Modelling uniaxial non-uniform yeast colony growth: comparing an agent-based model and continuum approximations

Anthony J. Gallo

School of Mathematical Sciences, The University of Adelaide, Adelaide 5005, Australia

Hayden Tronnolone

College of Science and Engineering, Flinders University, GPO Box 2100, Adelaide 5001, Australia

J. Edward F. Green[1], Benjamin J. Binder

School of Mathematical Sciences, The University of Adelaide, Adelaide 5005, Australia

Abstract

Biological experiments have shown that yeast can be restricted to grow in a uniaxial direction, vertically upwards from an agar plate to form a colony. The growth occurs as a consequence of cell proliferation driven by a nutrient supply at the base of the colony, and the height of the colony has been observed to increase linearly with time. Within the colony the nutrient concentration is non-constant and yeast cells throughout the colony will therefore not have equal access to nutrient, resulting in non-uniform growth. In this work, an agent based model is developed to predict the microscopic spatial distribution of labelled cells within the colony when the probability of cell proliferation can vary in space and time. We also describe a method for determining the average trajectories or pathlines of labelled cells within a colony growing in a uniaxial direction, enabling us to connect the microscopic and macroscopic behaviours of the system. We present results for six cases, which involve different assumptions for the presence or absence of a quiescent region (where no cell proliferation occurs), the size of the proliferative region, and the spatial variation of proliferation rates within the proliferative region. These six cases are designed to provide qualitative insight into likely growth scenarios whilst remaining amenable to analysis. We compare our macroscopic results to experimental observations of uniaxial colony growth for two cases where only a fixed number of cells at the base of the colony can proliferate. The model predicts that the height of the colony will increase linearly with time in both these cases, which is consistent with experimental observations. However, our model shows how different functional forms for the spatial dependence of the proliferation rate can be distinguished by tracking the pathlines of cells at different positions in the colony. More generally, our methodology can be applied to other biological systems exhibiting uniaxial growth, providing a framework for classifying or determining regions of uniform and non-uniform growth.

Keywords: Uniaxial growth; Nonuniform growth; Agent-based model; Continuum approximation; Yeast

1. Introduction

Yeasts are unicellular microorganisms classified as part of the fungi kingdom [1]. They are a eukaryotic fungus and thus have organelles, such a cell nuclei and mitochondria, within their cells. This contrasts with prokaryotic microorganisms, such as bacteria, that do not have organelles. There are currently over 1500 recognised species of yeast [1]. The yeast species *Saccharomyces cerevisiae* was the first eukaryotic cell to have its genome fully sequenced [2] and is commonly used as a model organism for other eukaryotic cells (e.g. plants and animals)[3]. Yeast growth experiments and models are therefore important in investigating the genetic traits of diseases in humans such as cancer [4], and also have application in the understanding of the mechanisms and environmental factors that cause biofilm

[1]Corresponding author - email: edward.green@adelaide.edu.au

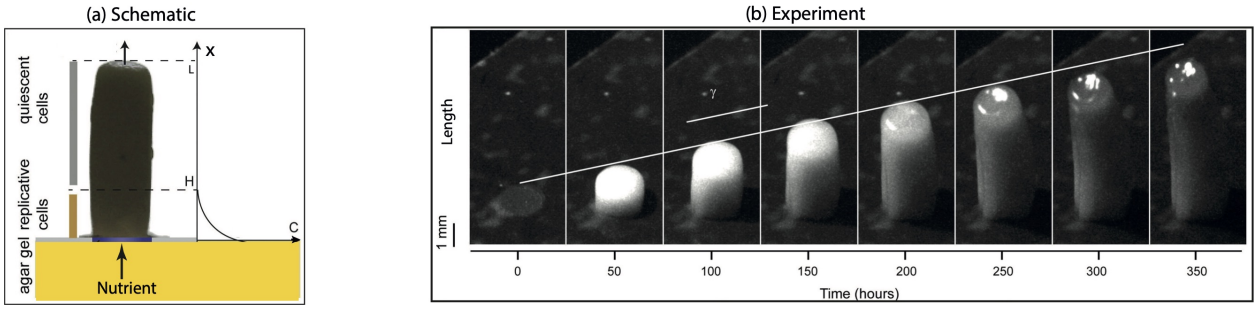

Figure 1: Modified images taken with permission from Vulin et al. [19]. (a) Schematic of uniaxial yeast growth. The nutrient concentration $C \geq 0$ is monotonically decreasing from the base of the colony until it reaches a distance H at which $C = 0$ mM. (b) Experiment of uniaxial yeast growth. The constant diameter of this cylindrical colony is approximately 1.5 mm and the constant flux of nutrient into the agar gel is 111 mM. The height or length of the colony, L, is observed to increase linearly with time, on the order of 1 mm per 50 hours.

formation on medical implants (e.g. catheters and stents) [5, 6]. In addition, yeast is used for brewing beverages [7, 8] and as a leavening agent when baking bread [9], with further biotechnological applications such as the production of biofuels [10] and the treatment of wastewater [11]. Therefore, understanding the response of yeast growth in non-constant nutrient environments is an important problem to consider.

Saccharomyces cerevisiae, like many other yeasts, typically reproduce asexually via mitosis in the presence of a nutrient. As a result, yeast grows by consuming nutrient and the behaviour of cells within yeasts can be modified by adjusting the nutrient supply [12]. In experiments, this nutrient typically comes from an agar medium within a Petri dish. Many yeast experiments, and mathematical models describing them, examine growth in the radial direction along the surface of the dish [5, 13, 14, 4, 15, 16, 17, 18]. However, Vulin et al. [19] showed that *Saccharomyces cerevisiae* can be restricted to grow in a uniaxial direction vertically upwards from an agar plate to form a cylindrical colony (see Figure 1). In the experiment, the culture system ensured nutrient was delivered into the agar gel directly beneath the colony, resulting in a monotonically decreasing nutrient concentration from the base to the top of the colony with both a replicative and quiescent region within the colony (see Figure 1a). In the replicative region there is sufficient nutrient for the cells to proliferate, whereas in the quiescent region there is insufficient nutrient for cell proliferation. This is due to the nutrient being consumed in the lower region of the colony at a greater rate than it can diffuse into the upper region. The distance H that the nutrient can reach up the colony depends on the nutrient consumption rate of the yeast, the amount of nutrient delivered and its diffusivity [20].

Time lapse images of a cylindrical yeast colony grown by Vulin et al. [19] are shown in Figure 1(b). The growth is uniaxial and the height or length L of the colony was observed to increase linearly with time. It was hypothesised by Vulin et al. [19] that this linear increase in the length is due to the fact that nutrient can only reach a certain distance up the colony. The purpose of this study is to validate this hypothesis using a discrete agent-based modelling approach. As observations of uniaxial growth are applicable to other biological systems, for example, embryonic gut tissue growth [21], it is important to understand the relationship between nutrient concentration and the spatial properties within the colony.

The non-constant nutrient concentration within the colony results in spatially dependent growth [22, 23]. As both quantities are difficult to measure during an experiment, little is known of their relationship and how this influences the colony height. We address this by developing a one-dimensional model of uniaxial growth that provides new insights into the influence of the nutrient distribution on the cell growth rates and colony morphology. We explore the growth rates by tracking the average trajectories or pathlines of labelled cells within a colony. This enables us to establish regions of macroscopic uniform and non-uniform growth within the colony.

In general, the growth of yeast colonies on agar gel is a problem in three spatial dimensions. Yeast colonies may grow into various different shapes and can experience radial growth. However, we observe in Figure 1(b) that the diameter of the colony is fixed to 1.5 mm throughout the duration of the experiment. As there is no radial growth, the cells will remain in a fixed radial and azimuthal position. We thus assume that the growth is in a uniaxial direction, which is consistent with experimental observations. Furthermore, Figure 1(a) suggests that the nutrient concentration is only varying in one spatial dimension — it is monotonically decreasing in the same direction as the uniaxial growth

of the colony. Thus, it is plausible to assume that two cells at the same height will have the same amount of access to nutrient. Cell proliferation, and hence cell displacement, is nutrient driven, so we expect two cells at the same height to have the same proliferation rate. As a result, it is a reasonable assumption to model the non-uniform growth of the colony in one spatial dimension.

The model we develop is an agent-based cellular automaton (CA). These models have previously been used to model proliferative tissue growth [24, 25, 26, 27], along with other intercellular interactions (e.g. cell motility) [28, 29, 30, 31]. Applications include tumour growth [32, 33], embryonic tissue growth [21], bone growth [34], fungal colonies [35] and yeast colonies [15].

Binder et al. [21] developed a CA model for uniaxial growth in the gut tissue of an embryonic quail. They found that uniform cell proliferation (implying that all cells have equal access to nutrient) results in uniform uniaxial growth at the macroscale. Binder et al. [21] also considered a uniaxial growth model where the tissue was split into three sections with three different constant growth rates. This adds position-dependent proliferation into the CA model, and it was shown to result in overall non-uniform growth of the tissue (although the growth in each section is uniform). A major limitation of this piecewise uniform model is that the non-uniform growth rate cannot be continuous across the domain, which is physically unrealistic in the majority of biological systems. Lai De Oliveira and Binder [36] overcame this limitation by introducing a probability mass function that prescribed the probability of cell proliferation as a function of position. This spatially dependent cell proliferation can account for a non-constant nutrient concentration and resulted in non-uniform growth. Although our work implements the same basic CA mechanism for cell proliferation found in these two studies, there is an important and fundamental distinction in the overall modelling approach.

Both Binder et al. [21] and Lai De Oliveira and Binder [36] assumed that the time evolution of the length of the domain (colony or tissue) was known in advance, and we remark that domain growth has also been commonly specified in continuum models [37, 38, 39, 40]. This assumption is reasonable in the sense that the length of the domain is a quantity that is easily observed, measured and recorded during an experiment (e.g. Figure 1b). The fact that the length of the domain was known in advance at any time then pre-determined the number of cell proliferation events to occur in a given time step of the CA models of Binder et al. [21] and Lai De Oliveira and Binder [36]. Uniform and position-dependent cell proliferation probabilities, reflecting constraints on the nutrient concentration, were considered and a combinatorial approach was taken to determine the cell displacement probability distribution, with the mean of the distribution being intuitively related to the average trajectories, or pathlines, of the labelled cells. For uniform cell proliferation, the mean of the cell displacement distribution was found in closed-form [21], whereas for position dependent cell proliferation the mean of the cell displacement distribution needed to be computed numerically [36]. Clearly, the latter numerical approach limits the potential for further analysis, but more importantly, neither study allowed the length of the domain to emerge in response to the specification of an input nutrient concentration. Crucially, the mechanism responsible for linear growth in the length of the colony observed in experiment of Vulin et al. [19] cannot be explained with their [21, 36] modelling approach (see Figure 1b).

Although the cylindrical yeast colonies (Figure 1) have been observed to increase linearly with time, it is not appropriate to specify linear domain growth as domain growth is dependent on nutrient concentration. In this work, we aim to model the emergent macroscale response of both the growth exhibited within regions of the domain and the evolving overall length of the domain for a given input nutrient concentration. Importantly, our results show that if there is only a fixed number of replicative cells at the base of the colony (and all remaining cells are quiescent), then the colony length increases linearly with time, consistent with the experimental observations of Vulin et al. [19]. Additionally, we derive an accurate continuum approximation for the pathlines (averaged positions of labelled cells), providing a relationship between the microscopic and macroscopic behaviours of the uniaxial yeast growth process. Through the examination and analysis of the pathlines for six different types of input nutrient concentrations, we are able to determine the properties of growth at the macroscale within the expanding domain. More broadly, we provide a framework for classifying or determining certain regions of uniform and non-uniform growth in biological systems exhibiting uniaxial growth.

2. Agent based model

We develop a one-dimensional CA model to simulate the microscopic stochastic behaviour of individual cells that are proliferating in a uniaxial direction. Due to the stochastic nature of the model, any single simulation will result in a

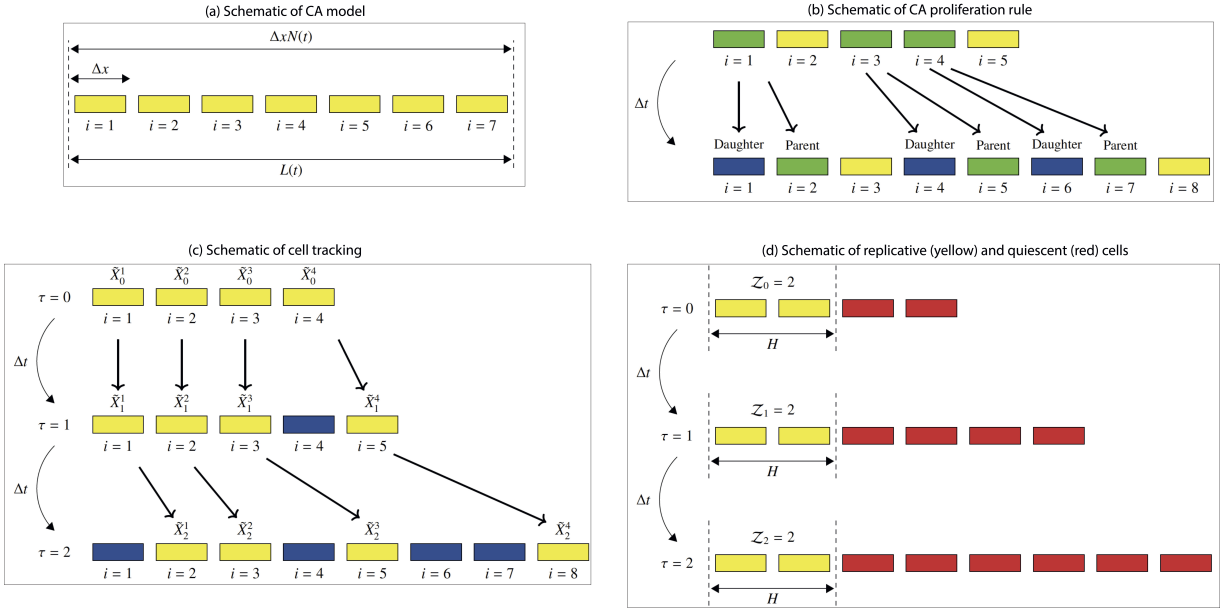

Figure 2: Schematics for the CA model. (a) One-dimensional CA model. The equal size cells (width Δx units) within the array are labelled by a dimensionless variable $i = 1, 2, \ldots, N(t)$, where $N(t)$ is the total number of cells at time t. The dimensional length of the domain at time t is then $L(t) = \Delta x N(t)$. (b) Proliferation rule. Three proliferation events have occurred in a single time step of the CA model. (c) Tracking $N_0 = 4$ initial cells after two time steps. The four trajectories are labelled $\tilde{X}^1_\tau$, $\tilde{X}^2_\tau$, $\tilde{X}^3_\tau$ and $\tilde{X}^4_\tau$. Cell trajectories do not intersect with each other and $\tilde{X}^4_\tau = \tilde{L}_\tau$. (d) Quiescent (red) and replicative (yellow) cells. The quiescent region is always assumed to be at the top of the colony. The number of cells in the replicative region is constant with $\mathcal{Z}_\tau = \mathcal{Z} = 2$, implying that the length of the replicative is fixed with $H = 2\Delta x$.

non-uniform spatial distribution for the labelled cells. However, Binder et al. [21] showed that the ensemble average of many CA simulations can lead to uniform growth at the macro-scale if cell proliferation is spatially uniform (implying that there is a constant nutrient concentration throughout the entire domain). Here, we extend their CA model to include position-dependent cell proliferation based on the non-constant nutrient availability in the domain. If cell proliferation is spatially varying, we expect the ensemble average of many CA simulations to result in a non–uniform growth at the macro-scale.

To model the uniaxial growth, we first define a one–dimensional array of discrete cells (see Figure 2a), where each of the discrete (labelled) cells in the array represent one biological cell. We assume that all the cells are the same size and are evenly spread across the array. This is equivalent to assuming the cells are incompressible. Each of the parent cells in the array may proliferate to give birth to daughter cells. Furthermore, we assume that there will be no cell motility and no cell death. These are reasonable assumptions for yeasts as the cells are sessile and cell death is negligible for the time-scale of the experiments [16]. The number of cells in the array, and length, will hence be increasing with time.

The width of a cell is denoted as Δx. At time t, each cell in the colony is labelled with the index i such that the physical position of the cell is given by $x = i\Delta x$. The leftmost cell is indexed as $i = 1$ and the second cell from the left is indexed $i = 2$ until the rightmost cell. The physical position x can be interpreted as the distance from the base of the colony. The length of the colony at time t is given by $L(t) = \Delta x N(t)$, where $N(t)$ is the total number of cells in the colony at time t (see Figure 2a). We impose that the initial number of cells in the colony is $N(0) = N_0$ and initial length of the colony is $L(0) = L_0$.

We choose to split up time into discrete time steps of size Δt, and define τ as the number of time steps since the start of the simulation. This means that the temporal variable can be written as $t = \tau\Delta t$, which will be continuous in the limit $\Delta t \to 0$. We define $\mathcal{T}$ to be the total number of time steps and $N(\tau\Delta t)$ to be the number of cells after τ time steps, noting the length after τ time steps will be $L_\tau = L(\tau\Delta t) = \Delta x N_\tau$. In each time step of length Δt, each of the cells will have the opportunity to proliferate. We introduce a proliferation probability p^i defined as the probability that the cell located at position i proliferates in a given time step. It is important to note that this probability p^i will depend on the

size of the time step. We also define a proliferation rate $\hat{p}^i$ as the rate in which a cell located at position i proliferates. As the proliferation rate $\hat{p}^i$ is a physical quantity depending on the biology of the system, it will not change as the size of the time step Δt varies. The proliferation probability and proliferation rate are related by $p^i = \hat{p}^i \Delta t$. At the τ^{th} time step we then have $p^i_\tau = \Delta t \hat{p}^i_\tau$.

We now define the cell proliferation rule for the CA model. As aforementioned, each of the cells may proliferate to produce a daughter cell in one time step of length Δt. If a cell in position i proliferates, the parent cell is displaced to the right, to position $i + 1$, and a daughter cell is inserted in the parent cell's original position i. The cells to the left of the proliferating cell remain in the same position while the cells right of the proliferating cell are displaced one position to the right. This displacement occurs because of the insertion of the daughter cell. Now consider the displacement of a cell when multiple proliferation events have occurred in one time step (see Figure 2b). A cell is displaced n positions to the right if n proliferation events have occurred to the left of it. Furthermore, if a cell is chosen to proliferate, its daughter cell will be birthed n positions to the right and the parent cell will be displaced $n + 1$ positions to the right if n proliferation events have occurred to the left of the parent cell.

We note that in our model, a newly born daughter cell is always able to displace all the cells to the right of it, no matter how large the colony. Hence, the length of the colony can grow indefinitely. In reality, we might expect that as the colony gets large, cells near $x = 0$ would need to produce a large force to achieve this displacement. This may result in the cells becoming compressed, which in turn may stop them proliferating. However, to the best of our knowledge, this phenomenon has not been reported in the experiments considered here, possibly because the duration of the experiment is not sufficiently long. Hence, we neglect the possible effect of compression on cell proliferation in this paper; we have considered this possibility in [41] (using a continuum model) and, in a different context, in a recent individual-based lattice-free model [42].

One of the main biological factors used to determine the proliferation rate is the nutrient concentration, $C(x, t)$, which is a function of both space and time. For simplicity, in this paper we simply prescribe the nutrient concentration. In our CA model, the concentration will have support on the spatial domain $x = \Delta x, 2\Delta x, \ldots, L_\tau$ and the temporal domain $t = \Delta t, 2\Delta t, \ldots, \mathcal{T}\Delta t$. Therefore we can define the nutrient concentration of the i-th cell after τ time steps by $C^i_\tau = C(i\Delta x, \tau\Delta t)$. For ease of interpretation, we assume that the rate of proliferation is proportional to the nutrient concentration: $\hat{p}^i_\tau = kC^i_\tau$, where k is a constant chosen such that the resulting probabilities $p^i_\tau = \Delta t k C^i_\tau$ are between zero and one. (Note that since the nutrient concentration is prescribed, this is equivalent to prescribing the proliferation rates / probabilities.) Having defined all the key parameters and mechanisms for the CA model, we now consider the trajectories of labelled cells.

2.1. Cell trajectories

We use the CA algorithm to track the trajectories of labelled cells within the colony. The cell trajectories can be defined as the path of an individual cell as the colony grows. These trajectories depend on the number of cell proliferation events and their positions, and are therefore stochastic. We thus expect the trajectories to be different in each simulation and are interested in the average trajectories, or pathlines. In particular, we are interested in the pathlines of the N_0 initial cells, as these allow us to analyse the spatial distribution of the cells, infer where in the colony the proliferation events are occurring and how this evolves over time. From these pathlines, we can determine whether uniform or non–uniform growth has occurred.

The pathlines of the N_0 initial cells in each simulation are shown in Figure 2c. Suppose we label the trajectory $\tilde{X}^j_\tau$ where τ is the number of time steps and j is the initial position of the tracked cell. At $t = 0$, we thus have $\tilde{X}^j_0 = j\Delta x$ for $j = 1, 2, \ldots, N_0$. We note that $\tilde{X}^j_\tau$ is the dimensional position in terms of physical length. We also observe that the trajectory $\tilde{X}^{N_0}_\tau = \tilde{L}_\tau$ is the length of one simulation.

A naïve method of estimating the pathlines would be to run multiple simulations and calculate an ensemble average for the trajectory. However, this is not only computationally expensive, but it also precludes further closed-form analysis. We therefore consider an alternative method of obtaining accurate continuum approximations for the pathlines of the initial cells, which connects the stochastic microscopic behaviour to the macro-scale colony growth.

We derive an expression for the evolution of the pathlines over time by considering cell displacement in a single realisation of the CA model. The cell displacement depends on the number of cell proliferations to the left of it. Specifically, the cell displacement will be $n\Delta x$, where n is the number of cell proliferations to the left of, and including, the specific cell. We thus need to determine the value of n for each of the initial cells at each time step. Recall that $\tilde{X}^j_\tau$

is the trajectory of the cell initially in position j. Suppose we define n_τ^j as the number of cells proliferating to the left of, and including, $\tilde{X}_\tau^j$ during time step τ. We can hence derive the difference equation for one simulation to be

$$\underbrace{\tilde{X}_{\tau+1}^j}_{\substack{\text{Location of cell at}\\ \text{next time step}}} = \underbrace{\tilde{X}_\tau^j}_{\substack{\text{Location of cell at}\\ \text{current time step}}} + \underbrace{n_\tau^j \Delta x.}_{\substack{\text{Displacement during}\\ \text{time step } \tau}} \tag{1}$$

As the CA model is stochastic, the quantity n_τ^j is a random variable and hence we can derive the average trajectories, or pathlines, by using the expected value of n_τ^j. The difference equation for the pathlines of the initial cells is then

$$X_{\tau+1}^j = X_\tau^j + \Delta x \mathbb{E}\left[n_\tau^j\right], \tag{2}$$

where $X_\tau^j = \mathbb{E}\left[\tilde{X}_\tau^j\right]$ is the expected value of the cell initially in position j after τ time steps. We note that $\tilde{X}_\tau^j$ can only be a multiple of Δx but the average trajectory X_τ^j can be any real positive number. Intuitively, $\mathbb{E}\left[n_\tau^j\right]$ will depend on the proliferation probabilities p_τ^i. Suppose we refer to a cell proliferating as a success and a cell not proliferating as a failure. We can therefore view a cell proliferation as a Bernoulli trial with success probability p_τ^i, which means the expected value of cell proliferation for a cell with proliferation probability p_τ^i is simply p_τ^i. We also note that each of the cells proliferate independently of each other, so we can view the n_τ^j as a sum of independent Bernoulli trials. The expected value for n_τ^j is thus the cumulative sum of the proliferation probabilities up to, and including, the cell X_τ^j,

$$\mathbb{E}\left[n_\tau^j\right] = \sum_{i=1}^{X_\tau^j/\Delta x} p_\tau^i. \tag{3}$$

It is important to note that the i in the expression for the proliferation probabilities refers to the dimensionless index of the cell. As X_τ^j is the dimensional position of the cell, we need to convert it to a nondimensional value. This is why we divide X_τ^j by Δx in Equation (3). We can hence simplify the difference equation for the pathlines

$$X_{\tau+1}^j = X_\tau^j + \Delta x \sum_{i=1}^{X_\tau^j/\Delta x} p_\tau^i. \tag{4}$$

2.2. Replicative and quiescent regions

As the proliferation probabilities depend on the nutrient concentration in the colony, $p_\tau^i = k\Delta t C_\tau^i$, we expect to see different pathlines for different nutrient concentration profiles. As we shall see, a constant nutrient concentration leads to uniform growth and a non–constant nutrient concentration leads to non–uniform growth at the macro-scale.

Motivated by the experimental observations of Vulin et al. [19], we consider the situation when there is a replicative region and quiescent region of cells (see Figure1a). In the replicative region, cells have access to a non-zero nutrient concentration and can proliferate, whereas cells in the quiescent region do not, due to a lack of nutrient (*i.e.*, in the quiescent region $p_\tau^i = k\Delta t C_\tau^i = 0$). The size of the replicative region is denoted by the dimensionless variable $\mathcal{Z}$, representing the number of cells. The physical length of the replicative region is then $H = \mathcal{Z}\Delta x$ (see Figures 1a & 2d).

We now consider how the size of the replicative region evolves with time by defining $\mathcal{Z}_\tau$ as the number of cells in the region at time step τ, with the length being given by $H_\tau = \mathcal{Z}_\tau \Delta x$. Intuitively, we require $\mathcal{Z}_\tau \leq N_\tau$ (and equivalently $H_\tau \leq L_\tau$) for all time steps τ. Other than this inequality, we may arbitrarily define $\mathcal{Z}_\tau$. In the following section, we consider a number of prototypical scenarios (Cases I-VI) in which the nutrient distributions and the lengths of the proliferative regions are prescribed in ways which qualitatively re-create a number of possible biological scenarios, but are sufficiently simple to be analytically tractable.

3. Results

We choose to focus on the results for six specific cases (see Table 1) to demonstrate how the CA model can identify regions of uniform and nonuniform growth by determining and examining the pathlines in each case. The main input

	Constant nutrient concentration in replicative region	Linearly decreasing nutrient concentration in replicative region
No quiescent region, $\mathcal{Z}_\tau = N_\tau$	**I**	**II**
Fixed height of replicative region, $\mathcal{Z}_\tau = \mathcal{Z} \leq N_0$	**III**	**IV**
Variable height of replicative region, $\mathcal{Z}_\tau = hN_\tau \leq N_\tau$	**V**	**VI**

Table 1: The six modelling scenarios, **I–IV**

into the CA model is the cell proliferation probabilities, $p^i_\tau = \Delta t k C^i_\tau$, which are proportional to the prescribed nutrient concentration, $C^i_\tau = C(i\Delta x, \tau\Delta t)$. For cases **I** and **II** there is no quiescent region, with $\mathcal{Z}_\tau = N_\tau$ and $H_\tau = L$. These situations could represent the early growth of a yeast colony, when it is still small enough that all cells have access to nutrient. Cases **III** and **IV** correspond to situations where the number of cells in the replicative region is constant at all times, and we set $\mathcal{Z}_\tau = \mathcal{Z}$, implying that the length of the proliferative region is constant, with $H_\tau = H$. These cases mimic the situation in a larger colony where nutrient is transported by diffusion and consumed by the cells, hence only being present in appreciable concentrations in a fixed region at the bottom of the colony. For cases **V** and **VI** the variable length of the replicative region is a constant fraction of the overall length, $H_\tau = hL_\tau$, implying that the number of cells in the replicative region is increasing with time, so $\mathcal{Z}_\tau = hN_\tau$, where $0 < h < 1$. This could represent a situation in which nutrient is transported mainly by advection.

In the replicative regions of these six cases, we assume for simplicity that the nutrient concentration is either constant (**I**, **III** & **V**) or linear (**II**, **IV** & **VI**), with the corresponding cell proliferation probabilities for each of the six cases illustrated in Figure 3. Whilst these two nutrient profiles are not biologically realistic, their simplicity facilitates analysis and comparison with previous work *e.g.* [21]. The linearly decreasing distribution provides a qualitative representation of the expected monotonically decreasing nutrient concentration seen in Figure 1a.

3.1. The length of the colony

We begin by deriving some general results regarding the length of the colony. We recall from earlier that Vulin et al. [19] observed the yeast colony's length to grow linearly in time; however, in other models of biological growth, exponential growth of the domain is often found, or assumed.

The basic equation for the length of the colony is (4) with $j = N_\tau$. Recalling that $X^{N_\tau}_\tau = L_\tau$ we have

$$L_{\tau+1} = L_\tau + \Delta x \sum_{i=1}^{N_\tau} p^i_\tau. \tag{5}$$

It is also useful at this point to define the mean proliferation probability, and the mean proliferation rate within the proliferative region at timestep, τ, which are respectively given by:

$$\overline{p}_\tau = \frac{1}{\mathcal{Z}_\tau} \sum_{i=1}^{\mathcal{Z}_\tau} p^i_\tau, \qquad \overline{\hat{p}}_\tau = \frac{1}{\mathcal{Z}_\tau} \sum_{i=1}^{\mathcal{Z}_\tau} \hat{p}^i_\tau = \frac{1}{\Delta t} \overline{p}_\tau. \tag{6}$$

Since we have $p^i_\tau = 0$ for $i > \mathcal{Z}_\tau$ (there is no proliferation outside the proliferative zone) then we can re-write equation (5) as

$$L_{\tau+1} = L_\tau + \Delta x \mathcal{Z}_\tau \overline{p}_\tau = L_\tau + \Delta t \Delta x \mathcal{Z}_\tau \overline{\hat{p}}_\tau. \tag{7}$$

We now consider two important cases.

Proliferative region of length proportional to the colony length

Let us first assume that the proliferative region is proportional in length to the length of the entire colony. In this case, $Z_\tau = hN_\tau$ for some constant value, h, or equivalently, the length of the replicative region H_τ is given by $H_\tau = hL_\tau = \Delta x h N_\tau = \Delta x \mathcal{Z}_\tau$. The, equation (7) can be written as

$$L_{\tau+1} = L_\tau + \ hL_\tau \overline{p}_\tau. \tag{8}$$

Provided that $\overline{p}_\tau$ is independent of time (*i.e.* $\overline{p}_\tau = \overline{p}$ (constant)) - which is true for cases I, II, V and VI - then (8) has solution

$$L_\tau = L_0(1 + h\overline{p})^\tau. \tag{9}$$

Alternatively, since we equivalently have $\overline{\hat{p}}_\tau = \overline{\hat{p}}$ (*i.e.* independent of time) we can write (7) as

$$\frac{L_{\tau+1} - L_\tau}{\Delta t} = hL_\tau\overline{\hat{p}}. \tag{10}$$

On recalling that $L_\tau = L(\tau\Delta t) = L(t)$ and taking the limit $\Delta t \to 0$ we obtain the continuous time approximation

$$\frac{dL}{dt} = h\overline{\hat{p}}L \quad \Rightarrow \quad L(t) = L_0 e^{h\overline{\hat{p}}\,t}. \tag{11}$$

Hence, in this case, equations (9) and (11) clearly show the growth in the length of the colony is exponential, and its rate depends on the size of the proliferative region and the mean rate of proliferation. This generalises the results of *e.g.* [21] which demonstrated that uniform proliferation gives rise to exponential growth of the length of the tissue. Here, we see that the result still holds when proliferation can vary in space and time, as long as the mean proliferation rate is time-independent.

Proliferative region of fixed length

We now consider the situation where the proliferative region is of fixed length. This implies that $Z_\tau = \mathcal{Z}$ (constant) o $H_\tau = H$ (constant). Using the fact that $H = \Delta x\mathcal{Z}$ equation (7) now becomes

$$L_{\tau+1} = L_\tau + H\,\overline{p}_\tau = L_\tau + \Delta t H\overline{\hat{p}}_\tau. \tag{12}$$

We assume as before that the mean proliferation probability and rate are independent of time (which is true for cases III and IV) and write $\overline{p}_\tau = \overline{p}$ and $\overline{\hat{p}}_\tau = \overline{\hat{p}}$). Then, proceeding as above, we obtain the following discrete and continuous time approximations for the colony length

$$L_\tau = L_0 + \overline{p}H\tau, \tag{13}$$

$$L(t) = L_0 + H\overline{\hat{p}}\,t. \tag{14}$$

In this case, equations (13) and (14) show that the colony length grows linearly, which is consistent with the observations of Vulin et al. [19].

In the following section, we consider the six proliferation scenarios in detail, including the determination of the pathlines for the cells.

*3.2. Case **I**: No quiescent region and constant nutrient concentration*

We first consider the situation of a constant nutrient concentration with no quiescent region. This corresponds to a physical system in which an unlimited nutrient supply can diffuse fast enough (relative to the rate of consumption by the cells) to maintain a constant nutrient concentration, $C^i_\tau = C$, throughout the entire domain. It then follows that the proliferation rate will be constant, $\hat{p}^i_\tau = \hat{p}$, and as the size of the time step Δt is fixed, the cell proliferation probabilities will also be constant,

$$p^i_\tau = p \qquad \text{for } i = 1 \ldots N_\tau. \tag{15}$$

(see Figure 3a). The difference equation (4) then becomes

$$X^j_{\tau+1} = X^j_\tau + \Delta x \sum_{i=1}^{X^j_\tau/\Delta x} p = (1 + p)\,X^j_\tau, \tag{16}$$

which can be solved to give the discrete expression

$$X^j_\tau = X^j_0\,(1 + p)^\tau\,, \tag{17}$$

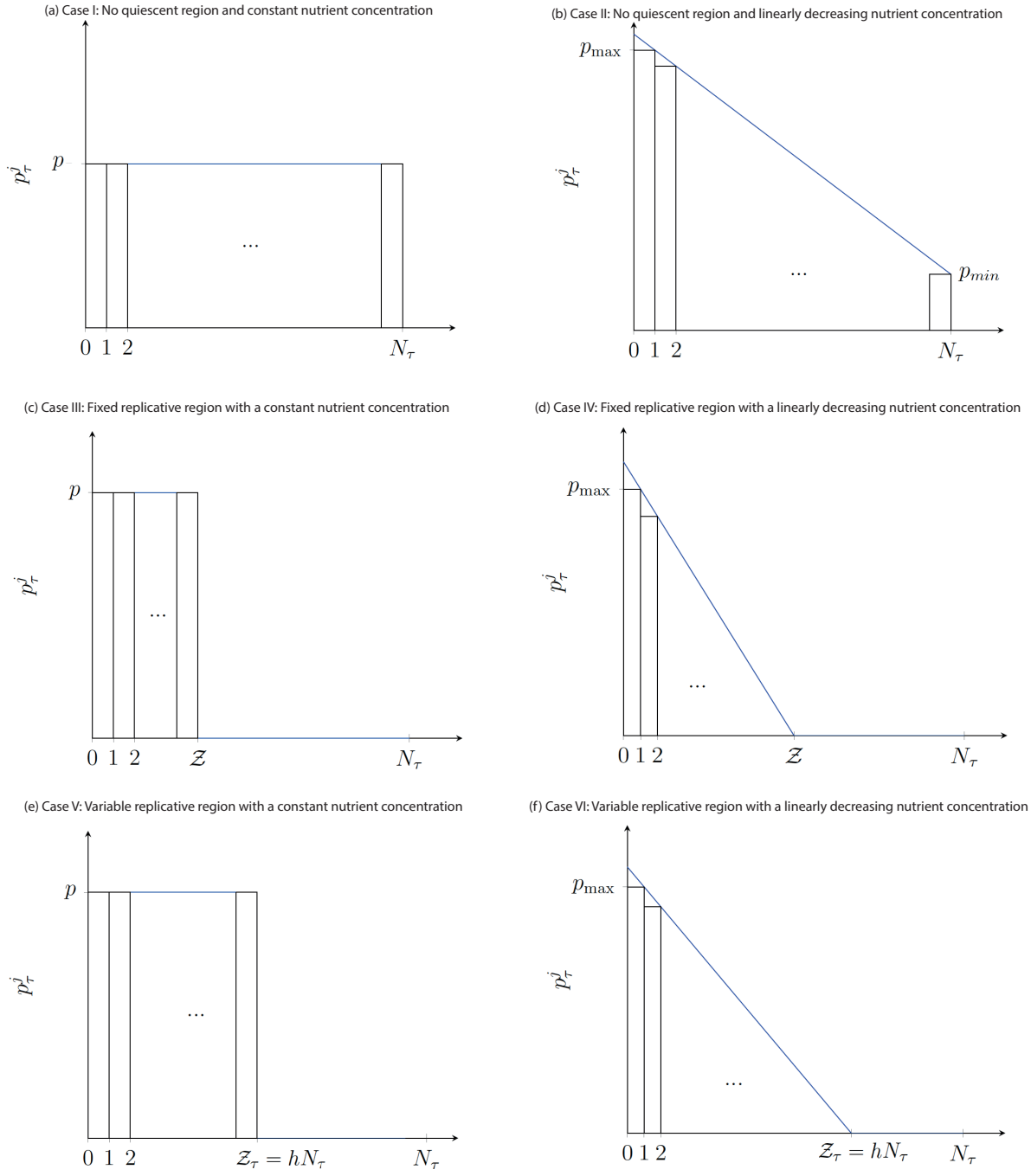

Figure 3: Schematics for the cell proliferation probabilities, $p_\tau^j = \Delta t k C_\tau^j$, for the six modelling scenarios. (a) Case **I**: No quiescent region and constant nutrient concentration. (b) Case **II**: No quiescent region and linearly decreasing nutrient concentration. (c) Case **III**: Fixed length replicative region with a constant nutrient concentration. (d) Case **IV**: Fixed length replicative region with a linearly decreasing nutrient concentration. (e) Case **V**: Variable length replicative region with a constant nutrient concentration. (d) Case **VI**: Variable length replicative region with a linearly decreasing nutrient concentration.

for the pathlines. The length of the domain is then

$$L_\tau = X_\tau^{N_0} = L_0 (1 + p)^\tau . \tag{18}$$

We note that the solution to the difference equation (17) is discrete in time, and is given in terms of the uniform proliferation probability, p. Suppose we now wish to express this in terms of continuous time and uniform proliferation rate, $\hat{p}$. We recall that our continuous time will be given by $t = \Delta t\tau$ and that $p = \hat{p}\Delta t$. We can therefore obtain a *continuum path* for each pathline by taking the limit as the time step $\Delta t \to 0$,

$$X^j(t) = \lim_{\Delta t \to 0} X_\tau^j = \lim_{\Delta t \to 0} X_0^j (1 + \hat{p}\Delta t)^{t/\Delta t} = X_0^j e^{\hat{p}t}, \tag{19}$$

with the length of the domain being given by

$$L(t) = L_0 e^{\hat{p}t}. \tag{20}$$

It is important to note that the continuum paths (19) will give the best agreement with their discrete equivalent when the change in length for each proliferation event is small (i.e. as $\Delta x \to 0$) when compared to X_τ^j. This is equivalent to the limit $X_\tau^j/\Delta x \to \infty$ and will occur when there is a large number of cells in the domain.

In Figure 4, we plot both the continuum paths (solid curves) given by (19) and average cell trajectories (markers) from 1000 simulations of the CA model, for $N_0 = 100$ initial cells and constant cell proliferation probability, $p = 0.05$. The comparison validates our derivation of the continuum paths, and we observe exponential growth at the macro-scale. This makes intuitive sense as every cell in the domain has equal opportunity to proliferate at each time step. Furthermore, we observe that the distance between adjacent continuum paths is equal at any fixed time. We also note that this equidistance is increasing with time. Importantly, we may use this increasing equidistance between adjacent continuum paths to classify uniform growth in the domain.

As previously mentioned, our results are consistent with those of Binder et al. [21], but there is a fundamental difference in the two modelling approaches. In our work, we assume a known constant nutrient concentration (or constant cell proliferation probability) across entire domain which produces the emergent property of uniform exponential growth at the macro-scale. In contrast, Binder et al. [21] assumed the length of the domain was known and growing exponentially to determine the number of cell proliferation events at any time step. This gives the same emergent property of uniform growth at the macro-scale, and implicitly determines the constant cell proliferation probability (or constant nutrient concentration). In essence, our approach may be considered as the forward problem of determining properties of the growth for a given constant nutrient concentration. The approach of Binder et al. [21] is then the inverse problem of determining the constant value of the nutrient concentration for known properties of the growth.

We now turn our attention to the problem of a non-constant (linearly decreasing) nutrient concentration with no quiescent region.

*3.3. Case **II**: No quiescent region and linearly decreasing nutrient concentration*

In this example we consider a linearly decreasing nutrient concentration, implying that $C_\tau^i \geq C_\tau^j$ if $i < j$, in a domain with no quiescent region. It is assumed that the flux (from an unlimited nutrient source) and nutrient diffusion properties within the domain can support constant maximum and minimum values of the nutrient concentration at the two boundaries of the domain, for all time. As the nutrient concentration C_τ^i is linearly decreasing, the proliferation rates $\hat{p}_\tau^i$, and hence the proliferation probabilities p_τ^i, will also be linearly decreasing. Therefore, for all time steps τ, the maximum proliferation probability will be $p_\tau^1 = p_{\max}$ and the minimum probability will be $p_\tau^{N_\tau} = p_{\min}$, with the corresponding proliferation rates $\hat{p}_\tau^1 = \hat{p}_{\max} = p_{\max}/\Delta t$ and $\hat{p}_\tau^{N_\tau} = \hat{p}_{\min} = p_{\min}/\Delta t$ (see Figure 3b). The linearly decreasing proliferation probabilities can then be defined by

$$p_\tau^i = p_{\max} + \frac{p_{\max} - p_{\min}}{N_\tau - 1} - \frac{p_{\max} - p_{\min}}{N_\tau - 1} i \qquad \text{for } i = 1 \ldots N_\tau. \tag{21}$$

In order to determine the continuum paths for the labelled cells, we first substitute the probability function (21) into the difference equation (4), and use the relationship $\Delta x\,(N_\tau - 1) = L_\tau - \Delta x$ to obtain:

$$X_{\tau+1}^j = X_\tau^j + p_{\max} X_\tau^j - \left(X_\tau^j - \Delta x\right) \frac{p_{\max} - p_{\min}}{2\Delta x\,(N_\tau - 1)} X_\tau^j = X_\tau^j + p_{\max} X_\tau^j - \left(X_\tau^j - \Delta x\right) \frac{p_{\max} - p_{\min}}{2\,(L_\tau - \Delta x)} X_\tau^j. \tag{22}$$

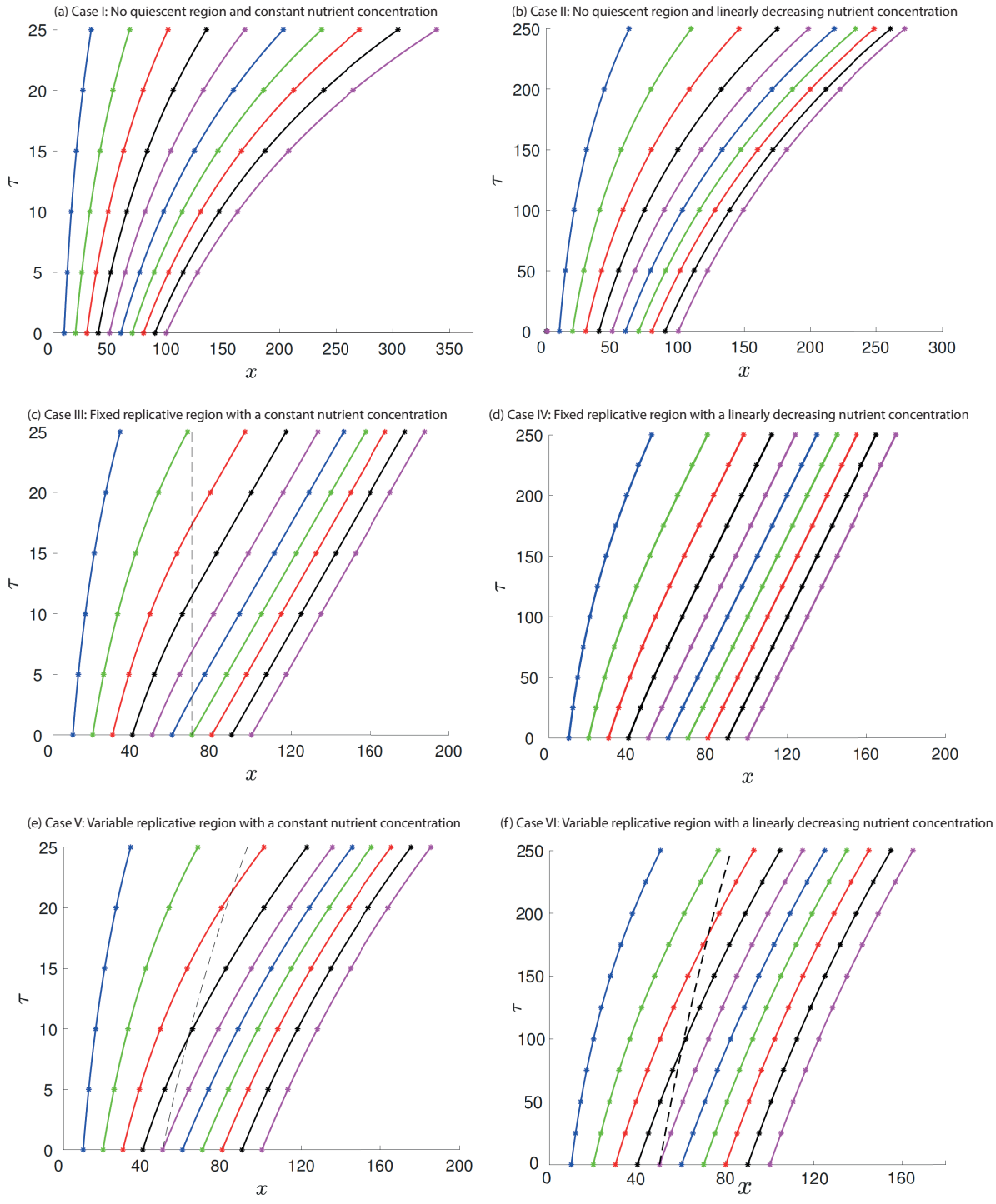

Figure 4: Pathlines (solid curves) for the four modelling scenarios, $N_0 = 1000$, $\Delta x = 0.1$, $L_0 = 100$, $\Delta t = 1$. The markers are averaged results from 1000 simulations of the CA model. (a) Case **I**: No quiescent region and constant nutrient concentration, with $p = 0.05$. (b) Case **II**: No quiescent region and linearly decreasing nutrient concentration, $p_{max} = 0.008$ and $p_{min} = 0$. (c) Case **III**: Fixed replicative region, $\mathcal{Z} = 700$ or $H = 70$ (broken vertical line), with a constant nutrient concentration, $p = 0.05$. (d) Case **IV**: Fixed replicative region, $\mathcal{Z} = 750$ or $H = 75$ (broken vertical line), with a linearly decreasing nutrient concentration, $p_{max} = 0.008$. (e) Case **V**: Variable replicative region, $\mathcal{Z}_\tau = hN_\tau$ and $h = \frac{1}{2}$ (broken curve), with a constant nutrient concentration, $p = 0.05$. (f) Case **VI**: Variable replicative region, $\mathcal{Z}_\tau = hN_\tau$ and $h = \frac{1}{2}$ (broken curve), with a linearly decreasing nutrient concentration, $p_{max} = 0.008$.

We note that (22) is dependent on the length of the colony L_τ. From equation (9) we have:

$$L_\tau = L_0\left(1 + \frac{p_{\max} + p_{\min}}{2}\right)^\tau, \tag{23}$$

whilst (11) yields the continuum path as

$$L(t) = L_0 \exp\left\{\frac{\hat{p}_{\max} + \hat{p}_{\min}}{2} t\right\}. \tag{24}$$

To determine if the growth is uniform or non-uniform, we need to solve (22) for the remaining pathlines, X_τ^j using the expression for L_τ above. Unfortunately, the difference equation for the pathlines cannot be solved analytically. We thus choose to take the continuum limit to find a corresponding differential equation for the continuum paths. First we rearrange equation (22) and divide by Δt to find

$$\frac{X_{\tau+1}^j - X_\tau^j}{\Delta t} = \hat{p}_{\max} X_\tau^j - \left(X_\tau^j - \Delta x\right)\frac{\hat{p}_{\max} - \hat{p}_{\min}}{2\left(L_\tau - \Delta x\right)} X_\tau^j, \tag{25}$$

and taking the continuum limits $\Delta x \to 0$ and $\Delta t \to 0$ gives

$$\frac{dX^j}{dt} = \hat{p}_{\max} X^j - \frac{\hat{p}_{\max} - \hat{p}_{\min}}{2}\frac{\left(X^j\right)^2}{L(t)} = \hat{p}_{\max} X^j - \frac{\hat{p}_{\max} - \hat{p}_{\min}}{2L_0} \exp\left\{-\frac{\hat{p}_{\max} + \hat{p}_{\min}}{2} t\right\}\left(X^j\right)^2, \tag{26}$$

where we substituted for $L(t)$ using (24). Equation (26) is a Bernouilli equation with the closed form solution

$$X^j(t) = \frac{L_0 \exp\left\{\frac{\hat{p}_{\max} + \hat{p}_{\min}}{2} t\right\}}{1 + \left[\frac{L_0}{X_0^j} - 1\right]\exp\left\{-\frac{\hat{p}_{\max} - \hat{p}_{\min}}{2} t\right\}}, \tag{27}$$

for the continuum paths.

The continuum paths (27) are plotted in Figure 4(b) and illustrate the key characteristic of non-uniform growth. At any fixed non-zero time the distances between adjacent paths are no longer equal as was previously found in case **I**, but are instead decreasing in the direction of increasing x. This observation is consistent with the linearly decreasing nutrient concentration. However, we observe each distance between any two adjacent paths is increasing exponentially with time. Therefore, we classify the emergent macro-scale growth as non-uniform exponential growth.

The results for case **II** are consistent with those of Lai De Oliveira and Binder [36], who obtained the continuum paths numerically, unlike the closed-form formulas for the paths found in our work. Furthermore, and similar to the approach of Binder et al. [21], Lai De Oliveira and Binder [36] assumed that the length of domain was known in advance, and this approach differs to the one presented in this work, where the length of the domain emerges in response to the given input nutrient concentration.

*3.4. Case **III**: Fixed replicative region with a constant nutrient concentration*

We now consider the situation of a fixed length replicative region, $H = \mathcal{Z}\Delta x$, with a constant nutrient concentration, and a quiescent region that increases in length, $L(t) - H$, as time increases. We hence prescribe the proliferation probabilities as

$$p_\tau^i = \begin{cases} p & \text{if } i \le \mathcal{Z}, \\ 0 & \text{if } i > \mathcal{Z}, \end{cases} \tag{28}$$

(see Figure 3c).

The pathlines within the replicative region are obtained in a similar way to case **I**, giving

$$X_\tau^j = X_0^j (1 + p)^\tau \qquad \text{for } X_\tau^j \le H. \tag{29}$$

For cells in the quiescent region, equation (4) simplifies to

$$X_{\tau+1}^j = X_\tau^j + p\mathcal{Z}\Delta x \qquad \text{for } X_0^j > H, \tag{30}$$

which can be solved to yield

$$X_\tau^j = pH\tau + X_0^j \qquad \text{for } X_0^j > H. \tag{31}$$

However, equations (29) & (31) do not account for the situation when cells cross over from the replicative region to the quiescent region.

Let's suppose we define the value $\tau_j^\star$ as the time at which the cell with pathline X_τ^j crosses from the replicative region to the quiescent region. As the cell is initially in the replicative region, its pathline for $\tau < \tau_j^\star$ will be given by equation (29). We can find $\tau_j^\star$ by solving

$$X_0^j (1+p)^{\tau_j^\star} = H, \tag{32}$$

because H is the location at which the cell's pathline crosses into the quiescent region. Solving (32) gives

$$\tau_j^\star = \frac{\log\left(H/X_0^j\right)}{\log(1+p)}. \tag{33}$$

For $\tau \geq \tau_j^\star$, the cell will be in the quiescent region. The pathline for $\tau \geq \tau_j^\star$ will thus be of the form

$$X_\tau^j = pH\tau + c, \tag{34}$$

where c is a constant that accounts for the time taken for the cell to leave the replicative region. We can calculate c by considering the point $\left(H, \tau_j^\star\right)$ where the pathline transitions from the replicative region to the quiescent region. Substituting $\left(X_\tau^j, \tau\right) = \left(H, \tau_j^\star\right)$ into (34) and rearranging, we find that $c = H - pH\tau_j^\star$. The pathline for $\tau \geq \tau_j^\star$ of an initially replicative cell that has crossed into the quiescent region at time $\tau = \tau_j^\star$ is thus

$$X_\tau^j = pH\left(\tau - \tau_j^\star\right) + H. \tag{35}$$

The full expression for the pathlines is then

$$X_\tau^j = \begin{cases} X_0^j (1+p)^\tau, & \text{if } X_0^j \leq H, \tau < \tau_j^\star, \\ pH\left(\tau - \tau_j^\star\right) + H, & \text{if } X_0^j \leq H, \tau \geq \tau_j^\star, \\ pH\tau + X_0^j, & \text{if } X_0^j > H. \end{cases} \tag{36}$$

Once again, we wish to derive continuum paths for the pathlines of the labelled cells. Recalling that $t = \tau\Delta t$ and $p = \hat{p}\Delta t$, the continuum paths are obtained by taking the limit as $\Delta t \to 0$ in (36), yielding

$$X^j(t) = \begin{cases} X_0^j \exp\{\hat{p}t\}, & \text{if } X_0^j \leq H, t < t_j^\star, \\ \hat{p}H\left(t - t_j^\star\right) + H, & \text{if } X_0^j \leq H, t \geq t_j^\star, \\ \hat{p}Ht + X_0^j, & \text{if } X_0^j > H, \end{cases} \tag{37}$$

where

$$t_j^\star = \frac{\log\left(H/X_0^j\right)}{\hat{p}}. \tag{38}$$

Note that $t_j^\star$ is found in a similar way to $\tau_j^\star$, by equating $X^j(t_j^\star) = H$.

Equation (37) then gives the length,

$$L(t) = X^{N_0}(t) = \hat{p}Ht + X_0^{N_0} = \hat{p}Ht + L_0, \tag{39}$$

of the domain. Notably, the length is increasing linearly with time which is in contrast to the exponential growth found in case **I** (see Figures 4a and 4c). Intuitively, this result may at first appear surprising, as both cases **I** & **III** have the same constant nutrient concentration in the replicative region of the domain. The important difference is that the nutrient concentration is only non-zero in the fixed length replicative region of case **III**, and this leads to the emergent

property of linear growth in the overall length of the domain. This is due to the fact that the proportion of the length of the replicative region to the overall length of the domain is decreasing as time increases.

Further properties on the type of growth within the domain can be inferred from examining the space-time diagram in Figure 4(c). Consistent with the results shown in Figure 4(a), we observe uniform exponential growth in the replicative region of Figure 4(c) (to the left of the vertical broken line). As expected, the distances between adjacent continuum paths for cells initially located in the quiescent region are equal and do not increase with time (to the right of the green continuum path that intersects with the vertical broken line and x axis), an important characteristic which indicates that no growth has occurred in this portion of the quiescent region. Also, the distances between adjacent continuum paths that have transitioned into the quiescent region (to the right of the vertical broken line and to the left of the green continuum path that intersects with the vertical broken line and x axis) do not increase with time, but at any fixed time, these distances are decreasing in the direction of x, indicating non-uniform growth (see Figure 4b and analysis in Section 3.3). Therefore, we have the global property of non-uniform linear growth of the domain, with uniform exponential growth in the replicative region and no growth in the quiescent region.

While it is tempting to infer that the analysis can explain the observed linear growth in the experiments of Vulin et al. [19], we remind ourselves of the fact that the nutrient concentration is not decreasing linearly in this case. This then motivates us to consider our next example.

*3.5. Case **IV**: Fixed replicative region with a linearly decreasing nutrient concentration*

In this example, we consider a fixed replicative region, $H = \Delta x \mathcal{Z}$, with a linearly decreasing nutrient concentration to approximate the monotonically decreasing nutrient concentration in the experiments of Vulin et al. [19] (see Figures 1a & 3d). We recall that the nutrient concentration is zero in the quiescent region with corresponding cell proliferation probabilities $p_\tau^i = 0$ for $i > \mathcal{Z}$. For a nutrient concentration that is continuous across the entire the domain we need to set $p_\tau^{\mathcal{Z}} = p_{\min} = 0$ at the end of the replicative region which is adjacent to the quiescent region. At the other end of the replicative region (i.e. the base of the colony), with maximum nutrient concentration, we choose $p_\tau^1 = p_{\max}$, with the corresponding proliferation rate $\hat{p}_{\max}$. The linear cell proliferation probability function is then

$$p_\tau^i = \begin{cases} p_{\max} + \frac{p_{\max}}{\mathcal{Z}-1} - \frac{p_{\max}}{\mathcal{Z}-1} i & \text{if } i \le \mathcal{Z} \\ 0 & \text{if } i > \mathcal{Z}, \end{cases} \tag{40}$$

(see Figure 3d).

As we done previously, the aim now is to determine and continuum paths, so we substitute (40) into the difference equation (4) and obtain

$$X_{\tau+1}^j = X_\tau^j + \Delta x \sum_{i=1}^{X_\tau^j/\Delta x} \left(p_{\max} + \frac{p_{\max}}{\mathcal{Z}-1} - \frac{p_{\max}}{\mathcal{Z}-1} i \right). \tag{41}$$

For the replicative region we have $i \le \mathcal{Z}$ and $X_\tau^j \le H$, and simplifying equation (41) then gives

$$X_{\tau+1}^j = X_\tau^j + X_\tau^j p_{\max} - \frac{\left(X_\tau^j - \Delta x\right) p_{\max}}{2\left(H - \Delta x\right)} X_\tau^j \qquad \text{for } X_\tau^j \le H, \tag{42}$$

which is of the same form as (22) in Section 3.3. Therefore, the solution for the continuum paths in the replicative region is

$$X^j(t) = \frac{2H}{1 + \left[\frac{2H}{X_0^j} - 1\right] \exp\{-\hat{p}_{\max} t\}} \qquad \text{for } X_\tau^j \le H. \tag{43}$$

For the quiescent region we have $i > \mathcal{Z}$ and $X_\tau^j > H$. We can thus simplify (41) to obtain

$$X_{\tau+1}^j = X_\tau^j + \Delta x \sum_{i=1}^{\mathcal{Z}} \left(p_{\max} + \frac{p_{\max}}{\mathcal{Z}-1} - \frac{p_{\max}}{\mathcal{Z}-1} i \right) = X_\tau^j + H \frac{p_{\max}}{2}, \tag{44}$$

with solution

$$X_\tau^j = H \frac{p_{\max}}{2} \tau + X_0^j. \tag{45}$$

Taking the limits as $\Delta t \to 0$ and $\Delta x \to 0$, we obtain

$$X^j(t) = H\frac{\hat{p}_{\max}}{2}t + X_0^j, \tag{46}$$

for the continuum paths in the quiescent region, with the length of the domain being given by

$$L(t) = H\frac{\hat{p}_{\max}}{2}t + L_0, \tag{47}$$

which shows that the domain is elongating linearly.

Similarly to case **III**, we need to consider the situation when cells from the replicative region transition into the quiescent region at some time $t_j^\star$, defined by $X^j(t_j^\star) = H$. Using (43), we find that

$$t_j^\star = \frac{-1}{\hat{p}_{\max}}\log\left(\frac{X_0^j}{2H - X_0^j}\right) \qquad \text{for } X_0^j \le H. \tag{48}$$

We again note that the pathlines of cells initially in the replicative region that transition into quiescent region will be of the form

$$X^j(t) = H\frac{\hat{p}_{\max}}{2}t + c \qquad \text{for } t \ge t_j^\star. \tag{49}$$

We can determine the constant, $c = H - H\frac{\hat{p}_{\max}}{2}t_j^\star$, by again noting that the path must pass through the point $\left(H, t_j^\star\right)$.

Hence, the continuum paths are given by

$$X^j(t) = \begin{cases} \frac{2H}{1+\left(\frac{2H}{X_0^j}-1\right)\exp\{-\hat{p}_{\max}t\}}, & \text{if } X_0^j \le H, t < t_j^\star, \\ H\frac{\hat{p}_{\max}}{2}\left(t - t_j^\star\right) + H, & \text{if } X_0^j \le H, t \ge t_j^\star, \\ H\frac{\hat{p}_{\max}}{2}t + X_0^j, & \text{if } X_0^j > H. \end{cases} \tag{50}$$

We have just seen that the domain is increasing linearly with time (47), consistent with what was observed in the experiments of Vulin et al. [19] (see Figure 1b). We also recall that in this case the nutrient concentration is decreasing linearly over a fixed replicative region of the domain (see Figure 3d), approximating the hypothesised monotonically decreasing nutrient concentration in the experimental work of Vulin et al. [19] (see Figure 1b). Therefore, we believe case **IV** is a suitable model for yeast colonies growing in an uniaxial direction, validating the hypothesis of Vulin et al. [19].

To establish the properties of growth within the domain (i.e. the yeast colony), we examine the space-time diagram in Figure 4(d), for the continuum paths (50). Similar features are found to that observed in case **III**, there is non-uniform growth across the entire domain and no growth in the quiescent region of the domain. The main difference between cases **IV** & **III** is in the replicative region of the domain, where we see that there is non-uniform (to the left of the vertical broken line in Figure 4d) instead of uniform growth (to the left of the vertical broken line in Figure 4c), respectively. Therefore, we classify the growth as non-uniform linear growth, with non-uniform growth in the replicative region and no growth in the quiescent region.

*3.6. Cases **V** & **VI**: Variable replicative region with a constant or linearly decreasing nutrient concentration*

In our last two cases, we consider the scenario of when the replicative region is a fraction or proportion of the overall length of the domain with either a constant, case **V**, or linearly decreasing, case **VI**, nutrient concentration in the now variable length replicative region (see Figures 3e & 3f). In both cases, we prescribe the number of replicative cells in the region to be $\mathcal{Z}_\tau = hN_\tau$, where $h \in [0, 1]$ is the constant of proportionality. The length of the replicative region is then $H_\tau = hL_\tau$. Note that a value of $h = 0$ implies that there are no replicative cells and no growth can occur in the domain. However, if $h = 1$ then all the cells in the domain are replicative cells and the cases **V** & **VI** reduce to the previously considered cases **I** & **II**, respectively. Furthermore, the cell proliferation probabilities for the cases **V** & **VI** are prescribed in a similar way to cases **III** & **IV** (28, 40), respectively.

For case **V**, with a constant nutrient concentration in the variable length replicative region, we have

$$p_\tau^i = \begin{cases} p & \text{if } i \le hN_\tau, \\ 0 & \text{if } i > hN_\tau, \end{cases} \tag{51}$$

and in case **VI**, with a linearly decreasing nutrient concentration in the variable length replicative region, we have

$$p_\tau^i = \begin{cases} p_{\max} + \frac{p_{\max}}{hN_\tau - 1} - \frac{p_{\max}}{hN_\tau - 1} i & \text{if } i \le hN_\tau, \\ 0 & \text{if } i > hN_\tau, \end{cases} \tag{52}$$

(see Figures 3e & 3f).

The pathlines and length of the domain for the cell proliferation probabilities (51, 52) can be determined in a similar way to the cases **III** & **IV** (28, 40). Therefore, and for the sake of brevity, we simply provide the results below, and remark that the full derivation can be found in [41].

For case **V**, the pathlines and length of domain are gievn by:

$$X^j(t) = \begin{cases} X_0^j \exp\{\hat{p}t\}, & \text{if } X_0^j \le H_0, t < t_j^\star, \\ L_0 \exp\{\hat{p}t\} - L_0 \exp\{h\hat{p}t_j^\star\} + hL_0 \exp\{h\hat{p}t_j^\star\}, & \text{if } X_0^j \le H_0, t \ge t_j^\star, \\ X_0^j - L_0 + L_0 \exp\{h\hat{p}t\}, & \text{if } X_0^j > H_0, \end{cases} \tag{53}$$

where H_0 is the initial length of the replicative region, and $\hat{p}$ is the constant proliferation rate in the variable length replicative region, with

$$t_j^\star = \frac{\log\left(\frac{X_0^j}{hL_0}\right)}{h\hat{p} - \hat{p}}. \tag{54}$$

The length of the domain is then

$$L(t) = L_0 \exp\{h\hat{p}t\}. \tag{55}$$

For case **VI**, the pathlines and length of domain are given by:

$$X^j(t) = \begin{cases} \frac{hL_0 \exp\left\{\frac{h\hat{p}_{\max}}{2}t\right\}}{\frac{1}{2-h} + \left[\frac{hL_0}{X_0^j} - \frac{1}{2-h}\right] \exp\{(\frac{h}{2}-1)\hat{p}_{\max}t\}}, & \text{if } X_0^j \le H_0, t < t_j^\star, \\ L_0 \exp\left\{\frac{\hat{p}_{\max}}{2}t\right\} + hL_0 \exp\left\{\frac{\hat{p}_{\max}}{2}t_j^\star\right\} - L_0 \exp\left\{\frac{\hat{p}_{\max}}{2}t_j^\star\right\}, & \text{if } X_0^j \le H_0, t \ge t_j^\star, \\ X_0^j - L_0 + L_0 \exp\left\{\frac{h\hat{p}_{\max}}{2}t\right\}, & \text{if } X_0^j > H_0, \end{cases} \tag{56}$$

where $\hat{p}_{\max}$ is the maximum proliferation rate in the variable length replicative region, and $t_j^\star$ is determined by solving

$$X^j(t_j^\star) = hL_0 \exp\left\{h\frac{\hat{p}_{\max}}{2}t_j^\star\right\} \tag{57}$$

numerically. The length of the domain in this case is then

$$L(t) = L_0 \exp\left\{\frac{h\hat{p}_{\max}}{2}t\right\}. \tag{58}$$

The pathlines (53, 56) for cases **V** & **VI** are shown in Figures 4(e) & 4(f), respectively. We observe similar macroscale growth properties as those shown in Figures 4(c) & 4(d), for cases **III** & **IV**–globally we have non-uniform growth. However, and importantly, we see that the length of domain is increasing exponentially (55, 58) in the cases **V** & **VI**. More specifically, in the cases **V** & **VI**, where the replicative region is a constant proportion of the length, the length of the domain is growing exponentially. This is in contrast to the cases **III** & **IV**, where the length of the replicative region is constant, and the domain is growing linearly.

4. Discussion

In this paper, we have considered six input nutrient concentrations (equivalently, cell proliferation probabilities) in our CA model, and determined closed-form expressions for the continuum paths of cells initially labelled within the domain. We also showed that, for sufficiently large colonies, the continuum approximation is accurate for predicting the expected discrete cell positions as time evolves. Of course, there are an endless number of distributions for the cell proliferation probabilities, but the six cases considered in this work capture the qualitative features of uniform and non-uniform growth emerging at the macro-scale. In principle, it is possible to determine solutions for the paths, at least numerically, for any input nutrient concentration using equation (4). Whether the paths are given in closed form or numerically, they can be used to examine the properties of growth at the macro-scale, without recourse to the time consuming averaging of data from a large number of simulations of the CA.

We find that domain growth is exponential if a constant proportion of all cells in the domain are replicative (*e.g.* cases **I**, **II**, **V** & **VI**)). However, if there is only a fixed number of replicative cells, or fixed-length replicative region (and all remaining cells are quiescent) then the domain grows linearly (*e.g.* cases **III** & **IV**). This is an important result as it identifies that it is necessary to have a fixed-length replicative region to produce linear growth. Our results therefore show that the fixed length replicative region is the primary reason why the cylindrical yeast colonies are growing linearly in the experiments of Vulin et al. [19]. With the additional assumption of a monotonically decreasing nutrient concentration within the replicative region, which is approximated by a linearly decreasing nutrient concentration in case **IV**, the growth will be non-uniform. Importantly, in our model the overall length of the colony emerges in response to the input nutrient concentration, and this is in contrast to previous studies in which the length of the domain has been assumed to be an input of the system rather than a response [21, 36, 37, 38, 39, 40].

Based on our analysis, we have also identified the characteristics of growth within a domain that is increasing in length uniaxially. This provides the basis of a classification system for the type of growth in certain regions of a domain, with potential application in future experiments of yeast growth to ascertain if our model prediction of non-uniform growth within the replicative region is correct.

The first classification is *uniform growth*, e.g. Figure 4(a) and replicative region of Figure 4(c). There are two conditions to determine whether uniform growth has occurred within a region. Firstly, we require proliferation to occur within the region and thus we require the distances between adjacent paths to increase. Secondly, we require proliferation to be uniform across the region and thus we expect the distances between adjacent paths to increase uniformly across the region. Suppose we define the distance function $d_\tau^{n,i}$ to be the distance between paths X_τ^i and X_τ^{i+n} at time step τ. We thus have

$$d_\tau^{n,i} = X_\tau^{i+n} - X_\tau^i. \tag{59}$$

We note that the cells X_τ^i and X_τ^{i+n} are initially n cells apart. It will thus always be the case that $d_0^{n,i} = n\Delta x$ and so, for uniform growth, we find that

$$d_\tau^{n,i} = d_\tau^{n,j} \quad \text{for all } i, j, n, \tau, \tag{60}$$

$$d_{\tau+1}^{n,i} > d_\tau^{n,i} \quad \text{for all } i, n, \tau. \tag{61}$$

The second classification is *non-uniform growth*, e.g. Figure 4(b) and replicative region of Figure 4(d). We again require proliferation to occur within the region and thus require the distances between adjacent paths to be increasing. However, we require proliferation to be non-uniform across the region and thus expect the distances between initially adjacent paths to be different as the number of time steps increases. The conditions for non-uniform growth are thus

$$d_\tau^{n,i} \neq d_\tau^{n,j} \quad \text{for all } i, j, n, \tau, \tag{62}$$

$$d_{\tau+1}^{n,i} > d_\tau^{n,i} \quad \text{for all } i, n, \tau. \tag{63}$$

The last classification we have is *quiescence*, e.g. quiescent region of Figures 4(c) and (d), which occurs in regions where no proliferation events occur. We thus expect the distances between adjacent paths to remain the same for all time steps τ and hence the condition for quiescence is

$$d_{\tau+1}^{n,i} = d_\tau^{n,i} \quad \text{for all } i, n, \tau. \tag{64}$$

We can use equations (60-64) to create a classification system of the three types of growth in a region. Suppose we have empirical data for the displacement of initial cells in some colony exhibiting uniaxial growth. This data can be obtained by labelling individual cells with unique fluorescent protein colour combinations [43], and green fluorescent proteins (GFP) have previously been used as a marker for tumor cells to detect and predict single-cell behaviour [44]. Fluorescent ubiquination-based cell cycle indicators (FUCCI) have also been used to track cell progression [45]. Using the empirical data, we can determine the distances between adjacent paths $d_\tau^{n,i}$. We can thus analyse the distances $d_\tau^{n,i}$ to classify regions of growth using equations (60-64). Although this procedure is purely a classification system, we may still make some inference about cell proliferation within the colony. If a region in the colony is classified as quiescent, we know there is no cell proliferation in that region. Furthermore, if a region in the colony is classified as having uniform growth, we know the proliferation rate in that region will be some constant. However, if a region in the colony is classified as having non-uniform growth, we cannot make any inference about the cell proliferation for these regions without applying further analysis. This could involve comparing experimentally-determined pathlines with theoretical results from this paper. However, the question of whether the qualitative differences in the shape of the pathlines under the different hypotheses could be detected in noisy experimental data remains open.

If we know the relationship between the cell proliferation rate and the nutrient concentration, our model can be used to infer the nutrient distribution given experimentally-observed cell pathlines. For simplicity, in our interpretations in this paper, we have assumed the proliferation rate is directly proportional to the nutrient concentration. However, there are a variety of different functional forms that are biologically plausible *e.g.*, cells might proliferate at a constant rate once the nutrient concentration is above a certain threshold, or proliferation rate could be a sigmoidal function of nutrient concentration. Theoretically, the relationship could be determined by measuring cell proliferation rates in experiments where cells are grown in uniform nutrient conditions, at different nutrient concentrations. Alternatively, if both experimentally determined pathlines, and measurements of nutrient concentration are available in uniaxial growth experiments, then it should be possible to infer the relationship between the nutrient level and the proliferation rate.

For simplicity, in this paper, we have neglected cell death, and breakdown of dead cells. Although we believe this assumption is reasonable for yeast cells, for other cell types it is likely to be inaccurate. If cell death and necrosis was non-negligible, this would have a potentially significant effect on the pathlines of the cells, since it would allow for a local contraction of the tissue (there could be retrograde movement of cells to 'fill in' the gap left by the dead cell) whilst our model assumes that the pathlines are always non-decreasing in time. The other major simplifying assumption in the model is that the nutrient concentration (or, equivalently, the cell proliferation rate) is prescribed, but in reality, of course, the two processes are coupled. Cell proliferation requires nutrients, which must diffuse through, and are advected with, the cells in the colony and hence we expect the nutrient distribution to evolve in space and time in a much more complex way than is considered here. In future work, we aim to analyse the pathlines produced by such fully-coupled models of cell growth and nutrient transport. It is also worth noting that many factors influence cell proliferation, in addition to nutrient levels. These might include the build-up of toxic metabolites, or resides from necrotic cells, or a reduction in cell proliferation rate as a result of local overcrowding. We hope to consider the impact of at least some of these effects in the future.

Acknowledgements

AJG acknowledges funding from the Australian Government under the Research Training Program. HT and BJB acknowledge funding from the Australian Research Council (grant number DP160102644).

References

[1] C. Rosa, G. Peter, Biodiversity and Ecophysiology of Yeasts, Springer, 2006.

[2] N. Williams, Yeast genome sequence ferments new research, Science 272 (1996) 481–481. doi:10.1126/science.272.5261.481. arXiv:https://science.sciencemag.org/content/272/5261/481.full.pdf.

[3] G. Botstein, D.and Fink, Yeast: An experimental organism for 21st century biology, Genetics 189 (2011) 695–704. doi:10.1534/genetics.111.130765. arXiv:https://www.genetics.org/content/189/3/695.full.pdf.

[4] L. Chen, J. Noorbakhsh, R. Adams, J. Samaniego-Evans, G. Agollah, D. Nevozhay, J. Kuzdzal-Fick, P. Mehta, G. Balazsi, Two-dimensionality of yeast colony expansion accompanied by pattern formation, PLOS Computational Biology 10 (2014) 1–14. doi:10.1371/journal.pcbi.1003979.

[5] T. Reynolds, G. Fink, Bakers' yeast, a model for fungal biofilm formation, Science 291 (2001) 878–881. doi:10.1126/science.291.5505.878.
[6] E. M. Kojic, R. O. Darouiche, Candida infections of medical devices, Clinical Microbiology Reviews 17 (2004) 255–267. doi:10.1128/cmr.17.2.255-267.2004.
[7] F. Priest, G. Stewart, W. Hardwick, Handbook of brewing, 2nd ed. ed., CRC/Taylor & Francis, Boca Raton, 2006.
[8] A. Teoh, G. Heard, J. Cox, Yeast ecology of kombucha fermentation, International Journal of Food Microbiology 95 (2004) 119–126. doi:10.1016/j.ijfoodmicro.2003.12.020.
[9] E. Moore-Landecker, Fundamentals of the fungi, Prentice Hall, 1996.
[10] N. A. Buijs, V. Siewers, J. Nielsen, Advanced biofuel production by the yeast saccharomyces cerevisiae, Current Opinion in Chemical Biology 17 (2013) 480 – 488. URL: http://www.sciencedirect.com/science/article/pii/S1367593113000598. doi:https://doi.org/10.1016/j.cbpa.2013.03.036, next generation therapeutics * Energy.
[11] M. Yang, S. Zheng, Pollutant removal-oriented yeast biomass production from high-organic-strength industrial wastewater: A review, Biomass and Bioenergy 64 (2014) 356 – 362. URL: http://www.sciencedirect.com/science/article/pii/S096195341400138X. doi:https://doi.org/10.1016/j.biombioe.2014.03.020.
[12] Z. Palkova, Multicellular microorganisms: laboratory versus nature, EMBO Rep 5 (2004) 470–476. doi:10.1038/sj.embor.7400145.
[13] S. Pirt, A kinetic study of the mode of growth of surface colonies of bacteria and fungi, J. Gen. Microbio. 47 (1967) 181–197. doi:10.1099/00221287-47-2-181.
[14] T. Reynolds, A. Jansen, X. Peng, G. Fink, Mat formation in saccharomyces cerevisiae requires nutrient and ph gradients, Eukaryotic Cell 7 (2008) 122–130. doi:10.1128/EC.00310-06.
[15] H. Tronnolone, J. M. Gardner, J. F. Sundstrom, V. Jiranek, S. G. Oliver, B. J. Binder, Quantifying the dominant growth of mechanisms of dimorphic yeast using a lattice–based model, J. R. Soc. Interface 14 (2017) 20170314. doi:http://dx.doi.org/10.1098/rsif.2017.0314.
[16] A. Tam, J. E. F. Green, S. Balasuriya, E. Tek, J. Gardner, J. Sundstrom, V. Jiranek, B. Binder, Nutrient-limited growth with non-linear cell diffusion as a mechanism for floral pattern formation in yeast biofilms, Journal of Theoretical Biology 448 (2018) 122–141. doi:10.1016/j.jtbi.2018.04.004.
[17] A. Tam, J. E. F. Green, S. Balasuriya, E. L. Tek, J. M. Gardner, J. F. Sundstrom, V. Jiranek, B. J. Binder, A thin-film extensional flow model for biofilm expansion by sliding motility, Proceedings of the Royal Society A 475 (2019) 20190175.
[18] A. Tam, Mathematical Modelling of Pattern Formation in Yeast Biofilms, Ph.D. thesis, School of Mathematical Sciences, The University of Adelaide, 2019.
[19] C. Vulin, J. Di Meglio, A. B. Lidner, A. Daerr, A. Murray, P. Hersen, Growing Yeast into Cylindrical Colonies, Biophysical Journal 106 (2014) 2214–2221.
[20] B. Nguyen, A. Upadhyaya, A. van Oudenaarden, M. Brenner, Elastic instability in growing yeast colonies, Biophysical Journal 86 (2004) 2740–2747. doi:10.1016/S0006-3495(04)74327-1.
[21] B. J. Binder, K. A. Landman, M. J. Simpson, M. Mariani, D. F. Newgreen, Modeling proliferative tissue growth: A general approach and an avian case study , Physical Review E 78 (2008).
[22] L. Minarikova, M. Kuthan, M. Ricicova, J. Forstova, Z. Palkova, Differentiated gene expression in cells within yeast colonies, Experimental Cell Research 271 (2001) 296–304. doi:10.1006/excr.2001.5379.
[23] L. Vachova, O. Chernyavskiy, D. Strachotova, P. Bianchini, Z. Burdikova, L. Fercikova, L. Kubinova, Z. Palkova, Architecture of developing multicellular yeast colony: spatio?temporal expression of ato1p ammonium exporter, Environmental Microbiology 11 (2009) 1866–1877. doi:10.1111/j.1462-2920.2009.01911.x.
[24] B. J. Binder, K. A. Landman, Exclusion Processes on a Growing Domain, Journal of Theoretical Biology 259 (2009) 541–551.
[25] B. J. Binder, K. A. Landman, Tissue Growth and the Polya Distribution, Australasian Journal of Engineering Education 15 (2009) 35–42.
[26] J. Hywood, E. Hackett-Jones, K. Landman, Modeling biological tissue growth: Discrete to continuum representations, Phys. Rev. E 88 (2013) 032704. doi:10.1103/PhysRevE.88.032704.
[27] R. Ross, R. Baker, C. Yates, How domain growth is implemented determines the long-term behavior of a cell population through its effect on spatial correlations, Phys. Rev. E 94 (2016) 012408. doi:10.1103/PhysRevE.94.012408.
[28] E. Codling, M. Plank, S. Benhamou, Random walk models in biology, Royal Society Interface 5 (2008) 813–834. doi:10.1098/rsif.2008.0014.
[29] J. Ross, B. Binder, Approximating spatially exclusive invasion processes, Phys. Rev. E 89 (2014) 052709. doi:10.1103/PhysRevE.89.052709.
[30] C. Yates, A. Parker, R. Baker, Incorporating pushing in exclusion-process models of cell migration, Phys. Rev. E 91 (2015) 052711. doi:10.1103/PhysRevE.91.052711.
[31] R. Ross, C. Yates, R. Baker, Inference of cell?cell interactions from population density characteristics and cell trajectories on static and growing domains, Mathematical Biosciences 264 (2015) 108–118. doi:10.1016/j.mbs.2015.04.002.
[32] A. Kansal, S. Torquato, G. Harsh IV, E. Chiocca, T. Deisboeck, Cellular automaton of idealized brain tumor growth dynamics, Biosystems 55 (2000) 119–127. doi:10.1016/S0303-2647(99)00089-1.
[33] A. Monteagudo, J. Santos, Treatment Analysis in a Cancer Stem Cell Context Using a Tumour Growth model Based on Cellular Automata, PLOS ONE 10 (2015) e0132306.
[34] J. S. Czarnecki, S. Jolivet, M. E. Blackmore, K. Lafdi, P. A. Tsonis, Cellular Automata Simulation of Osteoblast Growth on Microfibrous-Carbon-Based Scaffolds, Tissue Engineering Part A 20 (2014) 3176–3188.
[35] S. Matsuura, Random Growth of Fungal Colony Model on Diffusive and Non-Diffusive Media, Forma 15 (2000) 309–319.
[36] A. Lai De Oliveira, B. J. Binder, Modeling uniaxial nonuniform cell proliferation, Bulletin of Mathematical Biology 81 (2019) 2220–2238. doi:10.1007/s11538-019-00601-4.
[37] E. J. Crampin, E. A. Gaffney, P. K. Maini, Reaction and Diffusion on Growing Domains: Scenarios for Robust Pattern Formation, Bulletin of Mathematical Biology 61 (1999) 1093–1120.
[38] M. J. Simpson, Exact Solutions of Linear Reaction–Diffusion Processes on a Uniformly Growing Domain: Criteria for Successful Colonization, PLoS ONE 10 (2015) e0117949. doi:10.1371/journal.pone.0117949.
[39] M. Chaplain, M. Ganesh, I. Graham, Spatio-temporal pattern formation on spherical surfaces: numerical simulation and application to solid tumour growth, Journal of Mathematical Biology 42 (2001) 387–423. doi:10.1007/s002850000067.

[40] P. Mulesa, G. Cruywagen, S. Lubkin, P. Maini, M. Ferguson, J. Murray, On a model mechanism for the spatial patterning of teeth primordia in the alligator, Journal of Theoretical Biology 180 (1996) 287–296.
[41] A. J. Gallo, Modelling One-Dimensional Non-Uniform Growth with Applications to Cylindrical Yeast Colonies, Master's thesis, University of Adelaide, 2020.
[42] P. J. Brown, J. E. F. Green, B. J. Binder, J. M. Osborne, A rigid body framework for multi-cellular modelling, bioRxiv (2021). doi:https://doi.org/10.1101/2021.02.10.430170.
[43] S. Di Talia, K. D. Poss, Monitoring tissue regeneration at single-cell resolution, Cell Stem Cell 19 (2016) 428–431.
[44] M. Zimmer, Green fluorescent protein (gfp): Applications, structure, and related photophysical behaviour, Chem. Rev. 102 (2002) 759–782. doi:10.1021/cr010142r.
[45] S. T. Vittadello, S. W. McCue, G. Gunasingh, N. K. Haass, M. J. Simpson, Mathematical models incorporating a multi-stage cell cycle replicate normally–hidden inherent synchronisation in cell proliferation, J. R. Soc. Interface (2019). doi:20190382.