidal conditions e.g. when a cell exits the sector from the top, it will be introduced in the box from the bottom (see Figure 5A). We will choose $L = 50$ in arbitrary spatial units without loss of generality.

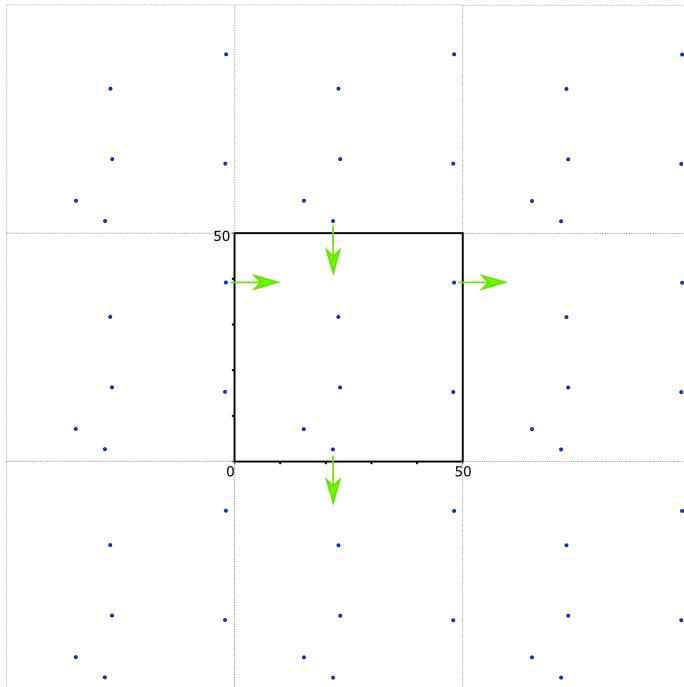


Figure 5: The diagram above illustrates the periodic boundary conditions. It shows two examples: a cell leaving the box by crossing the right-hand side boundary, and a cell leaving the box through its bottom. Both of them would re-appear in the opposite boundary.

To simulate the linear birth-death stochastic process compatible with the exponential process, we implemented stochastic acceptance or rejection of proliferation and death events by using a Monte Carlo method [25]. For each of the cells, we generate a random number X from a uniform probability distribution i.e. $X \sim \mathcal{U}(0, 1)$. If X is greater than the corresponding per capita rate (b or d), then the event takes place. Otherwise, the event is rejected.

In order to reduce the computational cost, we had to give up real-time information about the system, and “coarse-grain” our problem. This means data such as the number of cells or their positions was only recorded every certain number of timesteps.

Notice that the diffusive nature of the cells will not be determinant for the analysis of the population dynamics in this case. However, it will play an important role later on when cell-cell signalling is added to the model.

Results

In order to compare the stochastic and deterministic models for exponential growth, we performed different realizations of the stochastic agent based model with fixed values of b , d and N_0 . After that, we calculated the average number of cells at each timestep to give us an estimate of the expected number of cells at time t , $\langle N(t) \rangle$ (see Figure 6).

As expected from a linear birth-death process, the average of the stochastic model $\langle N(t) \rangle$ coincides with the deterministic description (see Figure 6A).

In addition, it is important to note that varying the value of the diffusion coefficient D for fixed b , d and N_0 does not affect our final results. We still recover the deterministic model. This is due to the absence of spatial interactions within the cells in this model. A more detailed analysis of the probability distribution of having N cells at time t can be carried out by solving the Master Equation of the birth-death process [26]. Nevertheless, we did not discuss this case further because of its lack of biological interest.

3.2 Competition

3.2.1 Continuous description (mean-field)

The simple exponential growth model can be adequate to describe cell population dynamics at early times. However, in a more realistic set-up, the population size of cells is limited and smaller than the intrinsic growth [27] due to competition for space and nutrients or direct cell-to-cell competition.

To accommodate restricted growth, Verhulst [28] proposed the logistic growth equation in 1838. It assumes there is an upper bound on the population size, whose notation Pearl and Reed popularised in the 1920s [27], called the carrying capacity (K). It can be seen as a saturation level for the

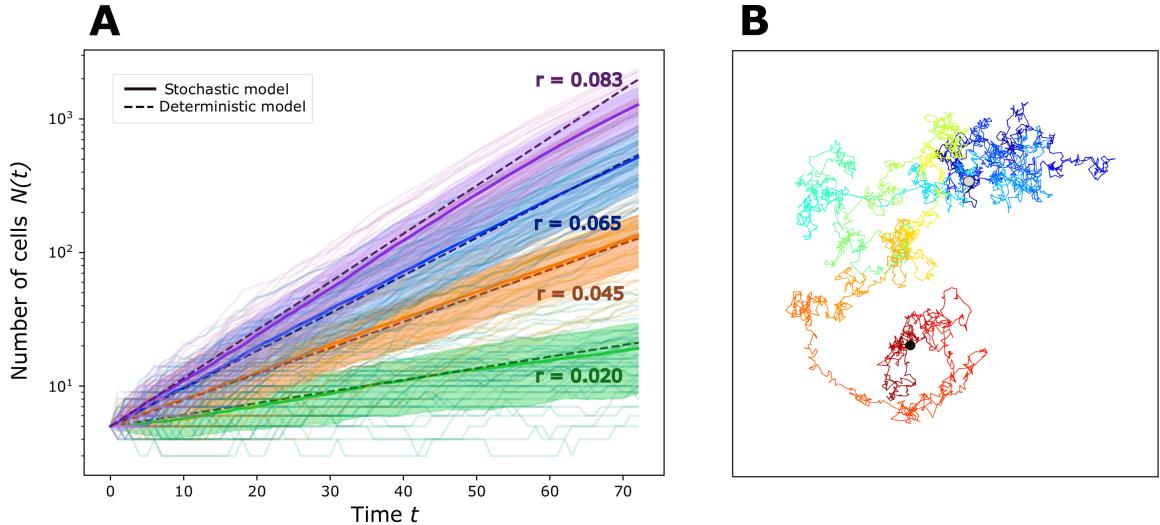


Figure 6: A) Evolution of the number of cells in time for different values of the intrinsic growth. For each condition, 100 stochastic trajectories (thin lines), their average in time (solid line) and the deterministic prediction (dashed line), are compared. B) Example of the trail in time (lines) of a cell of the agent-based model following a Brownian motion in arbitrary spatial units. The initial and final positions are indicated with a white and black circle respectively.

number of individuals an environment can sustain.

The logistic Verhulst-Pearl equation is normally expressed in the following form:

$$\frac{dN}{dt} = rN \left(1 - \frac{N}{K}\right). \quad (3)$$

This representation is equivalent to considering a linear correction of the intrinsic growth rate of the population that decays proportionally to the number of cells in the dish.

Equation (3) has solution

$$N(t) = \frac{KN_0}{(K - N_0) e^{-rt} + N_0}, \quad (4)$$

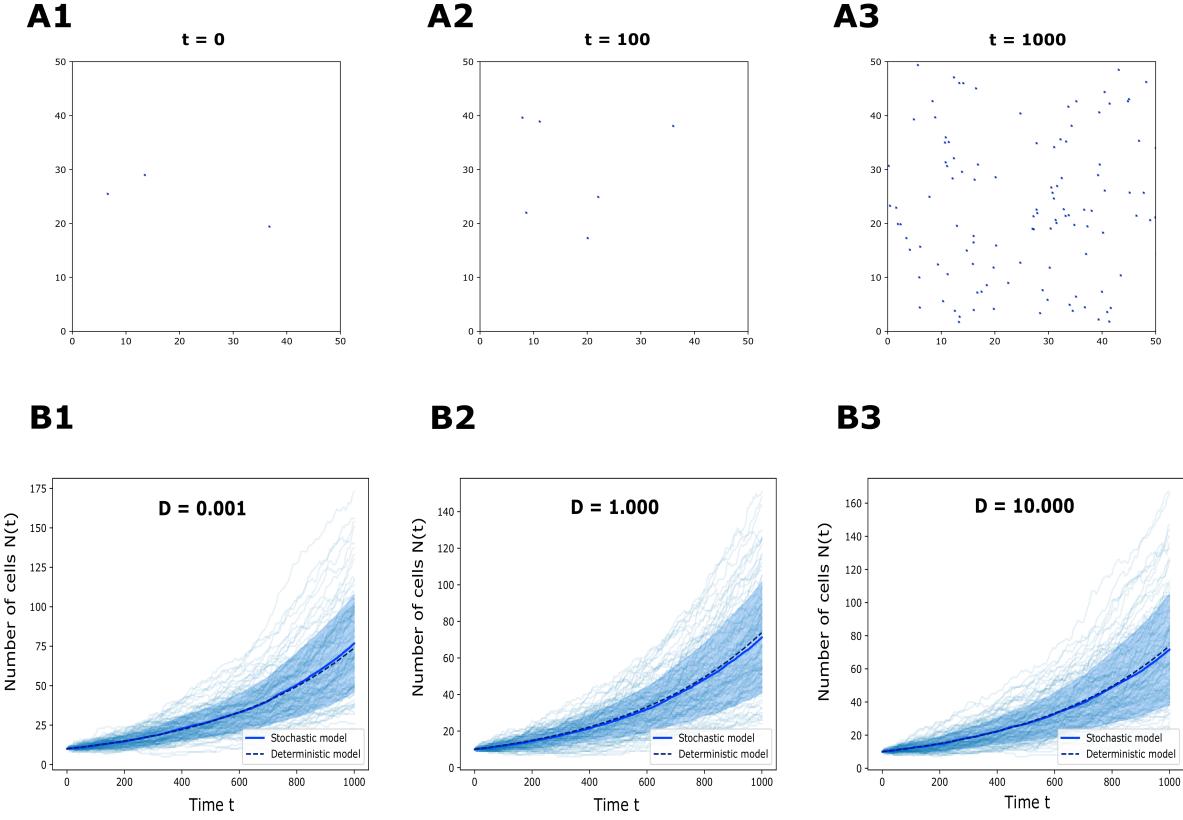


Figure 7: A1-A3 are snapshots at $t = 0$, $t = 100$ and $t = 1000$ respectively, of a stochastic exponential growth simulation with $b = 0.0050$ and $d = 0.0010$. B1-B3 compare the stochastic and deterministic models for $b = 0.0030$ and $d = 0.0010$ and initial number of cells $N_0 = 10$ in increasing diffusion D . The shaded part represents the standard deviation of the stochastic mean for $N(t)$.

where N_0 is the number of cells at $t = 0$.

Only the case where $r > 0$ is relevant in a proliferating cell culture, so we will restrict our analysis of the logistic growth equation to the $r > 0$ case from now on.

These are some key points that characterise logistic growth:

1. As time increases, the population size approaches its carrying capacity, i.e., $\lim_{t \rightarrow \infty} N(t) = K$.

2. Equation (3) has two equilibrium points: $N^* = 0$ and $N^* = K$.

- In the neighbourhood of $N^* = 0$, $\frac{N^2}{K}$ is small in comparison to N , so we can write

$$\frac{dN}{dt} \approx rN.$$

This means that, for $r > 0$, small perturbations about the origin grow exponentially [23]. Thus, we have that $N^* = 0$ is an unstable equilibrium point.

- In order to understand the behaviour of $N(t)$ in the vicinities of $N^* = K$, let us introduce a change of variable, $n = N - K$ [23]. This new variable measures the deviation of N from K . Rewriting Equation (3) in terms of n gives us

$$\frac{dn}{dt} = -r n - \frac{r}{K}n^2,$$

which can be approximated to

$$\frac{dn}{dt} \approx -r n,$$

for small n , i.e., as $N \rightarrow K$. Hence, small perturbations about $N^* = K$ decay exponentially, and we conclude $N^* = K$ is asymptotically stable.

3. Depending on the initial number of cells, N_0 , solutions to the logistic equation (with $r > 0$) will have different shapes. There are three main regimes:

- Solutions with $N_0 > K$ decay exponentially towards $N = K$.
- If $N_0 = K$, then $N(t) = K \forall t$. Similarly, if $N_0 = 0$, then $N(t) = 0 \forall t$.
- Solutions to the logistic equation with $N_0 < K$ have sigmoidal shape, and are fundamentally a combination of exponential growth close to zero and exponential decay close to K , as shown in 2. This means this sort of solutions are concave (upwards) just above $N^* = 0$ and convex (upwards) just below $N^* = K$ — there must be an inflection point lying between the origin and the carrying capacity. Setting $\frac{d^2N}{dt^2} = 0$, gives us the inflection point of equation (4). That is $N = K/2$.

All of the above is equivalent to considering birth and death rates do not longer remain constant as competition is introduced in the system — they will be dependent on N , and therefore, on time.

$$\frac{dN}{dt} = [\beta(N) - \delta(N)] N, \quad (5)$$

where $\beta(N)$ is the proliferation rate and $\delta(N)$, the apoptosis rate.

Experimental data suggests $\beta(N)$ can be taken to be constant, i.e., $\beta(N) \simeq b$. Therefore, by taking a linear approximation for $\delta(N)$, $\delta(N) \simeq d_0 + d_1 N$, with $d_0, d_1 \in \mathbb{R}^+/\{0\}$ constants, and plugging it in Equation (5), we arrive at

$$\frac{dN}{dt} = bN - (d_0 + d_1 N) N. \quad (6)$$

We can recover Equation (3), the logistic growth equation, from Equation (6) by letting $b - d_0 = r$ and $d_1 = r/K$.

Equation (6) still describes a mean-field interaction where the apoptosis rate depends on the total number of cells. For instance, when cells move at great speeds so there are no spatial correlations, or when the interactions are long range (e.g. through a diffusible molecule or hormone).

3.2.2 Agent based model (short-range interactions)

Method

Similarly to the continuous discrete model, we can implement competition for resources in the agent-based model by using the same rules displayed in 3.1.2, except for the birth and death rates, which will be allowed to depend with the population density.

We can rewrite Equation (6) in terms of the area density of cells within the square region, ρ_L :

$$\begin{aligned} \frac{dN}{dt} &= bN - \left(d_0 + d_1 \frac{N}{L^2} L^2 \right) N \\ \frac{dN}{dt} &= bN - (d_0 + d_1 \rho_L L^2) N \end{aligned}$$

,

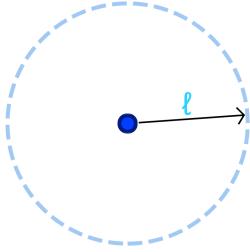


Figure 8: A cell with a radial range of interaction of ℓ . Neighbouring cells that move within the inner part of the circumference will contribute towards this cell's apoptosis rate.

where $\rho_L = \frac{N}{L^2}$.

This means $\delta(t) = d_0 + d_1 \rho_L L^2$, which is the same for all cells (mean-field approximation).

However, in a more realistic situation, each cell would have its own apoptosis rate, dependent on the number of neighbouring cells. Therefore, we introduce short-range interactions by assuming each cell only interacts with other cells within a circumference with center itself and radius ℓ with $0 < \ell < L$, i.e., cell-cell signalling distance is fixed to ℓ (see Figure that I still need to add).

The apoptosis rate for cell i would be

$$\delta_i(t) = d_0 + d_1 \rho_\ell^i L^2, \quad (7)$$

where $\rho_\ell^i = \frac{N_\ell^i}{\pi \ell^2}$, being N_ℓ^i the number of cells within a radial distance ℓ from cell i .

In summary, for the competition agent-based model, cells can perform the following actions:

1. Proliferate at a constant per capita birth rate b .
2. Die at an apoptosis rate $\delta(t)$ dependent on the density of cells in their vicinity, and which is computed as shown in Equation (7).

3. Perform a two-dimensional Brownian motion with diffusion coefficient D , in the same fashion as it was presented in section 3.1.2.

As it was also mentioned in section 3.1.2, real-time information about the system was given up in order to make the computations more time efficient — data was recorded every few timesteps, rather than at every single one. This is the case for each cell’s apoptosis rates, which were only updated after a specific number of iterations, rather than at each timestep, in some cases.

Results

As in the exponential growth model, we start our analysis by comparing the average number of cells at time t , $\langle N(t) \rangle$ for the mean-field interaction ($\ell \rightarrow \infty$) resulting from the agent based model.

After performing $n \gg 1$ stochastic experiments with mean-field competition, we see that varying the diffusion does not affect the final outcome, in analogy to section 3.1.2, as we expected. Once more, since cells have fixed rates of proliferation, apoptosis, and competition, this means they can ‘sense’ all cells in the plate and therefore, local spatial effects are negligible — we recover the deterministic scene.

However, when we turn on cell-cell signalling, the diffusive behaviour of cells plays an important role on the outcome of the model. There are two main regimes:

1. $\ell \ll L$.

For low diffusion (see A1 in Figure 9), cells eventually get extinct, which does not agree with the deterministic prediction. This is due to the short mean displacement of cells. Once cells divide, the daughter cells will stay within the interaction range of the mother cells for longer than they would for greater diffusions, and will induce a greater apoptosis rate. This will result in an increase of the effective apoptosis rate for both mother and daughter cells after division. For higher diffusions, nonetheless, the short-range model resembles more the deterministic prediction. Cells ‘notice’ more cells on the plate, but for shorter times, which lets the population increase steadily.

In Figure 9 (A2), the number of cells does not attain the deterministic saturation level, but in (A3), it surpasses it. Increasing the diffusion in

this regime speeds up the growth of the population.

2. $\ell \sim L$.

The range of the interactions is comparable to the size of the square quadrant we are analysing. In this limit, the mean-field model is a good approximation to describe the problem. B2 in Figure 9 is the case that resembles the deterministic prediction the most with a typical value of D . However, for low D , the theoretical prediction for the carrying capacity is surpassed by the stochastic model.

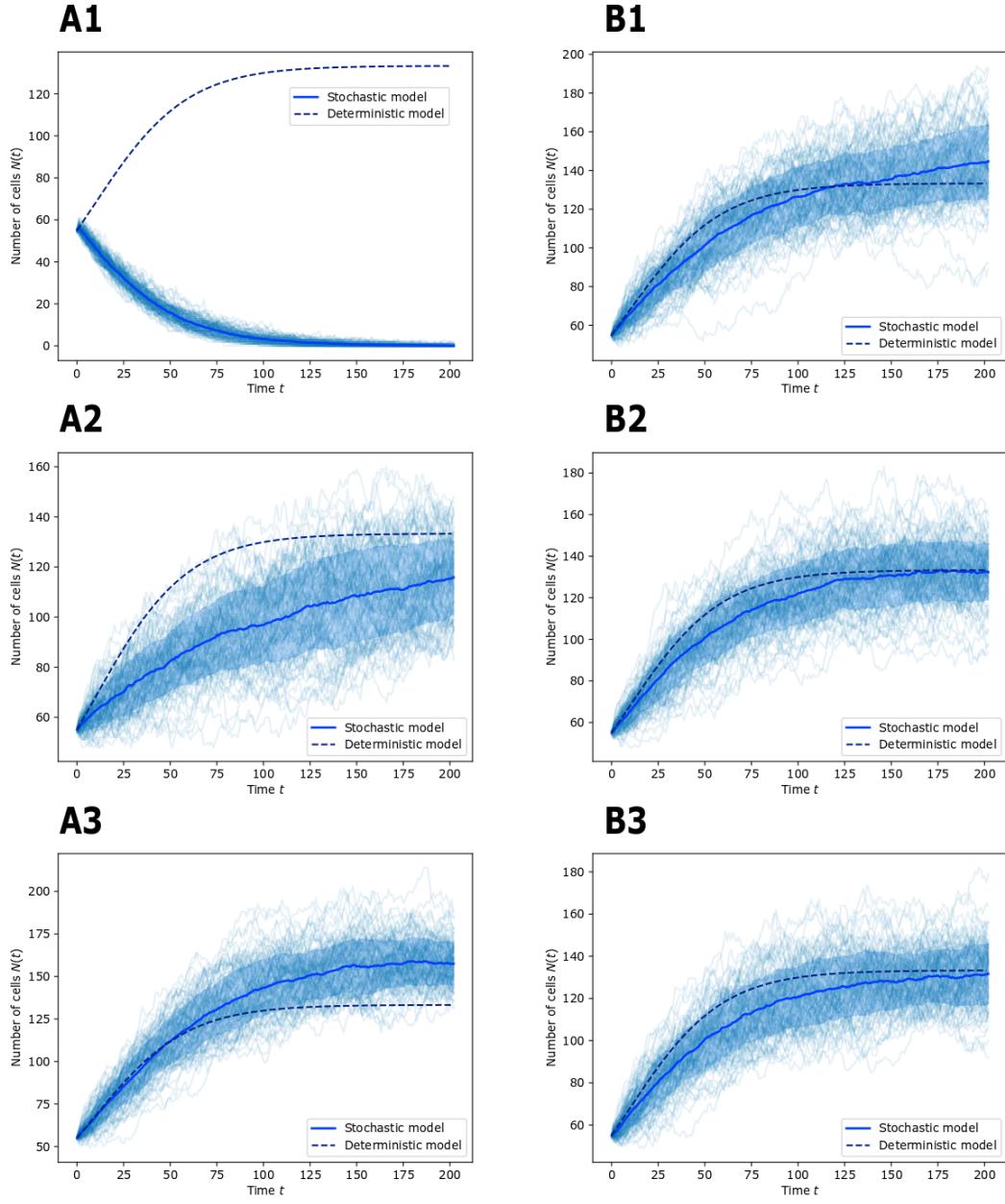


Figure 9: Graphs A1-A3 are plots of $N(t)$ vs t of the close-range interaction stochastic model for $\ell = 0.5$ with three different values of D and coefficients $b = 0.040$ and $d = 0.010$. B1-3 show the same model for $\ell = 25$.

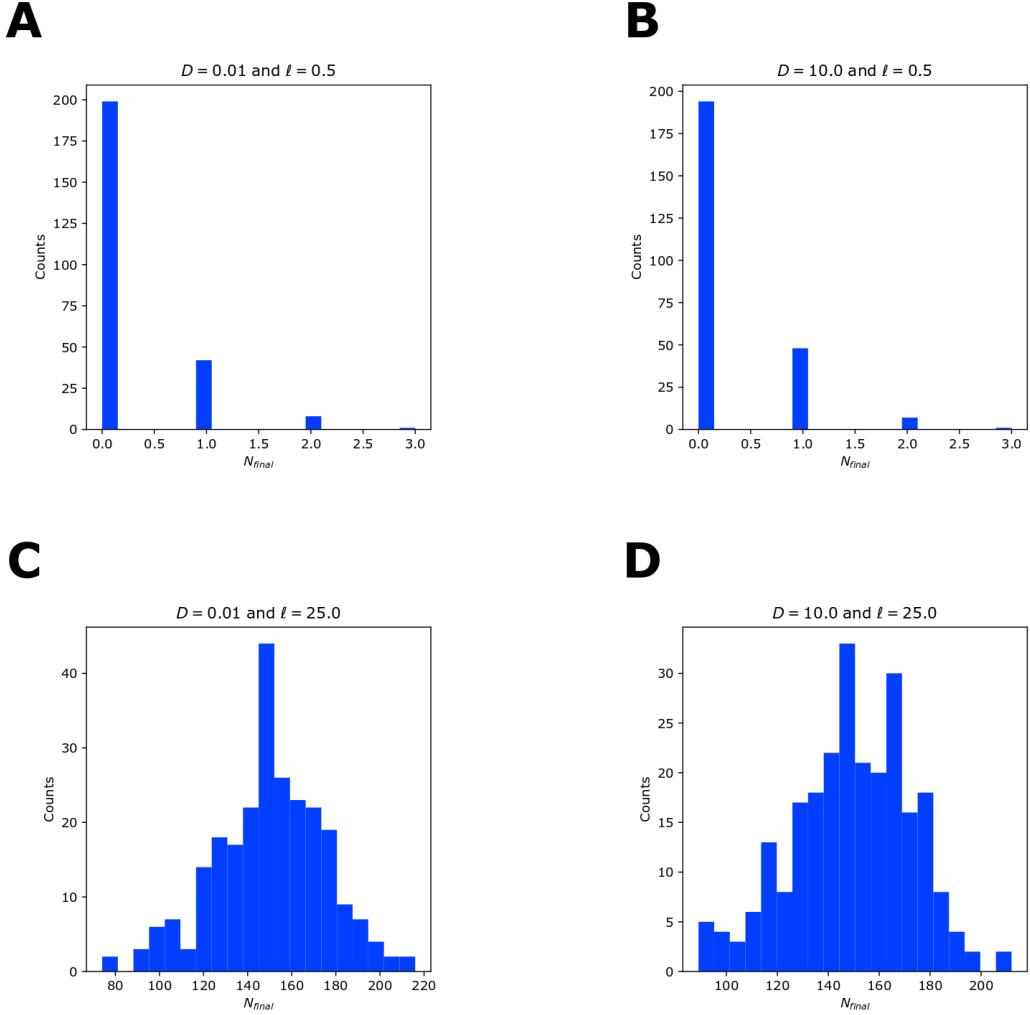


Figure 10: Histograms of the final number of cells, N_{final} , for 250 stochastic trajectories of $N(t)$. The values of D and ℓ for each of them are: (A) $D = 0.01$, $\ell = 0.5$; (B) $D = 10.000$, $\ell = 0.5$; (C) $D = 0.01$, $\ell = 25.0$; (D) $D = 10.000$, $\ell = 25.0$.

We performed a statistical analysis of the final number of cells N_{final} of the stochastic trajectories. We looked at the frequency of the different final numbers of cells. However, we did not find any significant changes in the mean of the histogram (see Figure 10 (C) and (D)).

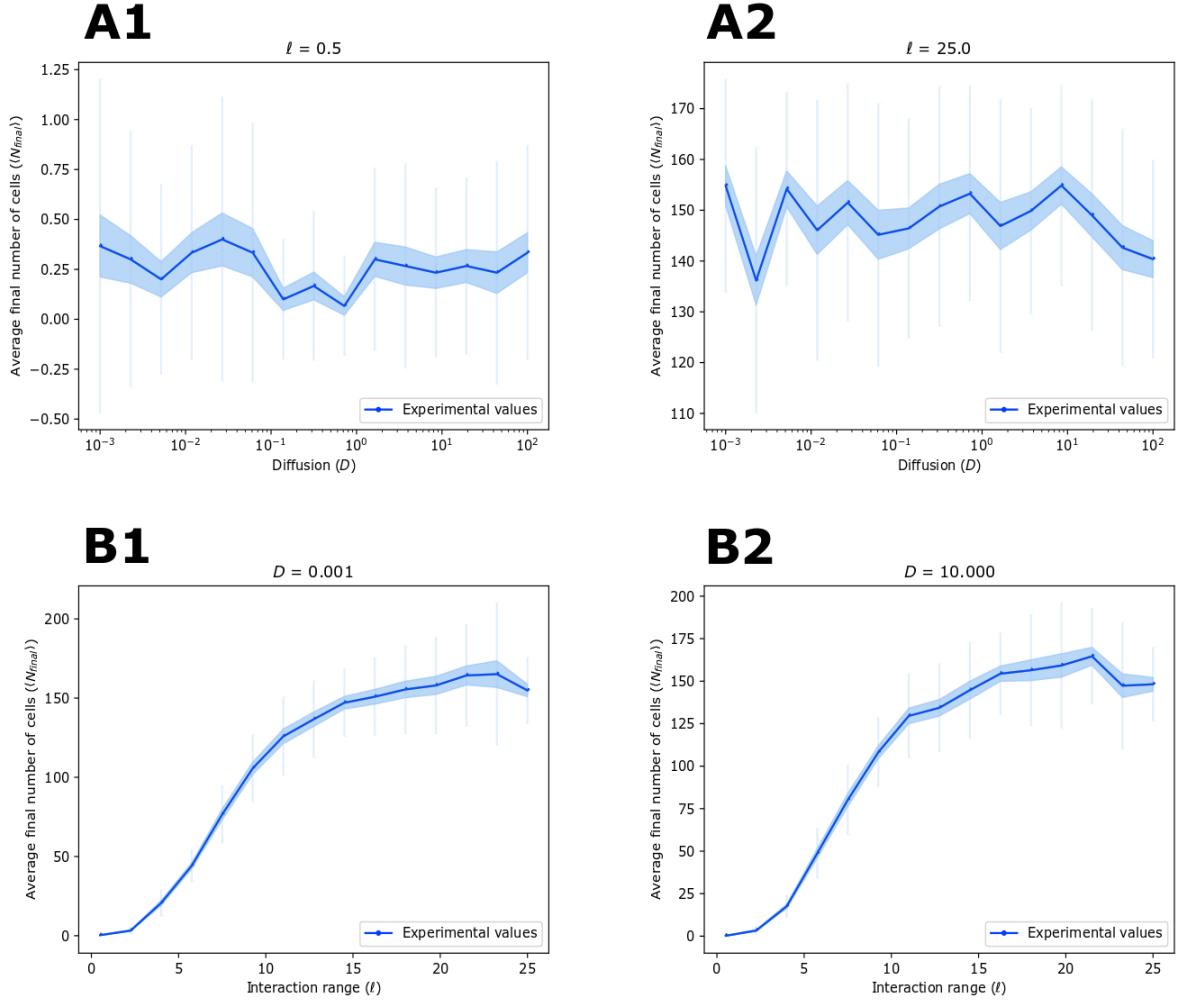


Figure 11: Plots of $\langle N_{final} \rangle$ vs D with a fixed value of ℓ : (A1) $\ell = 0.5$ and (A2) $\ell = 25.0$. Plots of $\langle N_{final} \rangle$ vs ℓ with a fixed value of D : (B1) $D = 0.001$ and (B2) $D = 10.000$.

In addition, we also looked at the final number of cells average when fixing the diffusion and varying the interaction range, and viceversa. Once more, we did not find any significant features or different conclusions from the ones we extracted from the plots in Figure 9. We can see that in Figure 11 (B1) and (B2), the average final number of cells on the plate is similar for the same values of ℓ even though we are at different regimes for D . This means that, to analyse the system further, we would have to look at the times where the

population's saturation level is attained, rather than at what the saturation level is itself.

4 Two species system

4.1 Continuous description (mean-field)

We start off with two decoupled populations of cells that grow logistically when bred in separate cultures. Let $N(t)$ be the number of WT cells at time t , and $P(t)$, the number of p53-/- cells. Then, we have

$$\frac{dN}{dt} = b_N N - (d_N + a_N N) N; \quad (8)$$

$$\frac{dP}{dt} = b_P P - (d_P + a_P P) P; \quad (9)$$

where b_N and b_P are the proliferation rates for WT and p53-/- cells respectively, and $\delta_N(N) = d_N + a_N N$ and $\delta_P(P) = d_P + a_P P$ are their respective apoptosis rates.

Now, let us consider the case scenario where both species are bred in the same plate — a co-culture of WT and p53-/- cells. We assume the presence of p53 mutants causes a decrease in the per capita growth of the WT species due to competition for resources, and vice-versa. This effect is called interspecific competition and it occurs amongst individuals of different species, as opposed to intraspecific competition, presented in section 3.2, and reintroduced in this section in Equations (8) and (9) in the form of the a_N and a_P coefficients.

In order to track the repercussions of this effect on the system's population dynamics, we introduce a pair of interspecific competition coefficients, c_N and c_P , which measure how strongly p53 mutants compete with WT cells, and how strongly WT cells compete with p53-/-, respectively. Thus, we arrive at the Lotka-Volterra competitive equations [23] (in phenomenological form)

$$\frac{dN}{dt} = b_N N - (d_N + a_N N + c_N P) N; \quad (10)$$

$$\frac{dP}{dt} = b_P P - (d_P + a_P P + c_P N) P; \quad (11)$$

where b_N and b_P are the proliferation rates of WT and p53-/ cells, and $\delta_N(N, P) = d_N + a_N N + c_N P$ and $\delta_P(N, P) = d_P + a_P P + c_P N$ are their respective apoptosis rates when in co-culture, with a_N and a_P being the inter-specific competition coefficients, and c_N and c_P , the intraspecific competition ones, and $b_N, b_P, d_N, d_P, a_N, a_P, c_N, c_P$ all $\in \mathbb{R}^+/\{0\}$.

Experiments show that the p53-/ mutants population drive WT cells to extinction [12], meaning there is no asymptotically stable coexistence state (see Figure 3B). Because of behaviour the p53 -/- cells exhibit, we expect that WT and p53-/ cells compete strongly with the other species than with their own species, and we will assume that WT cells compete more or equally strongly for resources amongst each other than p53 mutants. This leads to the following inequality for the competition coefficients:

$$c_P < a_P \leq a_N < c_N \quad (12)$$

Mathematically, we can address coexistence by looking at the point where the non-trivial N and P nullclines cross in the phase portrait of the system. If this point is within the first quadrant, i.e., if N and P are positive, then we could have either a stable steady state (coexistence) or unstable steady state. This crossing point being in the first quadrant is, therefore, a necessary condition for coexistence (but not a sufficient one).

The relevant system of equations that defines the interior state $(N, P) = (N^*, P^*)$ is

$$(b_N - d_N) = a_N N^* + c_N P^* \quad (13)$$

$$(b_P - d_P) = a_P P^* + c_P N^*. \quad (14)$$

Solving for P in Equation (12) gives

$$P^* = \frac{(b_N - d_N) - a_N N^*}{c_N}.$$

Plugging this last expression for P^* into Equation (13) leads to the following expression for N^* :

$$N^* = \frac{(b_P - d_P) - \frac{a_P}{a_N}(b_N - d_N)}{\frac{-a_P a_N}{c_N} + c_P}.$$

Assuming $b_N - d_N = b_P - d_P = r > 0$ leads to

$$N^* = \frac{r}{c_P} \left(\frac{1 - \frac{a_P}{c_N}}{1 - \frac{a_P a_N}{c_N c_P}} \right). \quad (15)$$

Since $r/c_P > 0$, for non-coexistence we need what is inside of the brackets in (15) to be negative. This can be achieved in two different ways:

- (a) $1 - \frac{a_P}{c_N} > 0$ and $1 - \frac{a_P a_N}{c_N c_P} < 0$, which implies $a_N < c_P$.
- (b) $1 - \frac{a_P}{c_N} < 0$ and $1 - \frac{a_P a_N}{c_N c_P} > 0$, which implies $c_P < a_N$.

Therefore, since (b) is the only case in agreement with (12), we will concentrate on in silico experiments with these coefficient constraints in our agent-based model.

4.2 Agent based model (short-range interactions)

Method

For the two species agent-based model, we will inherit the same behavioural and spatial rules displayed in sections 3.1.2 and 3.2.2. Nevertheless, now we have to take into account interspecific competition, i.e., coupling within the two different cell types — WT cells and p53-/- mutants.

This means that now each of the cells' apoptosis rate not only depends on the number of their same species' neighbouring cells, but also in the density of the other cell type within a fixed interaction radial distance ℓ .

Hence, we have that the apoptosis rate for a WT cell i (N) and for a p53-/- cell j (P) are

$$\delta_N^i = d_N + (a_N \rho_{NN}^i + c_N \rho_{NP}^i) L^2 \quad (16)$$

$$\delta_P^j = d_P + (a_P \rho_{PP}^j + c_P \rho_{PN}^j) L^2 \quad (17)$$

where:

- d_N , d_P , a_N , a_P , c_N and c_P are defined as in the previous section (4.1).
- $\rho_{NN}^i = \frac{N_\ell^i}{\pi \ell^2}$ is the density of WT cells in the neighbourhood of the WT cell i , with N_ℓ^i being the number of WT cells within that neighbourhood.
- $\rho_{NP}^i = \frac{P_\ell^i}{\pi \ell^2}$ is the density of p53-/- cells in the neighbourhood of the WT cell i , with P_ℓ^i being the number of p53-/- cells within that neighbourhood.

- $\rho_{PP}^j = \frac{P_\ell^j}{\pi\ell^2}$ is the density of p53-/- cells in the neighbourhood of the p53-/- cell j , with P_ℓ^j being the number of p53-/- cells within that neighbourhood.
- $\rho_{PN}^j = \frac{N_\ell^j}{\pi\ell^2}$ is the density of WT cells in the neighbourhood of the p53-/- cell j , with N_ℓ^j being the number of WT cells within that neighbourhood.

Results

After comparing the average number of cells at time t , $\langle N(t) \rangle$ and $\langle P(t) \rangle$ for the mean-field interaction ($\ell \rightarrow \infty$) resulting from the agent based model, we see that we recover the deterministic model when in the mean-field regime.

Just like in the one species model, we see that for short-range signalling (see A1 in Figure 12), both populations get extinct. As we increase the diffusion coefficient, we observe that now both populations grow more steadily and compete more strongly, since cells have greater average displacements, and therefore, encounter more cells in their paths. Eventually, for great diffusions (see A3 in Figure 12), the stochastic model with short-range signalling starts to catch up with the deterministic prediction.

On the other hand, when $\ell \sim L$, we still see that the stochastic model improves its fitting to the deterministic one as we increase D (see Figure 12 B1-3).

Overall, in both cases we observe that modifying the diffusion for a fixed signalling-range results in effectively controlling the timescale of the trajectories. However, due to the limited computational resources and the time available for this project, we could not tackle this feature further. It would proceed to look at the times at which the maximum for N is attained rather than at the difference between N and P , which is what follows in this section.

Another important feature of the diffusion is clustering. We explored this effect by carrying out video simulations of the plate as shown in Figure 13. For typical values of diffusion (see Figure 13 A1 and A2), cells are well-mixed in the plate, even in the presence of cell competition. However, for low diffusions (see Figure 13 B1 and B2, with no diffusion), we observe that both species form clusters and respect some virtual spatial barriers created by competition.

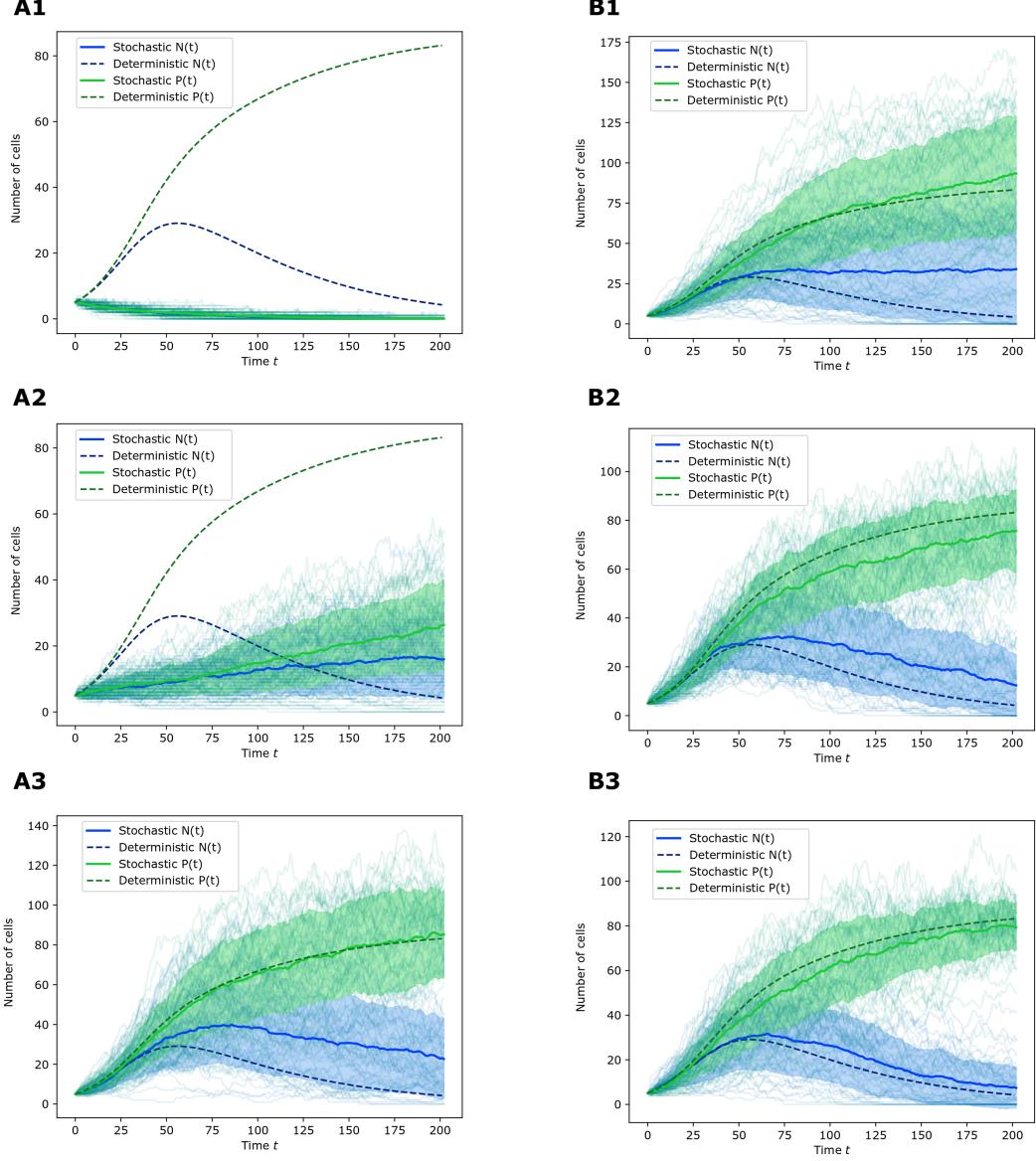


Figure 12: A1-A3 shows the stochastic trajectories of $N(t)$ and $P(t)$ (full line) and their deterministic values (dashed line) with $\ell = 0.5$ and (A1) $D = 0.001$, (A2) $D = 1.000$, (A3) $D = 10.000$, while B1-B3 has fixed $\ell = 25.0$ for (B1) $D = 0.001$, (B2) $D = 1.000$, (B3) $D = 10.000$. All of the graphs have coefficients $b_N = b_P = 0.0700$, $d_N = d_P = 0.0005$, $a_N = a_P = 0.0008$, $c_N = 0.0010$, and $c_P = 0.0006$.

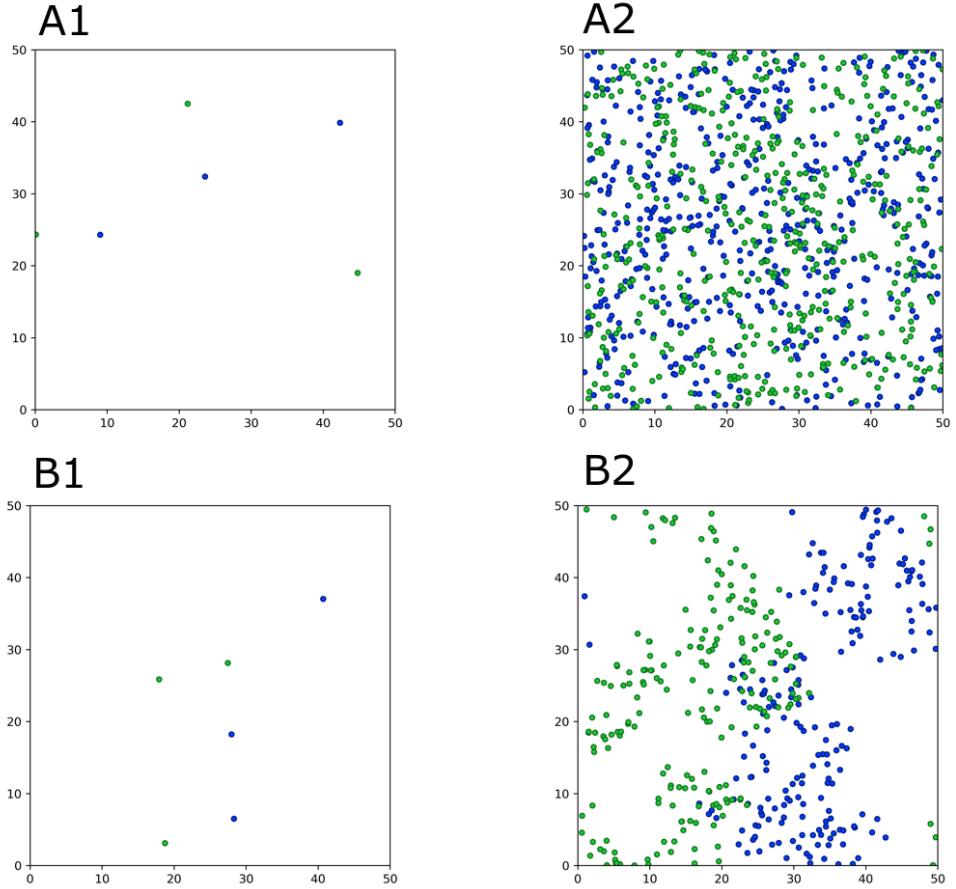


Figure 13: Snapshots at $t = 0$ and $t = 1000$ of a two-species simulation are shown for (A1-2) $D = 0$ and (B1-2) $D = 0.25$. WT cells are shown in blue, and p53 mutants, in green.

In order to understand further how diffusion controls the time scale of population growth in our two species model, we performed an analysis on the number of WT cells and p53 mutants at the time when we have a maximum for the population of WT cells.

We explored two different times — when the deterministic N attains its maximum, and when the stochastic N does.

In any of the cases, we did not see much difference in the averages of $N(t^*)$ and $P(t^*)$ for fixed values of ℓ (Figure 16 A B) and varying D versus the averages of $N(t_{det})$ and $P(t_{det})$

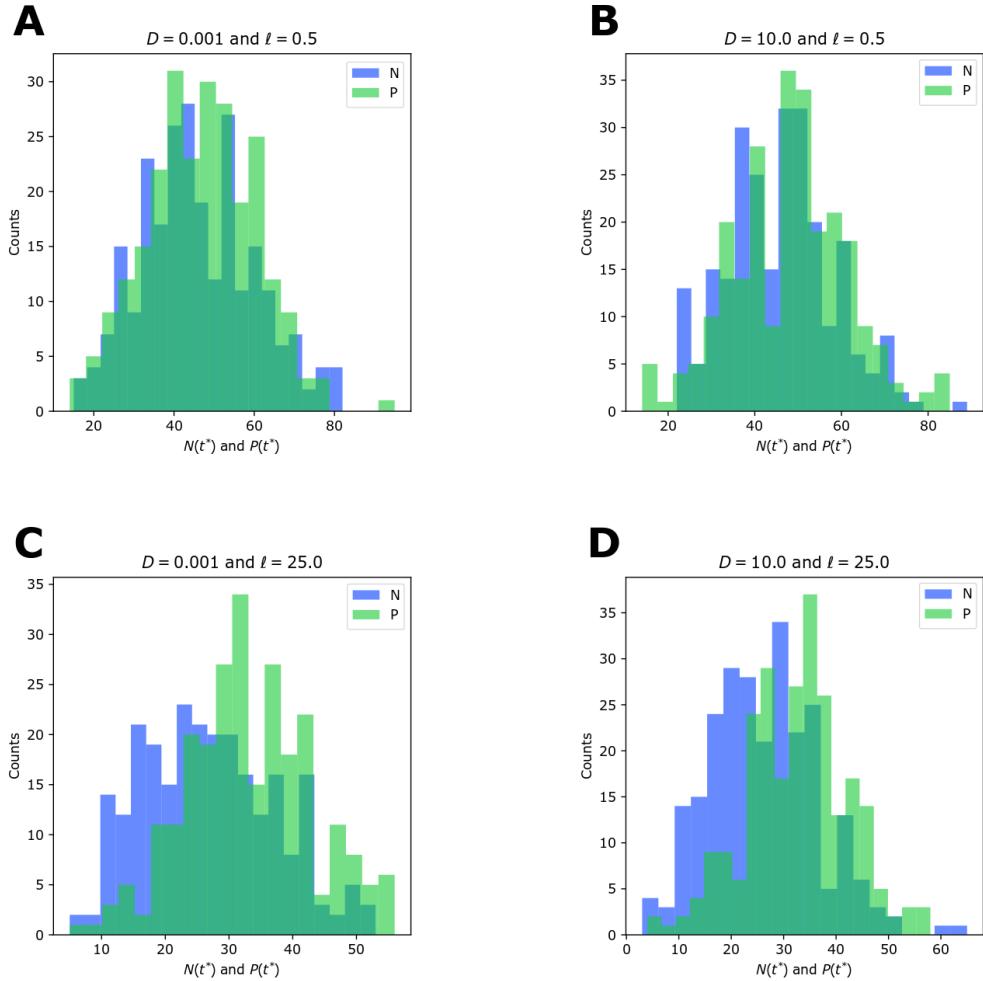


Figure 14: Histograms of the maximum stochastic value of N at time t^* , $N(t^*)$, and its corresponding value of P at that time, $P(t^*)$, for (A) $D = 0.001$ and $\ell = 0.5$, (B) $D = 10.000$ and $\ell = 0.5$, (C) $D = 0.001$ and $\ell = 25.0$, (D) $D = 10.000$ and $\ell = 25.0$.

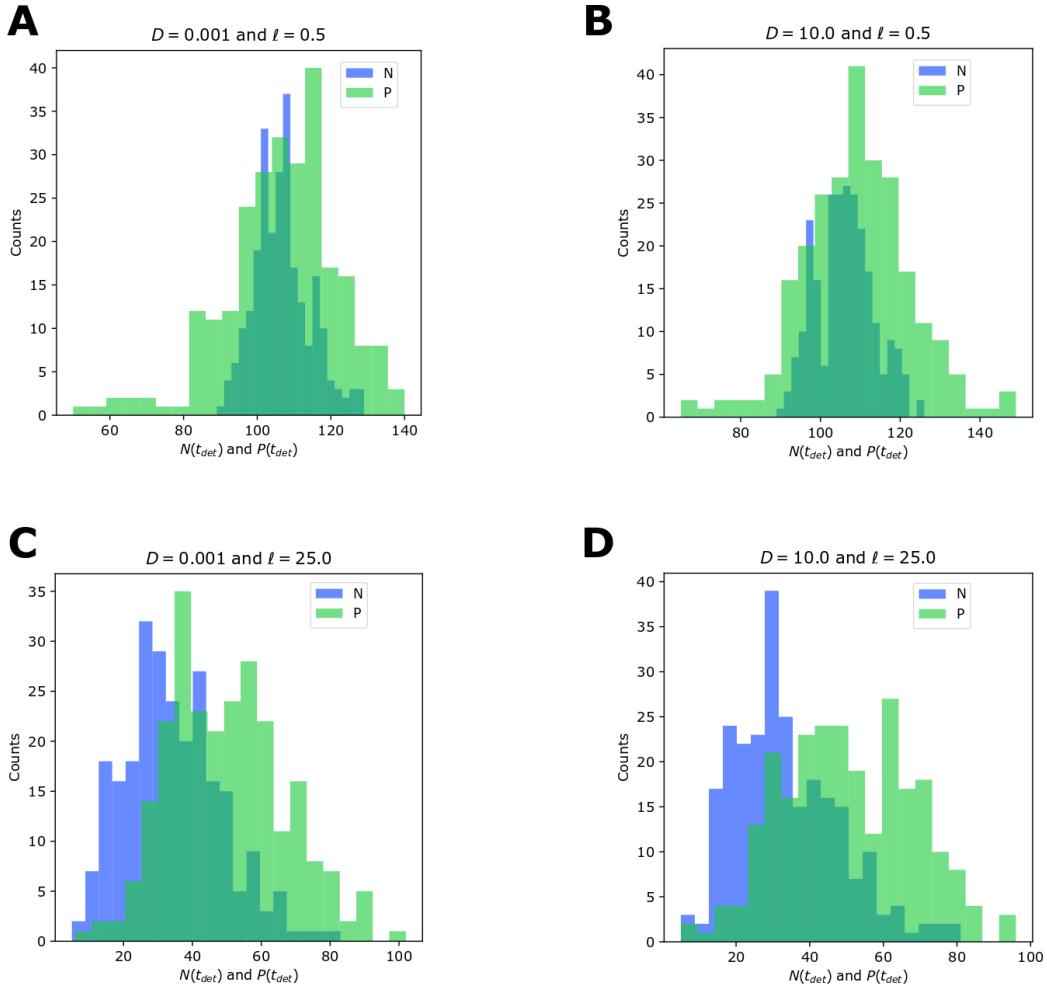


Figure 15: Histograms of the maximum deterministic value of N at time t_{det} , $N(t_{det})$ and its corresponding value of P at that time, $P(t_{det})$ for (A) $D = 0.001$ and $\ell = 0.5$, (B) $D = 10.000$ and $\ell = 0.5$, (C) $D = 0.001$ and $\ell = 25.0$, (D) $D = 10.000$ and $\ell = 25.0$.

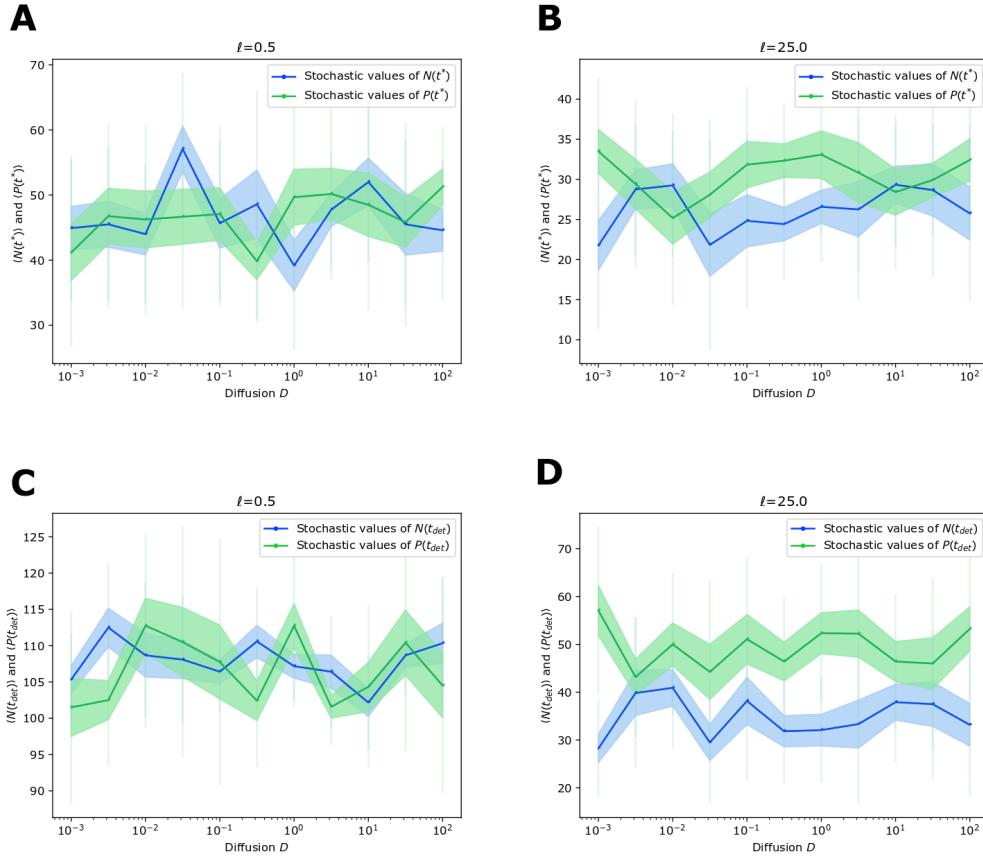


Figure 16: Plots of $\langle N(t^*) \rangle$ and $\langle P(t^*) \rangle$ vs D for (A) $\ell = 0.5$ and (B) $\ell = 25.0$, and of $\langle N(t_{det}) \rangle$ and $\langle P(t_{det}) \rangle$ vs D for (C) $\ell = 0.5$ and (D) $\ell = 25.0$.

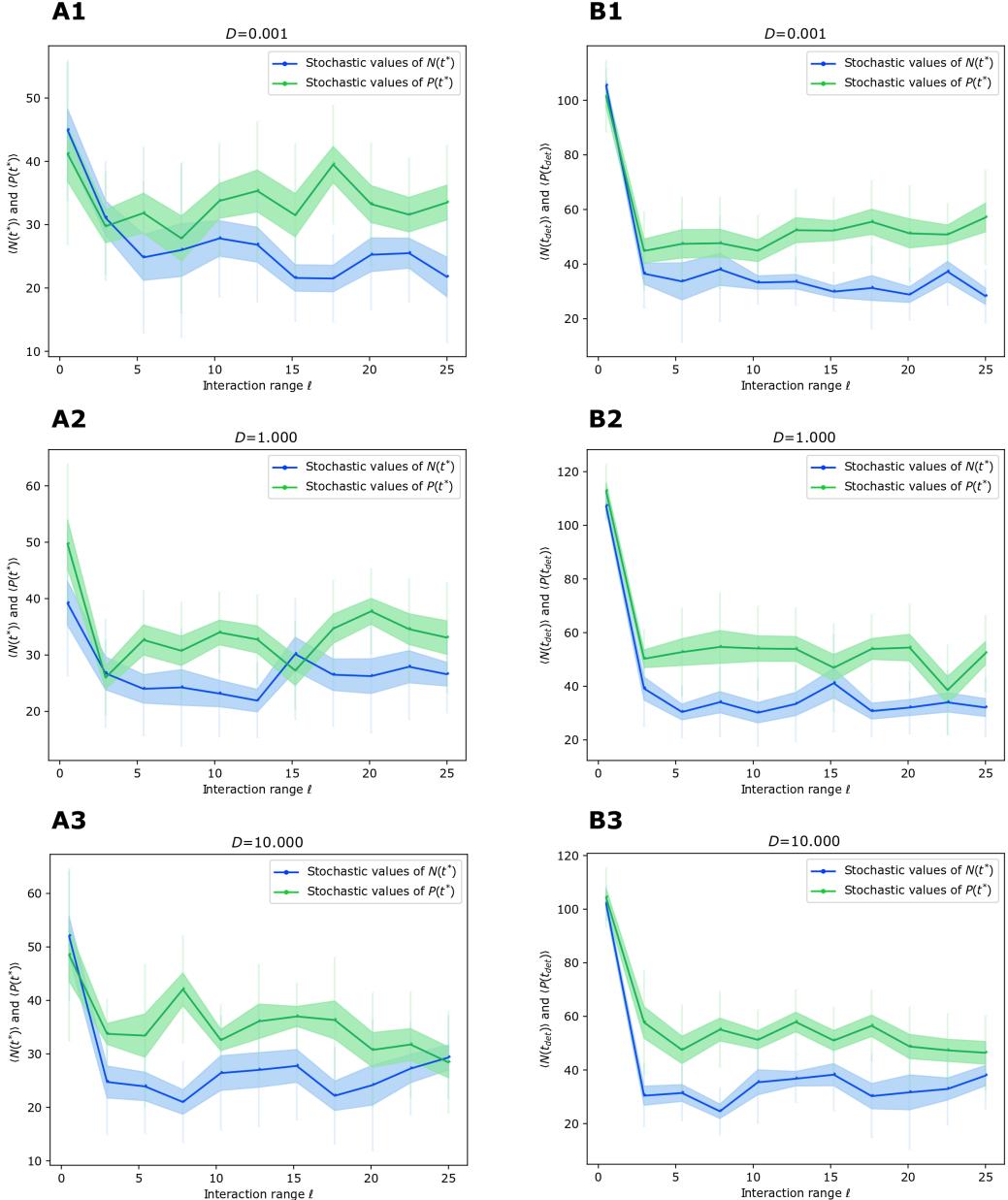


Figure 17: Plots of $\langle N^* \rangle$ and $\langle P^* \rangle$ vs ℓ for (A1) $D = 0.001$, (A2) $D = 1.000$, and (A3) $D = 10.000$, and of $\langle N_{det} \rangle$ and $\langle P_{det} \rangle$ vs ℓ for (B1) $D = 0.001$, (B2) $D = 1.000$, and (B3) $D = 10.000$.

5 Conclusions and discussion

In this paper, we have engineered an agent-based model to understand in vitro cell competition during early gastrulation within two different cell types — WT cells and p53-/ cells. All throughout, we have considered point-like cells moving with Brownian dynamics in a two-dimensional square region with periodic boundary conditions with arbitrary time and spatial units.

The modular structure of the project inherent to ABMs has helped us moving upwards in levels of complexity in an approachable way. However, this has been possible at the expense of computational power, which is limited in a typical academic setting like ours.

As mentioned in the main text, one of the limitations of agent based models is the requirement to compute interactions between a large number of individuals. Further research of this project should require more powerful computing resources, which would also make easier analysing the problem's statistical and asymptotic behaviour.

Whilst the model allowed to explore the role that cellular dynamics and range of interaction have on competition between different cellular subtypes, in order to have a more realistic description of the system, it is paramount to incorporate other possible cellular interactions, both physically and chemically.

From section 3.2.2 on, we implemented cell-cell signalling within a fixed radial distance ℓ . This scheme accounts for paracrine signalling (targeted to cells in the vicinity of the emitting cell) and endocrine signalling (targeted to more distant cells) in the limit of large ℓ . Nevertheless, there are other types of close-range signalling, like signalling via the secretion of hormones (autocrine and intocrine signalling), used to regulate intracellular events, that we did not include in the model. Some cells even interact with touching cells by using the proteins that are attached to the cell membrane (juxtacrine signalling). What is more, the rate of signalling transmission and propagation [29] is not necessarily constant.

It is also important to remark that, in the two species case, we focused on the case where the signalling distance is identical for both species. That is, if ℓ_N and ℓ_P are the interaction range of the WT cells and p53 -/- cells respectively, then we simplified the problem by assuming $\ell_N = \ell_P = \ell$, although we did implement different signalling ranges for each cell type in the

code.

Aside from cellular signalling, a more realistic scenario should include contact forces between the cells. On the one hand, repulsive forces were not implemented in our model, which allowed cells to "walk over" each other, or even "go through" each other without colliding. For this reason, the problem can be virtually seen as a three-dimensional one, rather than two-dimensional, setting the first steps towards a more accurate study of in vitro embryonic development. Additionally, a molecular dynamics (MD) approach could be taken in the future, where cells could be thought of as hard-spheres carrying out perfectly elastic collisions [30].

On the other hand, attractive forces within cells and between the cells and the plate result in cell adhesion [31]. However, by assuming periodic boundary conditions, we completely disregarded of cell-plate attractive forces on the edges of the petri dish. A higher spatial scale could be considered if in possession of better computational resources in order to account for cell-plate adhesion. Regarding cell-cell adhesion, an energetics and thermodynamical outlook could be taken, as considered in [31].

In addition to competition and mechanics, cells will also undergo differentiation, which will make them acquire different cellular identities corresponding to the progress of gastrulation. We expect that these upgrades change the physiology of the cell. This will incorporate variations in time to the different system parameters that had been considered constant during the work, such as proliferation or interaction terms.

Access to experimental data in the future would provide better contextualisation and more concreteness to the model. Bayesian inference could be preformed to parametrise the system more accurately, and find tighter and more precise coefficient constraints. In addition, experimental data would help us contrast the different time scales. In this aspect, understanding when the deterministic model is a good approximation for the behaviour of the agent based model can be an invaluable tool to accelerate the computer demand of such inference methods.

Another way of improving our model would be to consider a third agent — the substrate. Experimentally, the substrate concentration and the dissolution rate are relatively easy to control and measure in the lab. Furthermore, these two directly have an effect on the number of cells that can be sustained on the plate. This is why it would be useful to consider a so-called resource-

based or chemostat model [23].

Other examples of different approaches to the problem include evolutionary games and graph theory, where cells are compelled to occupy only the spaces available in a matrix [32], or to consider an age-structured population. Each "age" of the cell would correspond to a stage within the cell-cycle. There would be two different ways of escaping the cycle: by cell cycle arrest (i.e. by entering the so-called G0 phase), or by apoptosis, both of them depending on the density of neighbouring cells.

Finally, the most natural next step to take in this project would be to analyse the system using PDEs. We could study the concentration of cells in the plate with time and space being the dependent variables [33], or by using the Fokker-Plank equation [32].

Overall, in this work we have shown how combining mathematical and computational methods of different complexity can help us gain insights in cell competition during early embryo development. This emphasizes the thriving importance of mathematical modelling in biology, revealing not only potential analytical and computational tools to tackle the problem, but also guiding new experiments to achieve a mechanistic understanding of embryo formation.

6 Appendix

In this section we include the most relevant script of the project, on which all the other scripts are built upon. For the complete set of files and Python scripts used for plotting and creating video simulations, please visit <https://bitbucket.org/2piruben/cellcomp/src/master/>.

This script contains two main functions. The first one, *cells_evolution*, adds new cells to the system, removes dying cells, and diffuses cells. It includes the Monte Carlo method for acceptance of birth and death processes and the periodic boundary conditions to put the cells back in the box if they exit it.

The second function, *run_cells_linear*, contains the loop in time, and calls *cells_evolution* in every iteration. This is how the dynamics of the system arises. It has two modalities — you can either turn on short-range interactions, or use a mean-field approximation. Within the short-range mode, distances between cell pairs are computed in order to calculate the effective apoptosis rate of each cell.

You can accommodate this script for each of the cases described in this project by changing the values of each cell types and their corresponding coefficients.

```
1 # -*- coding: utf-8 -*-
2 """
3 Created on Wed Nov 27 10:48:12 2019
4
5 @author: casmp
6 """
7
8
9 import numpy as np
10 from random import choices
11
12 #This script contains the functions regarding two species
13     cell competition
14
15 #Time variables
16 dt = 0.025
17 timelapse = 1 # how often to record
18 Tfinal = 202
19 n = int(Tfinal/dt)    #timesteps
```

```

20
21
22 #Spatial variables
23 L = 50 #length of box
24 l_N = 0.5 #interaction range for species N
25 l_P = 20 #interaction range for species N
26 D = 1 # diffusion coefficient
27 diam = 0.1 # distance from each other at which cells divide
28
29
30 #Initial number of cells for each species (N and P)
31 N0 = 5
32 P0 = 5
33
34 # Equations:
35 # dN / dt = b_N * N - (d_N + a_N * N + c_N * P) * N
36 # dP / dt = b_P * P - (d_P + a_P * P + c_P * N) * P
37
38 #Proliferation rates
39 b_N = 0.070
40 b_P = 0.070
41
42 #Death rates
43 d_N = 0.00005 #for species N
44 d_P = 0.00005 #for species P
45
46
47 #Coupling constans
48
49 #N
50 a_N= 0.00080 #with N
51 c_N= 0.0010 #with P
52
53 # P
54 a_P = 0.00080 #with P
55 c_P= 0.00060 #with N
56
57 # a_x -> competition with its own species x
58 # c_x -> competition of x with the other species y
59
60
61
62 def cells_evolution(cellNumber, cellarray, a, d): #kills or
       gives birth to cells and moves them around
63
64     if cellNumber > 0:
65         u = np.random.rand(cellNumber)      #proliferation die
66         thrown cellNumber times
67         v = np.random.rand(cellNumber)      #death die thrown

```

```

cellNumber times

proliferatingCells = u < a*dt # a*dt = probability
of a cell proliferating in that time window
survivingCells = v > d*dt # b*dt = probability of a
cell dying in dt

new_d = d

# Limiting events to one event per dt (either one
cell dies or one cell proliferates)

if (np.sum(proliferatingCells)+np.sum(np.logical_not(
survivingCells)) > 0):

    prob_prol = np.sum(proliferatingCells)/(np.sum(
proliferatingCells)+np.sum(np.logical_not(survivingCells)))
)

    if prob_prol < np.random.rand(1):
        proliferatingCells = np.zeros(cellNumber,
dtype=bool) #no cells proliferating
        dyingCell = choices(population=np.arange(
cellNumber), k=1, weights=np.logical_not(survivingCells)/
np.sum(np.logical_not(survivingCells))) #cel that has
been randomly chosen to die
        survivingCells = np.ones(cellNumber, dtype=
bool)
        survivingCells [dyingCell[0]] = False

    else:
        survivingCells = np.ones(cellNumber, dtype=
bool)
        prolCell = choices(population=np.arange(
cellNumber), k=1, weights=proliferatingCells/np.sum(
proliferatingCells))
        proliferatingCells = np.zeros(cellNumber,
dtype=bool)
        proliferatingCells [prolCell[0]] = True

    # newcells_0 is the new cells that will be added to
the system
    newcells_0= cellarray[proliferatingCells]
    newcells_d = new_d[proliferatingCells]
    cellNumber += np.sum(proliferatingCells)

```

```

99      # If a proliferation event occurs, we add the
100     daughter cell to the system at an angle theta from its
101     mother cell
102     if np.sum(proliferatingCells) > 0:
103         theta = np.random.rand(np.sum(proliferatingCells))
104         )*2*np.pi
105         newcells_x = newcells_0[:,0]+diam*np.cos(theta)
106         newcells_y = newcells_0[:,1]+diam*np.sin(theta)
107
108         newcells=[]
109         for i, cellx in enumerate(newcells_x):
110             newcells.append([newcells_x[i], newcells_y[i]
111         ]))
112         newcells_0 = newcells
113
114
115
116     # removing dying cells
117     cellararray = cellararray[survivingCells]
118     new_d = new_d[survivingCells]
119     cellNumber -= np.sum(np.logical_not(survivingCells))
120
121     # adding newcells
122     cellararray = np.vstack((cellararray,newcells_0))
123     new_d = np.hstack((new_d,newcells_d))
124
125
126     # moving cells
127     cellararray[:,0] = cellararray[:,0] + np.sqrt(D*dt)*np.
128     random.normal(size=cellNumber) # diffusion x
129     cellararray[:,1] = cellararray[:,1] + np.sqrt(D*dt)*np.
130     random.normal(size=cellNumber) # diffusion y
131
132
133
134     # Back to the box (PBCs)
135
136     # cell exits the box from the left
137     exitBoxLeft = cellararray[:,0]<0
138     mLeft = abs(cellarray[exitBoxLeft,0]/L)
139     cellararray[exitBoxLeft,0] += np.ceil(mLeft)*L
140
141     # cell exits the box from the right
142     exitBoxRight = cellararray[:,0]>L
143     mRight = abs(cellarray[exitBoxRight,0]/L)
144     cellararray[exitBoxRight,0] -= np.floor(mRight)*L
145
146     # cell exits the box from the bottom
147     exitBoxBottom = cellararray[:,1]<0
148     mBottom = abs(cellarray[exitBoxBottom,1]/L)
149     cellararray[exitBoxBottom,1] += np.ceil(mBottom)*L
150
151

```

```

142
143     # cell exits the box from the top
144     exitBoxTop = cellarray[:,1]>L
145     mTop = abs(cellarray[exitBoxTop,1]/L)
146     cellarray[exitBoxTop,1] -= np.floor(mTop)*L
147
148     return(cellNumber, cellarray, new_d)
149
150 else:
151     # i.e. if cellNumber == 0
152     return(0,np.zeros(0),np.zeros(0))
153
154
155
156 def run_cells_linear(D,l_N, l_P, interaction="all_plate"):
157
158     # Initialitation of the variables
159     t = 0
160     history_cells = [] #N(t) & P(t)
161     initcelllistN = [] # list of cells N
162     initcelllistP = [] # list of cells P
163     nextrecordtime = 0 # next time to record time
164
165     # starting the population
166
167     # The information of the species N will be stored as an
168     # array with two columns, cellarrayN.
169     # Each row is a N cell and the two columns are its x and
170     # y coordinates, respectively
171
172     for k in range(N0):
173         initcelllistN.append([np.random.rand()*L,np.random.
174         rand()*L])
175     cellarrayN = np.array(initcelllistN)
176     cellNumberN = N0 # This variable will count the number of
177     # N cells at each time point
178
179     # Similarly, for P.
180     for k in range(P0):
181         initcelllistP.append([np.random.rand()*L,np.
182         random.rand()*L])
183     cellarrayP = np.array(initcelllistP)
184     cellNumberP = P0 # This variable will count the number of
185     # P cells at each time point
186
187     # Loop in time
188     for i in range(n):
189
190         # SHORT-RANGE SIGNALLING

```

```

185
186     if (interaction == "short_range"):
187
188         #Death coefficient for each cell depending on
189         #cell density every 10 iterations
190         if i%10==0:
191
192             cellNumberN_N_1 = np.array(np.zeros(
193             cellNumberN)) #number of cells of species N within range
194             l from cells of species N
195             cellNumberP_P_1 = np.array(np.zeros(
196             cellNumberP)) #number of cells of species P within range
197             l from cells of species P
198             cellNumberN_P_1 = np.array(np.zeros(
199             cellNumberN)) #number of cells of species P within range
200             l from cells of species N
201             cellNumberP_N_1 = np.array(np.zeros(
202             cellNumberP)) #number of cells of species N within range
203             l from cells of species P
204
205             #Calculate density of cells within l for each
206             #individual of species N and each death rate
207             for icell, celli in enumerate(cellarrayN):
208
209                 for jcell, cellj in enumerate(cellarrayN)
210                 :
211                     if (jcell > icell):
212                         dx_N = min(abs(cellj[0]-celli[0]))
213                         , L - abs(cellj[0]-celli[0]))
214                         dy_N = min(abs(cellj[1]-celli[1]))
215                         , L - abs(cellj[1]-celli[1]))
216                         dist_N_N = (dx_N**2 + dy_N**2)
217                         **0.5 #distance to each N cell
218
219                         if dist_N_N < l_N:
220                             cellNumberN_N_1[icell] += 1
221
222                         #number of N cells within range l surrounding N cell i
223                         cellNumberN_N_1[jcell] += 1
224
225                         #Density of cells: P vs P
226                         for icell, celli in enumerate(cellarrayP):
227
228                             for jcell, cellj in enumerate(cellarrayP)
229                             :
230                                 if (jcell > icell):
231                                     dx_P = min(abs(cellj[0]-celli[0]))
232                                     , L - abs(cellj[0]-celli[0)))
233                                     dy_P = min(abs(cellj[1]-celli[1]))
234                                     , L - abs(cellj[1]-celli[1]))
```

```

216                         dist_P_P = (dx_P**2 + dy_P**2)
217                         **0.5
218
219                         if dist_P_P < l_P:
220                             cellNumberP_P_1[icell] += 1 # number of P cells within range l surrounding P cell i
221                             cellNumberP_P_1[jcell] += 1
222
223                         # Density of cells: N vs P and P vs N
224                         for icell, celli in enumerate(cellarrayN):
225
226                             for jcell, cellj in enumerate(cellarrayP):
227
228                                 dx_N_P = min(abs(celli[0]-cellj[0]),
229                                 L - abs(cellj[0]-celli[0]))
230                                 dy_N_P = min(abs(celli[1]-cellj[1]),
231                                 L - abs(cellj[1]-celli[1]))
232                                 dist_N_P = (dx_N_P**2 + dy_N_P**2)
233                                 **0.5 #distance to each P cell
234
235                                 if dist_N_P < l_N:
236                                     cellNumberN_P_1[icell] += 1 # number of P cells around cell number i of N species
237
238                                 if dist_N_P < l_P:
239                                     cellNumberP_N_1[jcell] += 1 # number of N cells around cell number j of P species
240
241
242                         #Calculating death rate vector (each
243                         effective apoptosis rate) for each species
244                         dcell_N = d_N + a_N*(cellNumberN_N_1/(np.pi*
245                         l_N**2))*(L**2) + c_N*(cellNumberN_P_1/(np.pi*l_N**2))*(L
246                         **2)
247                         dcell_P = d_P + a_P*(cellNumberP_P_1/(np.pi*
248                         l_P**2))*(L**2) + c_P*(cellNumberP_N_1/(np.pi*l_P**2))*(L
249                         **2)
250
251
252                         #calculate surviving and dying cells and move
253                         surviving cells around
254                         cellNumberN, cellarrayN, dcell_N =
255                         cells_evolution(cellNumberN, cellarrayN, b_N, dcell_N)
256                         cellNumberP, cellarrayP, dcell_P =
257                         cells_evolution(cellNumberP, cellarrayP, b_P, dcell_P)
258
259                         else:

```

```

249         #calculate surviving and dying cells and move
250         #surviving cells around
251         cellNumberN, cellarrayN, dcell_N =
252             cells_evolution(cellNumberN, cellarrayN, b_N, dcell_N)
253             cellNumberP, cellarrayP, dcell_P =
254             cells_evolution(cellNumberP, cellarrayP, b_P, dcell_P)
255
256
257     # LONG-RANGE SINGALLING
258
259     elif (interaction == "all_plate"):
260
261         dcomp_N = d_N + a_N*cellNumberN + c_N*cellNumberP
262         dcomp_P = d_P + a_P*cellNumberP + c_P*cellNumberN
263
264         dcomp_N = np.ones(cellNumberN)*dcomp_N
265         dcomp_P = np.ones(cellNumberP)*dcomp_P
266
267
268     #calculate surviving and dying cells and move
269     #surviving cells around
270     cellNumberN, cellarrayN, dcomp_N =
271         cells_evolution(cellNumberN, cellarrayN, b_N, dcomp_N)
272         cellNumberP, cellarrayP, dcomp_P =
273         cells_evolution(cellNumberP, cellarrayP, b_P, dcomp_P)
274
275
276
277     # updating time and saving N(t) if necessary
278     t = t + dt
279
280
281     if t >= nextrecordtime:
282         # We record the information about the system
283         after a specific timelapse
284         history_cells.append([t,cellNumberN,cellNumberP])
285         nextrecordtime += timelapse
286
287
288
289     return np.array(history_cells), cellarrayN, cellarrayP

```

References

- [1] S. F. Gilbert, *Developmental Biology*. Sinauer Associates, 2016.
- [2] T. Mahajan, S. Ganguly, and N. Pagrut, “Embryogenesis: a comprehensive review,” *Journal of entomology and zoology studies*, vol. 6, pp. 1151–1153, 02 2018.
- [3] C. Brás-Pereira and E. Moreno, “Mechanical cell competition,” *Current Opinion in Cell Biology*, vol. 51, pp. 15 – 21, 2018. Cell Signalling.
- [4] A. Di Gregorio, S. Bowling, and T. Rodriguez, “Cell competition and its role in the regulation of cell fitness from development to cancer,” *Developmental Cell*, vol. 38, no. 6, pp. 621 – 634, 2016.
- [5] 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. 569, no. 7757, pp. 497–502, 2019.
- [6] G. Zhang, Y. Xie, Y. Zhou, C. Xiang, L. Chen, C. Zhang, X. Hou, J. Chen, H. Zong, and G. Liu, “p53 pathway is involved in cell competition during mouse embryogenesis,” *Proceedings of the National Academy of Sciences*, vol. 114, no. 3, pp. 498–503, 2017.
- [7] T. A. R. Aida Di Gregorio, Sarah Bowling, “Cell competition and its role in the regulation of cell fitness from development to cancer,” *Developmental Cell*, vol. 38, no. 6, pp. 621 – 634, 2016.
- [8] S. Venkatachalam, S. D. Tyner, C. R. Pickering, S. Boley, L. Recio, J. E. French, and L. A. Donehower, “Is p53 haploinsufficient for tumor suppression? implications for the p53+/- mouse model in carcinogenicity testing,” *Toxicologic Pathology*, vol. 29, no. 1_suppl, pp. 147–154, 2001. PMID: 11695551.
- [9] T. Ozaki and A. Nakagawara, “Role of p53 in cell death and human cancers,” *Cancers*, vol. 3, no. 1, pp. 994–1013, 2011.
- [10] F. Rodier, J. Campisi, and D. Bhaumik, “Two faces of p53: aging and tumor suppression,” *Nucleic Acids Research*, vol. 35, pp. 7475–7484, 10 2007.
- [11] X. Lou, M. Kang, P. Xenopoulos, S. Muñoz-Descalzo, and A.-K. Hadjantonakis, “A rapid and efficient 2d/3d nuclear segmentation method for analysis of early mouse embryo and stem cell image data,” *Stem Cell Reports*, vol. 2, no. 3, pp. 382 – 397, 2014.

- [12] S. de Beco, M. Ziosi, and L. A. Johnston, “New frontiers in cell competition†,” *Developmental Dynamics*, vol. 241, no. 5, pp. 831–841, 2012.
- [13] L. G. Shoemaker, L. L. Sullivan, I. Donohue, J. S. Cabral, R. J. Williams, M. M. Mayfield, J. M. Chase, C. Chu, S. H. W, A. Huth, J. HilleRis-Lambers, A. R. M. James, N. J. B. Kraft, F. May, R. Muthukrishnan, S. Satterlee, F. Taubert, X. Wang, T. Wiegand, Q. Yang, and K. C. Abbott, “Integrating the underlying structure of stochasticity into community ecology,” *Ecology*, vol. 101, no. 2, 2020.
- [14] G. A. Fox and B. E. Kendall, “Demographic stochasticity and the variance reduction effect,” *Ecology*, vol. 83, no. 7, pp. 1928–1934, 2002.
- [15] E. Bonabeau, “Agent-based modeling: Methods and techniques for simulating human systems,” *Proceedings of the national academy of sciences*, vol. 99, no. suppl 3, pp. 7280–7287, 2002.
- [16] V. Grimm, E. Revilla, U. Berger, F. Jeltsch, W. M. Mooij, S. F. Railsback, H.-H. Thulke, J. Weiner, T. Wiegand, and D. L. DeAngelis, “Pattern-oriented modeling of agent-based complex systems: Lessons from ecology,” *Science*, vol. 310, no. 5750, pp. 987–991, 2005.
- [17] G. An, Q. Mi, J. Dutta-Moscato, and Y. Vodovotz, “Agent-based models in translational systems biology,” *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, vol. 1, no. 2, pp. 159–171, 2009.
- [18] A. Mogilner and A. Manhart, “Agent-based modeling: case study in cleavage furrow models,” *Molecular biology of the cell*, vol. 27, no. 22, pp. 3379–3384, 2016.
- [19] C. Neuhauser, “Mathematical challenges in spatial ecology,” *Notices Amer, Math Soc*, vol. 48, no. 11, pp. 1304–1314, 2001.
- [20] L. J. Schumacher, P. K. Maini, and R. E. Baker, “Semblance of heterogeneity in collective cell migration,” *Cell Systems*, vol. 5, no. 2, pp. 119–127, 2017.
- [21] C. W. Reynolds, “Flocks, herds and schools: A distributed behavioral model,” *SIGGRAPH Comput. Graph.*, vol. 21, p. 25–34, Aug. 1987.
- [22] T. R. Malthus, D. Winch, and P. James, *Malthus:‘An Essay on the Principle of Population’*. Cambridge University Press, 1992.
- [23] M. Kot, *Elements of mathematical ecology / Mark Kot*. Cambridge: Cambridge University Press, 2001.

- [24] P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, vol. 23. Springer Science & Business Media, 2013.
- [25] J. E. Gubernatis, “Marshall rosenbluth and the metropolis algorithm,” *Physics of plasmas*, vol. 12, no. 5, p. 057303, 2005.
- [26] N. Van Kampen, *Stochastic Processes in Physics and Chemistry*. North-Holland Personal Library, Elsevier Science, 2011.
- [27] A. Tsoularis and J. Wallace, “Analysis of logistic growth models,” *Mathematical Biosciences*, vol. 179, no. 1, pp. 21 – 55, 2002.
- [28] P.-F. Verhulst, “Notice sur la loi que la population suit dans son accroissement,” *Corresp. Math. Phys.*, vol. 10, pp. 113–126, 1838.
- [29] R. E. Chen and J. Thorner, “Systems biology approaches in cell signaling research,” *Genome biology*, vol. 6, no. 10, p. 235, 2005.
- [30] P. Schwille, U. Haupts, S. Maiti, and W. W. Webb, “Molecular dynamics in living cells observed by fluorescence correlation spectroscopy with one- and two-photon excitation,” *Biophysical journal*, vol. 77, no. 4, pp. 2251–2265, 1999.
- [31] G. I. Bell, M. Dembo, and P. Bongrand, “Cell adhesion. competition between nonspecific repulsion and specific bonding.,” *Biophysical journal*, vol. 45, no. 6, p. 1051, 1984.
- [32] S. T. Johnston, M. J. Simpson, and E. J. Crampin, “Predicting population extinction in lattice-based birth-death-movement models,” *arXiv preprint arXiv:2002.05357*, 2020.
- [33] D. J. Jörg, Y. Kitadate, S. Yoshida, and B. D. Simons, “Competition for stem cell fate determinants as a mechanism for tissue homeostasis,” *arXiv preprint arXiv:1901.03903*, 2019.