

Multistage volume exclusion models for cell proliferation

John Carlo Dimaculangan , Cameron A. Smith , and Christian A. Yates 
Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, United Kingdom



(Received 22 April 2026; accepted 14 September 2026; published 13 October 2026)

Cell proliferation and cell movement are fundamentally stochastic processes which lead to variability in the growth and spatial structure of cell populations in many biological settings, such as cell invasion, wound healing, and tumor growth. Understanding how variability in cell-cycle progression and local spatial constraints interact to impact population-level behavior requires modeling frameworks that explicitly represent individual-cell behavior and elucidate how these affect emergent tissue-level behavior. We develop stochastic, on-lattice agent-based models (ABMs) which incorporate volume exclusion, random movement, and multistage representations of the cell cycle. The multistage framework enables a more realistic representation of true cell-cycle time distributions. We also introduce a form of myopic behavior, in which cells sense their local environment when attempting to proliferate. For each ABM, we derive a corresponding continuum partial differential equation (PDE) description under the mean-field approximation. Using numerical simulations, we investigate how different proliferation mechanisms influence population-level dynamics in both the discrete and continuum models. In particular, we consider biologically relevant contexts of growth-to-confluence assays (using uniform initial conditions) and traveling wave behavior associated with cell invasion. We examine how the PDE solutions compare with the behavior of the corresponding ABMs averaged over many realizations.

DOI: [10.1103/sn9r-l6ns](https://doi.org/10.1103/sn9r-l6ns)

I. INTRODUCTION

Cell proliferation is a fundamental biological process which underpins tissue development, immune responses [1], wound healing [2], and cancer growth [3]. Understanding the cell cycle, the framework which regulates when and how cell proliferation occurs, is crucial to developing biologically realistic models of such processes. The canonical eukaryotic cell cycle of actively dividing cells consists of four distinct phases: G_1 phase (Gap 1) where cells grow and prepare to replicate their DNA, S phase (Synthesis) where cells replicate their DNA, G_2 phase (Gap 2) where cells prepare for mitosis and continue to grow, and M phase (Mitosis) where a cell divides [4]. Cells are able to temporarily arrest at key checkpoints prior to progressing into the S phase or mitosis if they are sufficiently damaged or conditions for proliferation are unfavourable [5,6]. In particular, it has been shown that crowded environments can cause a cell to become quiescent [7,8]. Therefore, accurate mathematical models of such cell proliferation should investigate the impact that density-dependent behavior within the cell cycle has on global cell population dynamics.

In this paper we consider two complementary approaches to modeling the cell cycle. The first approach uses on-lattice ABMs which are able to capture individual cell behavior and stochasticity, which is inherent to cell-cycle progression and to cell movement. ABMs can be used to understand the spatial

heterogeneity which can arise within a growing population of cells. Furthermore, it is relatively simple and intuitive to implement specific cellular behavior or cell-cell interactions within these frameworks [9]. However, ABMs have the disadvantage of a high computational cost, and limited analytical tractability [10].

The ability to accurately capture cell-cycle time distributions is essential for constructing biologically realistic models of cell proliferation. Previous work often makes the simplifying assumption that cell-cycle times follow an exponential distribution [11–13]. This assumption is convenient as it allows for simulations to be carried out using Gillespie’s algorithm [14]. However, it can be shown that cell-cycle times are not well represented by the exponential distribution [15]. In contrast, the multistage model for the cell cycle partitions the cell cycle into multiple stages (distinct from the biological phases above) with exponential waiting times, which cells must progress through prior to dividing. This allows the multistage model to yield cell-cycle time distributions which accurately reflect experimentally observed cell-cycle times in mice fibroblasts [16] and retinal pigment epithelial cells [17]. Importantly, because these stages are modelled independently of each other, the multistage model can be implemented using Gillespie’s algorithm. We note that the multistage model is intended to provide an accurate description of eukaryotic cells which are actively dividing. For example, bacteria cells do not progress through the cell phases G_1 , S, G_2 , and M which we described previously [18]. The multistage model does not necessarily provide an accurate representation of bacterial cell division, or eukaryotic cells which do not divide by mitosis.

ABMs incorporating both cell proliferation and cell movement have also been widely studied. Some models assume that cell-cycle times follow an exponential distribution [19–21],

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

whereas other models adopt the multistage representation of the cell cycle [22,23]. Under standard multistage models, cells choose the site of placement of the daughter cell at random from its neighbors, with the event aborted if the site is occupied. In contrast, myopic behavior grants the ability for cells to sense their local environment. Previously, myopic behavior has been studied in the context of cell movement by Landman and Fernando [24]. In this paper we will continue to investigate spatially extended multistage ABMs and compare their behavior to spatial ABMs which use the simpler assumption of exponentially distributed cell-cycle times. We will also adapt the approach taken by Landman and Fernando and study myopic behavior in the context of cell proliferation. We will determine how incorporating myopic behavior into ABMs changes their behavior, and study the effects of myopic behavior together with the multistage model.

The second approach we take to modeling density-dependent cell proliferation and cell movement is to use continuum (nonspatial) ordinary and (spatial) partial differential equations (ODEs and PDEs) which can capture the macroscale dynamics of a system. There is well established theory (such as steady-state and bifurcation analysis) for many continuum models, and it is computationally cheaper to numerically solve a system of ODEs and PDEs in comparison to simulating an equivalent population of cells using ABMs. However, continuum models are unable to capture any stochastic behavior within a system and can neglect the local structure [19].

Nonspatial models of cell proliferation are typically formulated using coupled systems of ODEs [25,26]. These frameworks incorporate realistic cell-cycle time distributions by adopting an underlying multistage structure, in which each ODE represents the density of cells in each stage of the cell cycle. Delay differential equations (DDEs) have also been proposed [27], in which a distributed delay kernel is used to impose specific cell-cycle time distributions. DDEs are also able to capture realistic cell-cycle time distributions, but at the cost of a more complicated model. To account for spatial effects such as migration and crowding, these approaches must be extended. Classical spatial models often take the form of reaction-diffusion equations, such as FKPP-type equations with logistic growth [28]. However, in their simplest form, these models implicitly assume exponential waiting times between cell division attempts. More accurate representations of cell-cycle dynamics can be obtained by extending these formulations to systems of coupled PDEs, where each PDE represents cells in a different cell-cycle stage [29,30]. While such models are capable of capturing crowding effects and structured cell-cycle progression, they are typically limited to a relatively small number of stages, as the analysis and parameter inference for these systems becomes more challenging with increasing model complexity.

Cell-cycle regulation is not controlled solely by density dependent effects. Wang *et al.* [31] demonstrate that the population of quiescent skeletal muscle cells in the same spatial environment is heterogeneous. More general modeling frameworks such as those by Yao [32] for mammalian cells can be used to represent other causes of cell quiescence. However, a model which explicitly incorporates space allows for spatial effects to be quantified.

In this paper, we derive the continuum PDE limits of our discrete ABMs under various assumptions. When these assumptions are valid, we demonstrate that the resulting continuum PDEs provide an accurate description of the average behavior across many realizations of the corresponding ABM. We also investigate parameter regimes in which the mean-field assumption breaks down, and show that, in such cases, the growth dynamics predicted by the ABMs and their associated continuum PDE limits can differ significantly. This enables us to distinguish *a priori* between scenarios in which PDEs successfully capture the spatial structure of the underlying ABM and those in which they do not. Finally, we use the continuum PDE framework to recover known theoretical minimum wave speeds for traveling wave fronts [23], and perform numerical simulations to assess how well these theoretical predictions agree with the behavior observed in our ABMs.

The paper is organized as follows. In Sec. II we present the different implementations of ABMs and perform numerical simulations to investigate the differences in their behavior. In Sec. III we derive the continuum limits of each of our ABMs, and give conditions under which these continuum descriptions are valid. In Sec. IV we use pair correlation functions to identify parameter regimes where our continuum descriptions from Sec. III provide both good and poor agreement to the ensemble average behavior of corresponding discrete ABMs. In Sec. V we study cell invasion and compare numerical experiments of our models to known theoretical results on the speed of invasion. We conclude in Sec. VI with a discussion of our results and potential areas for future work.

II. AGENT-BASED MODELS FOR CELL PROLIFERATION INCORPORATING VOLUME EXCLUSION

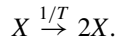
In this section, we will present several on-lattice ABMs incorporating density dependence through volume exclusion for cell proliferation. We will consider two different proliferation models (exponential and multistage) together with two mechanisms for handling failed proliferation events for the multistage model. We then extend these frameworks through the introduction of myopic behavior, which is a form of local sensing that has been previously studied in the context of movement [24]. For each of the cell-cycle models, we give the results of our numerical simulations and discuss the effects each model has on the growth and spatial structure of cell populations.

A. The multistage model and volume exclusion

We begin by introducing biologically motivated agent-based models for cell proliferation. We first introduce and give known results of proliferation models in a nonspatial context [16], then extend these to spatial models incorporating volume exclusion [21,33]. Through numerical simulations, we will demonstrate that the nonspatial theory does not explain the influence of density dependence on cell proliferation dynamics.

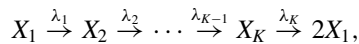
Throughout this section we assume that all cells have an average cell-cycle time of T . One method for simulating the cell cycle is to assume that cell-cycle times follow an exponential distribution with rate $1/T$. We represent this process

as the single reaction,



Under this assumption, we can use Gillespie's stochastic simulation algorithm (SSA) [14] to simulate cell proliferation. We will call this method of simulating the cell cycle the *Exponential model*.

The Exponential model is simple to implement and is analytically tractable. However, it has been shown that exponential distributions provide poor fits to experimentally observed cell-cycle times [16]. Consequently, the underlying assumption required to utilize Gillespie's algorithm is invalid. Instead, we introduce the *multistage model* which retains the ability to be simulated via Gillespie's algorithm, whilst also sampling from more realistic cell-cycle time distributions. For the multistage model, we decompose the cell cycle into multiple stages, each with exponential waiting times, through which a cell must progress before it may divide. Upon division, the two newly produced daughter cells return to the initial stage of the cell cycle. We can represent a K -stage model of the cell cycle with the following series of reactions,



where the transition rates satisfy $\sum_{s=1}^K 1/\lambda_s = T$ (so that the average time to cell division under this model is also T). In contrast to the exponential model, cell-cycle times under the multistage model follow a hypoexponential distribution, the sum of a number of independent exponential random variables. Throughout this paper, we will focus on the case for which all transition rates are equal $\lambda_s = \lambda$. In this special case, we have explicit rate $\lambda = K/T$ for a K -stage model. Under equal progression rates, cell-cycle times are Erlang distributed (the special case of a gamma distribution with integer shape parameter) with rate parameter $1/T$ and shape parameter K . This has been shown to provide a good match to real cell-cycle times [16]. Note that when $K = 1$, the multistage model reduces to the Exponential model.

These two models for cell proliferation have implications for cell population dynamics. We can use the probability master equation (describing the probability of observing the number of cells in each stage of the cell cycle) to derive ODEs which describe the evolution of the mean number of cells over time under a K -stage model. Given equal transition rates, the ODEs are analytically tractable and the mean number of cells can be shown to grow more slowly as the number of stages K increases [16,34]. This suggests that given an observation of cell counts after a known period of time, we would predict a higher average cell-cycle time under the exponential model than the multistage model (for $K > 1$), demonstrating the importance of using the correct cell-cycle model when dealing with experimental data.

We now formulate an on-lattice, volume exclusion model for cell proliferation which incorporates realistic cell-cycle times through the multistage representation of the cell cycle. This model is similar to a previous implementation in [21] which uses exponentially distributed cell-cycle times. Cells move and proliferate on a lattice and we denote these sites by (i, j) . The lattice sites are square with length $\Delta > 0$, and there are L_x and L_y sites along the horizontal and vertical axes

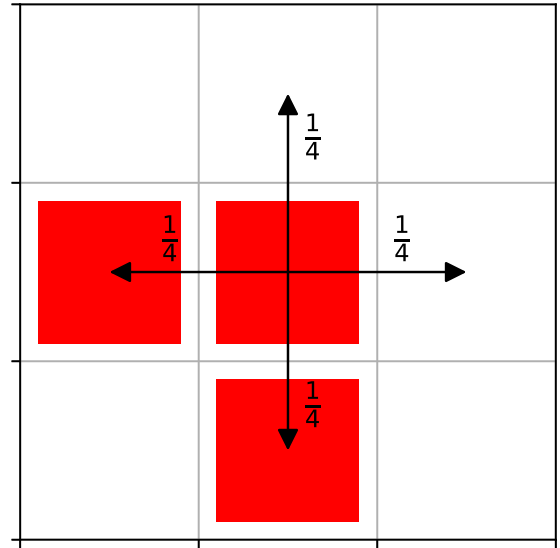


FIG. 1. Schematic of how a cell chooses a neighboring site during proliferation and movement events in our ABMs. Fractions denote the probability of moving in that direction.

respectively. Each lattice site can be occupied by at most one cell. Cells attempt to move according to a motility rate, r_m . When a cell is chosen to move, it selects a neighboring site (in the von Neumann sense) uniformly at random to attempt to move into (see Fig. 1). If the site is empty, the cell moves into the chosen site, otherwise the movement event is aborted and the simulation continues. Given an average cell-cycle time T , cells have proliferation rate, $r_p = 1/T$. Using this, we can write the rate of progression of a cell through each stage of a K -stage model as $\lambda = Kr_p$. When a stage- K cell is chosen to undergo a progression event, it chooses a neighboring site uniformly at random to attempt to proliferate into. If the site is empty, the mother cell places a stage 1 daughter cell in the chosen site and itself returns to the beginning of the cell cycle (stage 1). Otherwise, the proliferation attempt fails and we handle this in two distinct ways. In the *Reset model* the cell returns to stage 1, the beginning of the cell cycle. In the *Remain model* the cell remains in the final stage of the cell cycle. This means that cells under the Remain model are able to arrest in the final cell-cycle stage in high-density environments. We will see that this leads to a noticeable difference in cell-cycle dynamics between the Reset and Remain models.

When $K = 1$, the Reset and Remain models are equivalent as the first and last stage are the same. In simulations, we employ Gillespie's algorithm to generate the time to the next progression or movement event. We select an event to execute with probability proportional to the event rates r_m and λ , and then choose a cell on which to execute this event uniformly from the cell population. It is important to note that if a movement or proliferation attempt fails we continue the simulation without choosing a replacement event.

We demonstrate in Figs. 2(a) and 2(b) that under the Reset model all cells continue progressing through the cell cycle without arresting in high-density environments. This is unrealistic behavior as the implication of the Reset model is that cells which fail to proliferate would pass through all the cell-

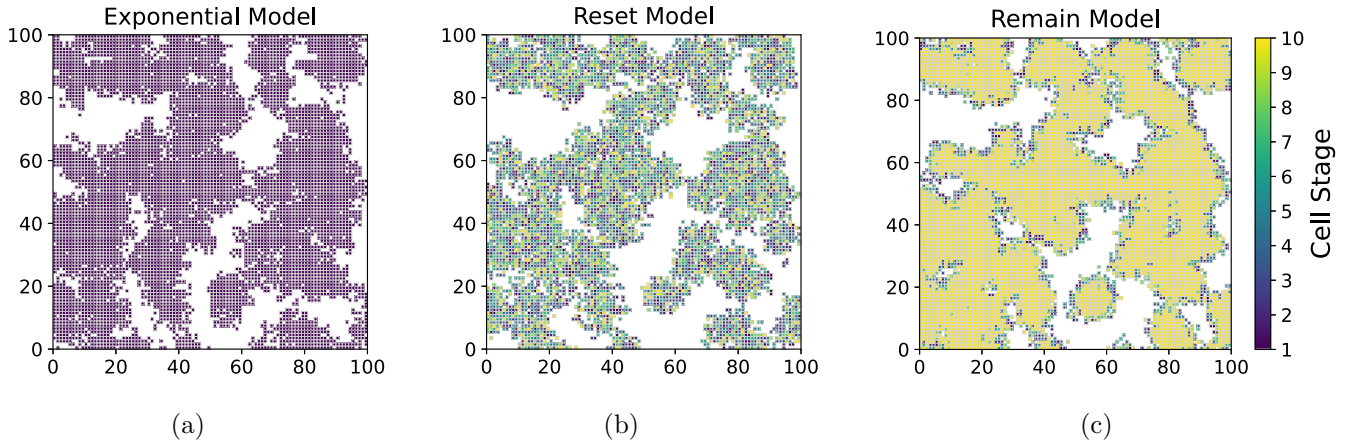


FIG. 2. Single realizations of various ABMs simulated over $t = 10$ time units with motility and proliferation rates $r_m, r_p = 1$ respectively. We simulate on a 100×100 lattice with lattice spacing $\Delta = 1$ and periodic boundary conditions. For each simulation we initialize with an identically seeded uniformly populated lattice with 1% density. Each cell begins in stage 1, the beginning of the cell cycle. We display the final state of each simulation at time $t = 10$ where occupied sites are colored according to the current stage of the cell. (a) Exponential model. (b) Reset model with $K = 10$ stages. (c) Remain model with $K = 10$ stages.

cycle phases again and, for example, in real biological cells, duplicate their DNA again. Furthermore, experiments have shown that cells arrest in high-density environments [35,36]. However, under the Reset model, cells fail to arrest in crowded environments.

In contrast, cells proliferating under the Remain model are able to arrest in high-density environments. We observe this behavior in the numerical simulations we carried out for cells proliferating under the multistage remain model. In Fig. 2(c) most cells which are in the interior of a cluster have all arrested their cell cycle in stage 10, the final stage. Although the Remain model provides a more realistic method of handling density dependence than the Reset and Exponential model, it is not a perfect representation of real cellular behavior either. If we interpret progress through the stages within the Remain model as representing the progress through the cell cycle, then the model implies that cells arrest at the end of the cell cycle, just prior to division. This behavior is biologically unrealistic, as experimental evidence suggests that cells typically arrest prior to the S phase in high-density environments. For example, Streichan *et al.* [35] and Uroz *et al.* [37] show through experiments that intracellular forces regulate progression from G_1 phase to S phase. It is more appropriate to view the final stage as a phenomenological mechanism which allows the Remain model to capture cell quiescence in high-density environments.

We note that previous studies have incorporated intermediate checkpoints within the cell cycle. Falcó *et al.* [30] explicitly model the cell invasion experiments performed by Streichan *et al.* [35], and implement a two-stage PDE model with stages corresponding to G_1 phase and S/ G_2 /M phases. In this model, progression from G_1 phase is density dependent, allowing for cells to arrest at an intermediate point of the cell cycle. Similarly, Höllring *et al.* [38] model a growing colony of epithelial cells and explicitly incorporate a transition probability from G_2 phase which depends on pressure exerted on a cell from neighboring cells relative to its size. If the pressure on a cell is sufficiently high, it cannot progress further through

the cell cycle. Both models allow for cells to arrest at the critical transition between G_1 and S phase. It is possible to extend the model we present in this study to include a checkpoint at an intermediate stage, where cells can halt progression through the cell cycle if local density is sufficiently high. The arguments we make regarding the population-level behavior would still be valid under this extended model.

The average density evolutions in Fig. 3(a) for the spatial multistage and Exponential models exhibit markedly different rates of growth. We observe that the mean number of cells grows more quickly at intermediate densities in the multistage Remain model than in the Exponential model. This is because cells in the Remain model that fail to proliferate do not have to progress through the entire cell cycle again prior to another proliferation attempt [16]. As the Remain model grows faster than the Exponential model in the spatial setting, fitting experimental data with an Exponential model would bias parameter estimates. If the true dynamics are multistage, the fitted cell-cycle time assuming the Exponential model would be shorter than the true cell-cycle time.

We confirm that more proliferation attempts occur in a given period of time under the Remain model in Fig. 3(b). When a cell fails to proliferate but still has empty neighboring sites, the Remain model confers a proliferative advantage compared to the Exponential model and, by extension, the Reset model. This advantage grows as we increase the number of stages K as a larger number of stages leads to a shorter amount of time between proliferation attempts in the Remain model (on average T/K time units compared to T for Exponential/Reset).

We have shown that the multistage Remain model is able to capture realistic cellular behavior where cells arrest in high-density environments. In contrast, the multistage Reset model continues to progress through the cell cycle. We have also observed that the multistage Remain model leads to quicker growth of cell populations than the Exponential model. This proliferative advantage is driven by cells at intermediate densities attempting to divide more frequently. However, an

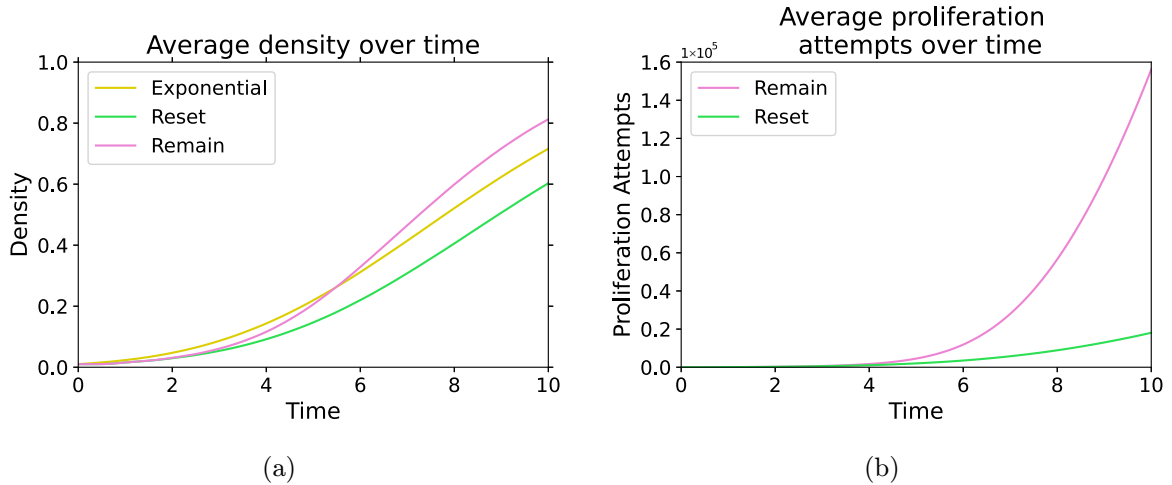


FIG. 3. (a) Density evolutions over $t = 10$ time units for the exponential and multistage models averaged over $M = 100$ realizations, with identical parameters to the simulations above. We initialize each realization with different uniformly populated lattices with 1% density as in Fig. 2. (b) Cumulative average number of proliferation attempts over time for the simulations in Fig. 3(a).

artifact of more proliferation attempts in the Remain model is more failed proliferation attempts. This is because, in our current model, cells are unaware of their local environment. This is biologically unrealistic as it has been observed that groups of neural crest cells are able to coordinate their migration [39]. We will now study an extension of these ABMs where we allow cells to sense their local environment. This behavior makes cells more likely to proliferate successfully when they reach the proliferation stage.

B. Myopic behavior

We now introduce myopic behavior into our ABMs which allows us to model the ability for cells to sense their local environment. Previously, myopic behavior has been studied in the context of movement [24]. In this paper we implement myopic behavior for proliferative events and study the consequences on cell populations. Our simulations show that myopic behavior leads to quicker growth of cell populations under all our cell-cycle models. We also observe, through numerical simulations, that cell populations proliferating under the Myopic Reset and Myopic Remain models grow at the same rate. This contrasts the behavior of the Reset and Remain models without myopic behavior, where we observed that cells grow more quickly under the Remain model than under the Reset model.

In the previous simulations (exemplified in Fig. 2), cells are blind to the occupancy of neighboring sites and choose between them randomly when attempting to proliferate. Now, when a final stage cell is chosen to proliferate, if there are unoccupied neighboring sites, the cell will choose between these unoccupied sites uniformly at random (see Fig. 4). Otherwise, the proliferation attempt fails if there are no empty neighbors. The way in which we handle cells which have failed proliferation attempts in the three different cell-cycle models stays the same. Consequently, we do not expect to see different cell arrest behavior at very high local densities i.e., where many cells are completely surrounded by other cells. Indeed, Figure 5 shows that cells that are completely

surrounded arrest at the final stage in the Remain model but continue progressing through the cell cycle without arresting in high-density environments in the Reset and Exponential models as was also seen under nonmyopic proliferation (see Fig. 2). This suggests that myopic behavior neither improves nor worsens the realism of a cell-cycle model at very high densities.

In Fig. 6 we observe that every ABM with myopic behavior leads to a higher population count than the corresponding ABMs without myopic behavior. Indeed, at sufficiently large times, even the myopic proliferation strategy with the lowest density (Reset) outperforms the “blind” proliferation strategy with the highest density (Remain). This implies that

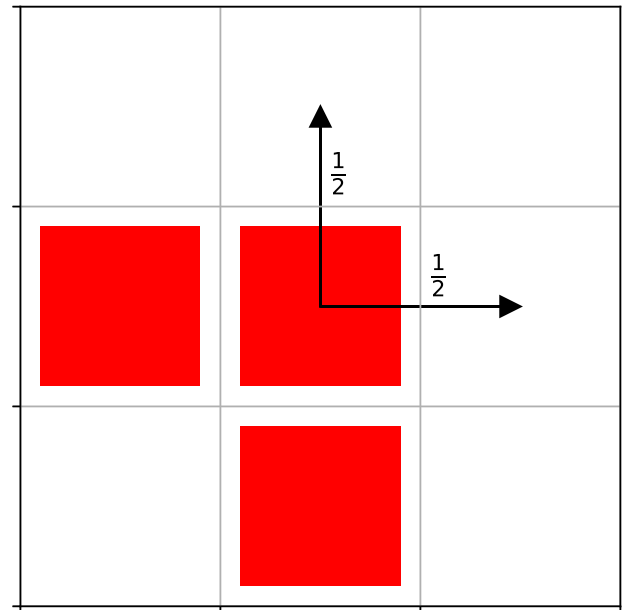


FIG. 4. Example showing how a myopic cell chooses a neighboring site during a proliferation event in our ABMs. Fractions denote the probability of choosing the site to proliferate into.

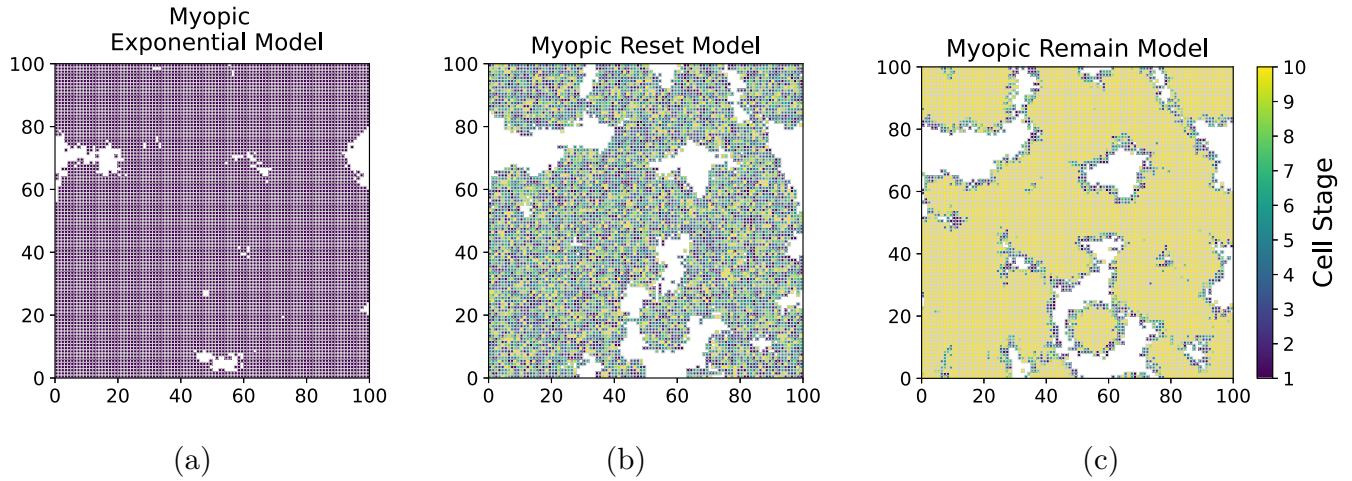


FIG. 5. Single realizations of various ABMs including myopic behavior simulated over $t = 10$ time units with identical parameters and identically seeded to Fig. 2. (a) Exponential model with myopic behavior. (b) Reset model with myopic behavior and $K = 10$ stages. (c) Remain model with myopic behavior and $K = 10$ stages.

myopic behavior provides a larger proliferative advantage than remaining in the last stage of the cell cycle after a failed proliferation attempt. We can interpret this as myopic behavior reducing the impact that high local densities have on cell proliferation.

Another observation we make is that, with myopic behavior, cells proliferating under the Exponential model now grow more quickly than cells proliferating under the multi-stage model. In the absence of failed proliferation events, cells which have just proliferated under the Exponential model are more likely to reattempt proliferation in short periods of time than under the multistage model. This is due to the larger variance in the exponential distribution compared to the Erlang distribution. Furthermore, myopic behavior effectively erodes

the difference in cell growth between the Reset and Remain models. The behavior of the Remain model explains these observations. The Remain model's main advantage over the Reset model is that cells can attempt to proliferate again more quickly than those in the Reset model following a failed attempt. However, under the myopic model, cells which attempt but fail to proliferate have all four neighbors occupied. For intermediate cell densities (in which failed proliferation events become more prominent) and low motility relative to the proliferation rate, successful proliferation events are restricted by the motility of neighboring cells and not by the frequency of proliferation attempts. A proliferating cell requires a neighbor to move away between proliferation attempts.

We have shown that myopic behavior confers a noticeable advantage to proliferating cell populations. This is because cells are now able to sense their local environment, which is more biologically realistic than ABMs where cells are unaware of their surroundings. Myopic behavior also maintains the ability for Remain model cells to arrest at high-density environments. This suggests that the Myopic Remain model successfully captures three key characteristics of proliferating cells: realistic cell-cycle durations, the ability to undergo cell-cycle arrest, and sensitivity to the local environment.

In this section we have introduced spatial ABMs incorporating volume exclusion which simulate cell diffusion and progression through the cell cycle. Through decomposing the cell cycle into multiple stages, each with exponential waiting times, we are able to capture realistic cell-cycle behavior. We then introduced the Reset model, where cells return to the first stage of the cell cycle after a failed proliferation attempt, and the Remain model, where cells stay in the final stage of the cell cycle after a failed proliferation attempt. This captured the ability of cells to arrest in high-density environments. Finally, we implemented a form of myopic behavior at the proliferative stage, capturing the ability for cells to sense their local environment. In Table I below, we summarize the important properties of these models. In the next section, we will derive continuum PDEs from the ABMs we presented in this section.

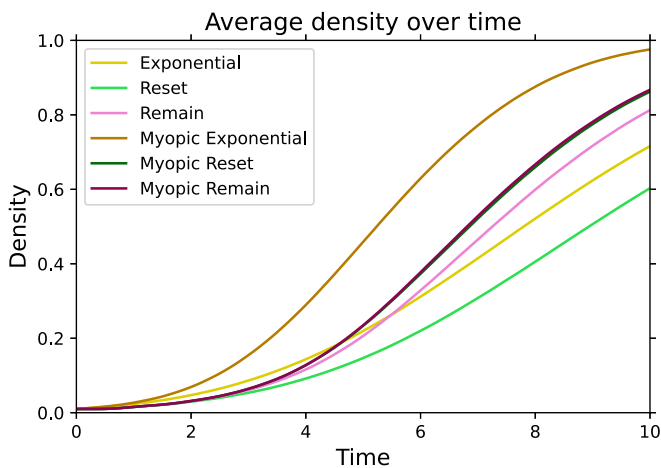


FIG. 6. Average density evolutions over $t = 10$ time units and 100 realizations for every proliferation model with equal motility and proliferation rates $r_m = r_p = 1$ as in Fig. 3(a). In each simulation, we initialize with a uniformly populated lattice with 1% density. We compare the density evolutions of models including myopic behavior to the evolutions without myopic behavior. We distinguish between different proliferation models using different colors.

TABLE I. Summary of the characteristics of the different ABMs explored in this paper.

Model	Cell-cycle time distribution	Choice of proliferation site	Handling of failed proliferation attempt
Exponential	Exponential	Uniformly at random	Cells return to the beginning of the cell cycle after failed proliferation attempts
Reset	Erlang	Uniformly at random	Cells return to the beginning of the cell cycle after failed proliferation attempts
Remain	Erlang	Uniformly at random	Cells remain in the final stage of the cell cycle after failed proliferation attempts
Myopic Exponential	Exponential	Uniformly between unoccupied neighbors	Cells return to the beginning of the cell cycle after failed proliferation attempts
Myopic Reset	Erlang	Uniformly between unoccupied neighbors	Cells return to the beginning of the cell cycle after failed proliferation attempts
Myopic Remain	Erlang	Uniformly between unoccupied neighbors	Cells remain in the final stage of the cell cycle after failed proliferation attempts

III. CONTINUUM PDE MODELS

In this section, we will derive corresponding continuum PDE limits for each ABM under the mean-field assumption. We seek a continuous time and space description for the average density of cells proliferating and diffusing under our various ABMs. Throughout this section, we assume that the average cell-cycle time is fixed. Let the rate of proliferation be r_p . Recall that lattice spacing is given by $\Delta > 0$, r_m is the motility rate, and that we have K stages and rate of cell stage progression $\lambda = Kr_p$. Define $C_{i,j}^{(s)}(t)$ to be the density of stage s cells in lattice site (i, j) at time t when averaged over many repeats over initially identically distributed simulations. We define $C_{i,j}(t) = \sum_{s=1}^K C_{i,j}^{(s)}(t)$ as the average occupancy of the site (i, j) at time t . As each site contains at most one cell, it is immediately apparent that for any stage, s , the densities are bounded as follows $0 \leq C_{i,j}^{(s)} \leq C_{i,j} \leq 1$.

The mean-field assumption is crucial to the methods we use to derive the continuum PDE limits of ABMs. We require lattice sites to be independent for our derivations to be valid. The independence of lattice sites is not guaranteed. In fact, we will show that we require the motility rate, r_m , to be high relative to the proliferation rate, r_p , for the continuum PDEs to be accurate descriptions of the related ABM. Later, in Sec. IV we will provide examples of parameter regimes where our continuum descriptions provide both good and bad approximations of the average behavior of the corresponding ABM.

We now give the master equations for a K -stage Reset model following similar derivations in Simpson *et al.* [20,21] where the authors consider cell proliferation and movement models which incorporate volume exclusion. Let τ be an infinitesimally small time step. By considering the probabilities that a cell can enter or exit a given site within the small time interval $[t, t + \tau)$, we derive the following expression for the average density of stage 1 cells at lattice site (i, j) at time $t + \tau$. Note we omit the dependence of right-hand-side terms on t for convenience:

$$C_{i,j}^{(1)}(t + \tau) = C_{i,j}^{(1)} + \frac{r_m \tau}{4} (1 - C_{i,j}) (C_{i-1,j}^{(1)} + C_{i+1,j}^{(1)} + C_{i,j-1}^{(1)} + C_{i,j+1}^{(1)})$$

$$\begin{aligned} & - \frac{r_m \tau}{4} C_{i,j}^{(1)} (4 - C_{i-1,j} - C_{i+1,j} - C_{i,j-1} - C_{i,j+1}) \\ & + \lambda C_{i,j}^{(K)} \tau - \lambda C_{i,j}^{(1)} \tau + \frac{\lambda \tau}{4} (C_{i-1,j}^{(K)} + C_{i+1,j}^{(K)} \\ & + C_{i,j-1}^{(K)} + C_{i,j+1}^{(K)}) (1 - C_{i,j}) + O(\tau^2). \end{aligned} \quad (1)$$

The second line on the right-hand side of Eq. (1) represents cells successfully moving into site (i, j) while the third line represents cells successfully moving out of site (i, j) . The terms in the fourth line of this equation represent cells moving into stage 1 from stage K , as a result of a successful proliferation attempt or as a result of a reset due to a failed proliferation attempt, and cells progressing out of stage 1 to stage 2. Note that there is no density-dependent term capturing progression for stage K cells as even after a failed proliferation event, a cell will return to stage 1 under the Reset model. The final line of the equation represents a successful proliferation into site (i, j) from a neighboring site.

Similarly for stage $s = 2, \dots, K$ cells, we have the following equation:

$$\begin{aligned} C_{i,j}^{(s)}(t + \tau) = & C_{i,j}^{(s)}(t) + \frac{r_m \tau}{4} (1 - C_{i,j}(t)) (C_{i-1,j}^{(s)}(t) \\ & + C_{i+1,j}^{(s)}(t) + C_{i,j-1}^{(s)}(t) + C_{i,j+1}^{(s)}(t)) \\ & - \frac{r_m \tau}{4} C_{i,j}^{(s)}(t) [4 - C_{i-1,j}(t) \\ & - C_{i+1,j}(t) - C_{i,j-1}(t) - C_{i,j+1}(t)] \\ & + \lambda C_{i,j}^{(s-1)} \tau - \lambda C_{i,j}^{(s)} \tau + O(\tau^2), \end{aligned} \quad (2)$$

which is similar to Eq. (1) but without the final line which resulted from proliferation.

In order to determine the continuum limit, we first apply a Taylor expansion about site (i, j) to the expressions for the occupancies of the neighboring sites, $C_{i-1,j}$, $C_{i+1,j}$, $C_{i,j-1}$, $C_{i,j+1}$. Next, we subtract $C_{i,j}^{(s)}(t)$ and divide through by τ in Eqs. (1) and (2), and finally take the limits $\tau, \Delta \rightarrow 0$ while maintaining $r_m \Delta^2$ and λ constant to

obtain the following system of PDEs,

$$\begin{aligned}\frac{\partial C^{(1)}}{\partial t} &= D[(1 - C)\nabla^2 C^{(1)} + C^{(1)}\nabla^2 C] + \lambda C^{(K)}(2 - C) - \lambda C^{(1)}, \\ \frac{\partial C^{(s)}}{\partial t} &= D[(1 - C)\nabla^2 C^{(s)} + C^{(s)}\nabla^2 C] + \lambda C^{(s-1)} - \lambda C^{(s)}, \quad \text{for } s = 2, \dots, K,\end{aligned}\quad (3)$$

where the diffusion constant D is given by the following limit:

$$D = \lim_{\tau, \Delta \rightarrow 0} \frac{r_m \Delta^2}{4}. \quad (4)$$

In the PDEs (3) above, the diffusion term is nonlinear as the movement of any subpopulation of cells is dependent on the density of total population of cells. If we sum over all the stages $s = 1, \dots, K$, we obtain the following PDE for the density of cells under the Reset model:

$$\frac{\partial C}{\partial t} = D\nabla^2 C + \lambda C^{(K)}(1 - C). \quad (5)$$

As opposed to the PDEs (3), the diffusion term for the total cell density is linear in the above PDE (5). Note that in the Exponential model (where $K = 1$), we recover the Fisher-Kolmogorov-Petrovsky-Piskunov (FKPP) equation [28], a well-studied PDE which is commonly used to model biological phenomenon such as wound healing [40] and species invasion [41].

We now give the master equations for the K -stage Remain model. For stage 1 cells we have the following equation:

$$\begin{aligned}C_{i,j}^{(1)}(t + \tau) &= C_{i,j}^{(1)} + \frac{r_m \tau}{4}(1 - C_{i,j})(C_{i-1,j}^{(1)} + C_{i+1,j}^{(1)} + C_{i,j-1}^{(1)} + C_{i,j+1}^{(1)}) \\ &\quad - \frac{r_m \tau}{4}C_{i,j}^{(1)}(4 - C_{i-1,j} - C_{i+1,j} - C_{i,j-1} - C_{i,j+1}) \\ &\quad + \frac{\lambda \tau}{4}C_{i,j}^{(K)}(4 - C_{i-1,j} - C_{i+1,j} - C_{i,j-1} - C_{i,j+1}) - \lambda C_{i,j}^{(1)}\tau \\ &\quad + \frac{\lambda \tau}{4}(1 - C_{i,j})(C_{i-1,j}^{(K)} + C_{i+1,j}^{(K)} + C_{i,j-1}^{(K)} + C_{i,j+1}^{(K)}) + O(\tau^2).\end{aligned}\quad (6)$$

Equation (6) differs from Eq. (1) only in the fourth line, where the term which represents stage K cells progressing into stage 1 is dependent on the density of neighboring sites. This is because if a final stage cell fails to proliferate, then it remains in the final stage of the cycle under the Remain model. Similarly, the master equation for stage K cells is altered due to the change in the way we handle failed proliferation events in the Remain model. This gives the following equation:

$$\begin{aligned}C_{i,j}^{(K)}(t + \tau) &= C_{i,j}^{(K)} + \frac{r_m \tau}{4}(1 - C_{i,j})(C_{i-1,j}^{(K)} + C_{i+1,j}^{(K)} + C_{i,j-1}^{(K)} + C_{i,j+1}^{(K)}) \\ &\quad - \frac{r_m \tau}{4}C_{i,j}^{(K)}(4 - C_{i-1,j} - C_{i+1,j} - C_{i,j-1} - C_{i,j+1}) \\ &\quad + \lambda C_{i,j}^{(K-1)}\tau - \frac{\lambda \tau}{4}C_{i,j}^{(K)}(4 - C_{i-1,j} - C_{i+1,j} - C_{i,j-1} - C_{i,j+1}) + O(\tau^2),\end{aligned}\quad (7)$$

where the term in the fourth line representing cell progression is density-dependent as opposed to Eq. (2). For stage $s = 2, \dots, K - 1$ cells, the equation for $C_{i,j}^{(s)}(t + \tau)$ is precisely the same as in the Reset model, which is given by Eq. (2). From these equations [(2), (6), and (7)] we can derive the continuum PDEs for the Remain model, using the same methodology and under the same assumptions as the Reset PDEs (3):

$$\begin{aligned}\frac{\partial C^{(1)}}{\partial t} &= D[(1 - C)\nabla^2 C^{(1)} + C^{(1)}\nabla^2 C] + 2\lambda C^{(K)}(1 - C) - \lambda C^{(1)}, \\ \frac{\partial C^{(s)}}{\partial t} &= D[(1 - C)\nabla^2 C^{(s)} + C^{(s)}\nabla^2 C] + \lambda C^{(s-1)} - \lambda C^{(s)}, \quad \text{for } s = 2, \dots, K - 1, \\ \frac{\partial C^{(K)}}{\partial t} &= D[(1 - C)\nabla^2 C^{(K)} + C^{(K)}\nabla^2 C] + \lambda C^{(K-1)} - \lambda C^{(K)}(1 - C).\end{aligned}\quad (8)$$

We observe that the PDE describing the dynamics of stage K cells has a density-dependent progression term. This means that for high densities of cells, we expect that cells should arrest in stage K , which matches the behavior observed in our ABM simulations of the Remain model.

If we sum over all stages, we obtain a PDE for the density of cells under the Remain model

$$\frac{\partial C}{\partial t} = D\nabla^2 C + \lambda C^{(K)}(1 - C). \quad (9)$$

Note that this PDE has the same functional form as the PDE (5) describing the evolution of the total cell population under the Reset model. However, we do not expect the cell densities under the Reset and Remain model PDEs to evolve identically as PDEs (5) and (9) are dependent on the population of stage- K cells. As the PDEs (3) and (8) which describe the dynamics of stage K cells are different, this will result in different solutions for PDEs (5) and (9). Under the Remain model, there is a density dependence in the proliferation term. Explicitly, proliferation depends on the density of stage- K cells. As already mentioned, we expect more cells to accumulate in stage- K in the Remain model than in the Reset model. Consequently, we expect to see cell counts growing more quickly under the Remain model PDEs, which matches our ABM simulations in Figs. 6 and 3(a).

Now, we derive continuum PDE limits for ABMs which incorporate myopic behavior. Let $P((i, j) | (i', j'))$ denote the probability that a myopic cell at site (i', j') chooses to attempt to proliferate into site (i, j) . Using this notation, we have the following equation for stage 1 cells under the Myopic Reset model

$$\begin{aligned} C_{i,j}^{(1)}(t + \tau) &= C_{i,j}^{(1)} + \frac{r_m \tau}{4} (1 - C_{i,j}) (C_{i-1,j}^{(1)} + C_{i+1,j}^{(1)} + C_{i,j-1}^{(1)} + C_{i,j+1}^{(1)}) \\ &\quad - \frac{r_m \tau}{4} C_{i,j}^{(1)} (4 - C_{i-1,j} - C_{i+1,j} - C_{i,j-1} - C_{i,j+1}) \\ &\quad + \lambda C_{i,j}^{(K)} \tau - \lambda C_{i,j}^{(1)} \tau + \lambda \tau (C_{i-1,j}^{(K)} P((i, j) | (i-1, j)) \\ &\quad + C_{i+1,j}^{(K)} P((i, j) | (i+1, j))) \\ &\quad + \lambda \tau (C_{i,j-1}^{(K)} P((i, j) | (i, j-1)) \\ &\quad + C_{i,j+1}^{(K)} P((i, j) | (i, j+1))) + O(\tau^2). \end{aligned} \quad (10)$$

The terms representing cell movement and stage progression in the above Eq. (10) are the same as in Eq. (1) for stage 1 cells under the Reset model without myopic behavior. However, the terms on the fourth and fifth lines of the right-hand side, corresponding to cells proliferating into site (i, j) , are generalized. In our derivations for continuum PDEs corresponding to ABMs without myopic behavior, $P((i, j) | (i', j')) = \frac{1}{4}(1 - C_{i,j})$ for any neighboring site (i', j') of site (i, j) . Under Myopic models, this probability is now dependent

on the occupancy of the neighbors of sites (i', j') i.e., the neighbors of the neighbors of site (i, j) . Below we give an equation for $P((i, j) | (i \pm 1, j))$:

$$P((i, j) | (i \pm 1, j)) = (1 - C_{i,j}) \times \sum_{l=0}^3 \mathcal{N}_l(i \pm 1, j), \quad (11)$$

where $\mathcal{N}_l(i \pm 1, j)$ represents the terms for the probability of proliferating into site (i, j) from $(i \pm 1, j)$ given that site $(i \pm 1, j)$ has l occupied neighboring sites. For $l = 0, 1, 2$, or 3 occupied neighbors of site $(i \pm 1, j)$ $\mathcal{N}_l = \mathcal{N}_l(i \pm 1, j)$ takes the following form:

$$\begin{aligned} \mathcal{N}_0 &= \frac{1}{4}((1 - C_{i\pm 2,j})(1 - C_{i\pm 1,j-1})(1 - C_{i\pm 1,j+1})) \\ \mathcal{N}_1 &= \frac{1}{3}C_{i\pm 2,j}(1 - C_{i\pm 1,j+1})(1 - C_{i\pm 1,j-1}) \\ &\quad + \frac{1}{3}C_{i\pm 1,j-1}(1 - C_{i\pm 2,j})(1 - C_{i\pm 1,j+1}) \\ &\quad + \frac{1}{3}C_{i\pm 1,j+1}(1 - C_{i\pm 2,j})(1 - C_{i\pm 1,j-1}) \\ \mathcal{N}_2 &= \frac{1}{2}C_{i\pm 2,j}C_{i\pm 1,j-1}(1 - C_{i\pm 1,j+1}) \\ &\quad + \frac{1}{2}C_{i\pm 2,j}C_{i\pm 1,j+1}(1 - C_{i\pm 1,j-1}) \\ &\quad + \frac{1}{2}C_{i\pm 1,j-1}C_{i\pm 1,j+1}(1 - C_{i\pm 2,j}) \\ \mathcal{N}_3 &= C_{i\pm 2,j}C_{i\pm 1,j-1}C_{i\pm 1,j+1}. \end{aligned}$$

We can define $P((i, j) | (i, j \pm 1))$ analogously by using Eq. (11). Substituting $P((i, j) | (i-1, j))$, $P((i, j) | (i+1, j))$, $P((i, j) | (i, j-1))$, and $P((i, j) | (i, j+1))$ into Eq. (10) using expressions similar to Eq. (11), then Taylor expanding and taking $\Delta, \tau \rightarrow 0$ yields the following PDE for stage 1 cells under the Myopic Reset model (see Appendix A for more details),

$$\begin{aligned} \frac{\partial C^{(1)}}{\partial t} &= D[(1 - C)\nabla^2 C^{(1)} + C^{(1)}\nabla^2 C] \\ &\quad + \lambda C^{(K)}(2 - C^4) - \lambda C^{(1)}, \end{aligned} \quad (12)$$

where D is given by the same limit as in Eq. (4). The PDEs describing the evolution of cells in stage $s = 2, \dots, K$ are the same as those in the Reset model without myopic behavior (3). This is because myopic behavior affects only the probability of a successful proliferation event. Below we give the full system of PDEs describing the Myopic Reset model:

$$\begin{aligned} \frac{\partial C^{(1)}}{\partial t} &= D[(1 - C)\nabla^2 C^{(1)} + C^{(1)}\nabla^2 C] + \lambda C^{(K)}(2 - C^4) - \lambda C^{(1)}, \\ \frac{\partial C^{(s)}}{\partial t} &= D[(1 - C)\nabla^2 C^{(s)} + C^{(s)}\nabla^2 C] + \lambda C^{(s-1)} - \lambda C^{(s)}, \quad \text{for } s = 2, \dots, K. \end{aligned} \quad (13)$$

Summing over all cell stages gives a PDE for the total cell density,

$$\frac{\partial C}{\partial t} = D\nabla^2 C + \lambda C^{(K)}(1 - C^4). \quad (14)$$

The proliferation term is now $\lambda C^{(K)}(1 - C^4)$, which differs from the PDE (5) describing the total density of cells under the Reset model without myopic behavior. As $0 \leq C \leq 1$, we have $\lambda C^{(K)}(1 - C^4) \geq \lambda C^{(K)}(1 - C)$, and we therefore expect a population of cells exhibiting myopic behavior to proliferate more quickly than populations of cells without myopic behavior. Indeed, we observed this behavior in our ABMs for nonmyopic and myopic behavior under the Reset model in Fig. 6.

We finish this section by deriving the continuum PDEs for the Myopic Remain model. For stage 1 cells we have the following equation:

$$\begin{aligned}
C_{i,j}^{(1)}(t + \tau) = & C_{i,j}^{(1)} + \frac{r_m \tau}{4} (1 - C_{i,j}) (C_{i,j+1}^{(1)} + C_{i,j-1}^{(1)} + C_{i+1,j}^{(1)} + C_{i-1,j}^{(1)}) \\
& - \frac{r_m \tau}{4} C_{i,j}^{(1)} \{4 - C_{i,j+1} - C_{i,j-1}(t) - C_{i+1,j} - C_{i-1,j}\} \\
& + \lambda C_{i,j}^{(K)} (1 - C_{i-1,j} C_{i+1,j} C_{i,j-1} C_{i,j+1}) \tau - \lambda C_{i,j}^{(1)} \tau \\
& + \lambda \tau (C_{i-1,j}^{(K)} P((i, j) | (i-1, j)) + C_{i+1,j}^{(K)} P((i, j) | (i+1, j))) \\
& + \lambda \tau (C_{i,j-1}^{(K)} P((i, j) | (i, j-1)) + C_{i,j+1}^{(K)} P((i, j) | (i, j+1))) + O(\tau^2).
\end{aligned} \tag{15}$$

This is the same as the master equation (10) for stage 1 cells under the Myopic Reset model except for the fourth line on the right-hand side. There is an extra density dependence (corresponding to the probability that there is an empty neighbor of site (i, j)) on the term representing stage- K cells returning to the beginning of the cell cycle. This is because cells proliferating under the Remain model only return to the beginning of the cell cycle after a successful proliferation.

Similarly to the derivation of the system of PDEs for the Remain model without myopic behavior, the master equations and corresponding PDEs for the Myopic Remain model are the same as in the Myopic Reset model for cells in stage $s = 2, \dots, K-1$ [see Eq. (2)]. The master equation for stage K cells is as follows:

$$\begin{aligned}
C_{i,j}^{(K)}(t + \tau) = & C_{i,j}^{(K)} + \frac{r_m \tau}{4} (1 - C_{i,j}) (C_{i-1,j}^{(K)} + C_{i+1,j}^{(K)} + C_{i,j-1}^{(K)} + C_{i,j+1}^{(K)}) \\
& - \frac{r_m \tau}{4} C_{i,j}^{(K)} (4 - C_{i-1,j} - C_{i+1,j} - C_{i,j-1} - C_{i,j+1}) \\
& + \lambda C_{i,j}^{(K-1)} \tau - \lambda \tau C_{i,j}^{(K)} (1 - C_{i-1,j} C_{i+1,j} C_{i,j-1} C_{i,j+1}) + O(\tau^2),
\end{aligned} \tag{16}$$

This is the same as Eq. (7) for stage K cells under the Remain model without myopic behavior except from the term on the fourth line of the right-hand side as the probability that a cell exhibiting myopic behavior successfully proliferates is the probability that a cell has an empty neighbor. Using Eqs. (2), (15), and (16), we can derive the following system of PDEs as we have done for our previous models:

$$\begin{aligned}
\frac{\partial C^{(1)}}{\partial t} &= D[(1 - C)\nabla^2 C^{(1)} + C^{(1)}\nabla^2 C] + 2\lambda C^{(K)}(1 - C^4) - \lambda C^{(1)}, \\
\frac{\partial C^{(s)}}{\partial t} &= D[(1 - C)\nabla^2 C^{(s)} + C^{(s)}\nabla^2 C] + \lambda C^{(s-1)} - \lambda C^{(s)}, \quad \text{if } s = 2, \dots, K-1, \\
\frac{\partial C^{(K)}}{\partial t} &= D[(1 - C)\nabla^2 C^{(K)} + C^{(K)}\nabla^2 C] + \lambda C^{(K-1)} - \lambda C^{(K)}(1 - C^4).
\end{aligned} \tag{17}$$

As is the case for the models without myopic behavior [see Eqs. (5) and (9)], the PDE describing the evolution of the total density of cells proliferating under the Myopic Remain model is the same as PDE (14), which describes the density of cells proliferating under the Myopic Reset model:

$$\frac{\partial C}{\partial t} = D\nabla^2 C + \lambda C^{(K)}(1 - C^4). \tag{18}$$

For the K -stage model, we have derived four corresponding systems of PDEs (3), (8), (13), and (17), which capture the full combination of reset versus remain and myopic versus nonmyopic behaviors. From these PDEs, we can immediately observe behavior that we had observed previously from their respective ABMs. The Remain model PDEs [(8) and (17)] contain a density-dependent proliferation term in the equation for stage- K cells. This means we expect cells to arrest in stage K , which agrees with our ABM simulations. From the PDEs (5) and (14) which describe the evolution of the total density of cells over time under the nonmyopic and myopic models, respectively, we see that cell numbers grow more quickly under the Remain model than the Reset model as we

expect to see more cells in stage K under the Remain model than in the Reset model. We further observe that myopic cells have a proliferative advantage over nonmyopic cells due to the weaker density dependence in the proliferative terms [$\lambda(1 - C)$ for PDEs (5) compared to $\lambda(1 - C^4)$ for PDEs (14)].

Although these PDEs have characteristics which match their respective ABMs, it is important to note that we have derived these continuum PDE limits under various assumptions. In particular, we assume that as $\tau, \Delta \rightarrow 0$ we have that $r_m \Delta^2$ and λ are held constant which implies that

$$\lim_{\tau, \Delta \rightarrow 0} \frac{\lambda}{r_m} = 0. \tag{19}$$

This condition requires that the motility and progression rate parameters to satisfy $r_m \gg \lambda$ in order to achieve good agreement between the ABM simulations and the corresponding continuum PDE limit.

In this section we have derived continuum PDEs (under the mean-field approximation) which correspond to the ABMs we presented in Sec. II. We observed how incorporating the

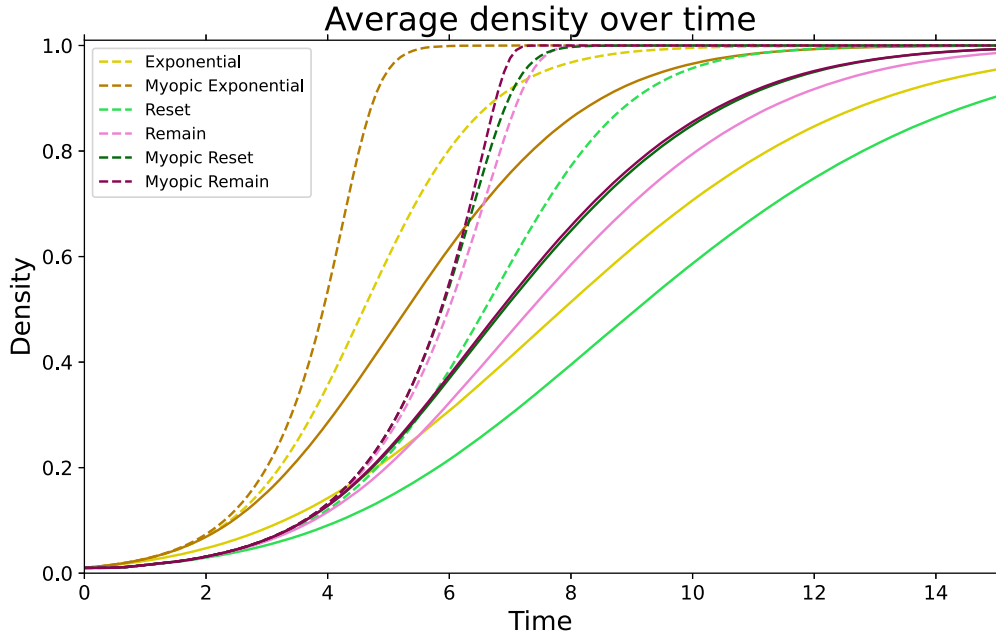


FIG. 7. Comparing the numerical solutions of the PDEs (5), (9), (14), (18) to the average ABM behavior over 15 time units. We initialize with the following uniform initial conditions $C^{(1)}(0) = C(0) = 0.01$ and $C^{(s)} = 0$ for $s = 2, \dots, K$, and parameters $r_m = r_p = 1$, so that the initial conditions and parameters for the PDEs and ABMs are identical to the ABM simulations in Fig. 6. We obtain the solutions to the corresponding ODE systems using the RK45 method from the `solve_ivp` function from the `Scipy` library in `Python`. PDE solutions are denoted using dashed lines, and we distinguish between different models by using different colors. For comparison to the ABMs, we plot the average density evolutions over time from Fig. 6 using the same color scheme to denote models as in the PDEs, but in solid lines.

multistage representation of the cell cycle changes the system of PDEs. Reset, remain, and myopic behavior within our ABMs lead to different density-dependent terms within the derived systems of related PDEs. In the next section, we will use the continuum PDEs to investigate the behavior of the various proliferation models we have presented under different parameter regimes. We also identify parameter values under which the continuum models yield close agreements with their corresponding ABMs.

IV. THE EFFECTS OF CELL MOTILITY

It is well known that mean-field models fail to accurately approximate the behavior of discrete ABMs when the motility rate of a system is insufficiently high compared to the proliferation rate [19,42]. We will demonstrate that this is the case by comparing the average ABM behavior over many realizations and their continuum limits at high densities. We investigate the behavior of numerical solutions of the derived PDEs under uniform initial conditions. Using the pair correlation function (PCF), a measure of the pairwise correlation of a lattice, we will qualitatively predict the degree of agreement between the ensemble averages of ABMs and their corresponding continuum limits. We will compare the average ABM behavior in our different proliferation models to the continuum limit, and demonstrate through numerical simulations that the PCF can be used to explain the observed discrepancies.

Throughout this section we consider a uniform initial condition with periodic boundary condition in a box $[0, L_x] \times [0, L_y]$ and compare the solutions of our PDEs to the behavior of their corresponding ABMs from Sec. II. Studying

uniform initial conditions is relevant to modeling growth-to-confluence assays [43,44]. Given a constant initial density of cells $C^{(s)}(x, y, 0) = \kappa^{(s)}$ for some $\kappa^{(s)} \in (0, 1)$ in each stage $s = 1, \dots, K$ and periodic boundary conditions, the solutions to the systems of PDEs (3), (8), (13), and (17) are not spatially dependent. This means that solutions are of the form $C^{(s)}(x, y, t) = C^{(s)}(t)$, and we may reduce the problem of solving a system of PDEs to solving a system of ODEs. We can do this by dropping the spatial terms from the PDEs (3), (8), (13), and (17). Notably, under a uniform initial condition, the solutions are independent of the diffusion coefficient D . This is because the uniform initial condition, combined with the translationally invariant mechanisms, preserves uniformity through all time.

For general initial conditions, as we increase the motility rate, r_m , in ABM simulations the mean-field assumption becomes more appropriate. Therefore, for a fixed progression rate λ and a given uniform initial condition, we may interpret the solutions of the continuum PDEs as describing the limiting behavior of the corresponding ABMs as we take the motility rate $r_m \rightarrow \infty$.

In Fig. 7 we compare the density evolutions for each PDE to their corresponding ABM using the simulations we performed in Fig. 6 where we set the motility rate to be $r_m = 1$. Every continuum PDE predicts a quicker rate of growth than the corresponding ABM. The poor agreements between our PDE models and ABMs can be explained by the low motility rate relative to the progression rate in the ABM. Due to the relatively low motility rate, pairwise spatial correlations build up during ABM simulations, which means that the neighboring sites of a cell are more likely to be occupied on a lattice

of the same density in which cell positions are uncorrelated. This results in populations of cells growing more slowly in our ABMs compared to our PDEs.

We further observe that the proliferative advantage over the nonmyopic Exponential model conferred by myopic and remain behavior in the multistage model becomes apparent only at higher densities (at around density 0.9) in the continuum PDEs when compared to the corresponding ABM simulations (at around density 0.2). Since the proliferative advantages of myopic and remain behavior arises exclusively when a cell attempts to move from stage K to stage 1, the impact of these models depend strongly on local cell density. Given a fixed cell density, cells in well-mixed systems (corresponding to ABMs with high motility rates) are less likely to have occupied neighboring sites than in ABMs initialized with low motility rates. This implies that proliferation is less restricted by local density in the continuum PDE framework than in the low motility ABM simulations. Consequently, the relative proliferative benefit of myopic and remain behavior is reduced compared with systems that exhibit pairwise spatial correlations. In the absence of volume exclusion effects, we know cell populations grow more rapidly under the Exponential model than in multistage models. We suggest that the Exponential model maintains a higher growth rate for a longer period of time in the continuum PDE compared to the corresponding ABM due to the weakened density dependence.

Another difference between the PDE solutions and ABM simulations is that, at higher densities, we are able to distinguish between the Myopic Reset and Myopic Remain models more easily in the PDE evolutions than in the average ABM density evolutions. Recall that when cells proliferate under myopic behavior, they fail to proliferate only if all adjacent sites are occupied. It is therefore unlikely that a cell fails to proliferate until reaching higher density environments when exhibiting myopic behavior. Under the assumption that there are no spatial correlations, cell proliferation is not limited by the ability for neighboring cells to move out of the way but by the global density of the environment. This means that the more frequent proliferation attempts under the Myopic Remain model will lead to quicker growth than in the Myopic Reset model but not until higher densities.

Despite these differences, we note there are also similarities when comparing the qualitative behavior of the PDE solutions to that of the average ABM behavior over many realizations. For example, we can see that myopic behavior leads to quicker population growth rates under all cell proliferation models in both the PDE models and average ABM behavior. Furthermore, we can see that the Remain model does offer a proliferative advantage over the Reset model in the continuum PDEs as well as in the ABMs. We also see that at low densities, cells grow more quickly under the Exponential model than in the multistage models in the solutions of PDEs and average ABM behavior.

We have demonstrated that under low motility rates, the continuum descriptions of our cell proliferation models provide poor agreements with their associated ABMs. This is due to pairwise spatial correlations which build up over time. We would like a method to quantify the degree of spatial correlation which arises from a given ABM. We will use this method to investigate how spatial correlations impact the degree of

agreement between ABMs and their continuum PDE solutions under a uniform initial condition.

A. Pair correlation function

Pair correlation functions (PCFs) can be used to determine the degree of spatial correlation in a lattice filled with cells. We define a PCF following previous notation from work analyzing different PCFs [45]. Given a $(X \times Y)$ lattice of cells, let M be the occupancy matrix whose entries denote whether a lattice site is occupied or not. If a site (x, y) is occupied, we set $M_{xy} = 1$, otherwise we set $M_{xy} = 0$. Given an occupancy matrix M , let ψ^M be the set of all pairs of cells in the lattice which can be written as follows:

$$\psi^M = \{(\mathbf{a}, \mathbf{b}) \in \mathbb{L} \times \mathbb{L} : \mathbf{a} \neq \mathbf{b}, M_{\mathbf{a}} = M_{\mathbf{b}} = 1\}, \quad (20)$$

where $\mathbb{L} = \{1, \dots, X\} \times \{1, \dots, Y\}$ denotes the set of possible sites of a lattice. In order to measure the degree of spatial correlations within a lattice, we need to determine a metric for our lattices. As our ABMs use the von Neumann neighborhood for both proliferation and movement, and they are periodic, the appropriate metric is given in the following equation. For $(\mathbf{a}, \mathbf{b}) \in \mathbb{L} \times \mathbb{L}$,

$$\begin{aligned} \|\mathbf{a} - \mathbf{b}\| &= \|(x_1, y_1) - (x_2, y_2)\| \\ &= \min\{|x_1 - x_2|, X - |x_1 - x_2|\} \\ &\quad + \min\{|y_1 - y_2|, Y - |y_1 - y_2|\}. \end{aligned} \quad (21)$$

This is not the Euclidean (L^2) metric, but the periodic L^1 metric instead. The periodic L^1 metric is more appropriate for calculating the PCF because our ABMs use von Neumann neighborhoods and periodic boundary conditions. This means that the standard Euclidean metric would not correctly capture the distances between sites in our on-lattice ABMs. This metric has been considered previously by Gavagnin *et al.* [45] where the authors derived PCFs that are more appropriate for lattice-based cell configurations.

Using definitions (20) and (21) we can define the subsets of cell pairs separated by distance d in matrix M to be

$$C^M(d) = \{(\mathbf{a}, \mathbf{b}) \in \psi^M : \|\mathbf{a} - \mathbf{b}\| = d\}. \quad (22)$$

We define the PCF for a given matrix M to be the ratio between the number of pairs of cells separated by distance d , and the expected number of pairs separated by distance d within a uniformly populated lattice at the same density. Let N be the number of occupied sites in the matrix M (i.e., the number of nonzero entries), and U be a random matrix with the same dimensions as M with entries $U_{xy} = 1$ for N sites chosen uniformly at random without replacement and $U_{xy} = 0$ otherwise. We define the PCF for a matrix M as follows:

$$f^M(d) = \frac{|C^M(d)|}{\mathbb{E}[C^U(d)]}. \quad (23)$$

Gavagnin *et al.* [45] show that the expected number of sites separated by distance d under metric (21) is $\mathbb{E}[C^U(d)] = \frac{2dN(N-1)}{L_x L_y - 1}$. This means we can rewrite the PCF (23) as

$$f^M(d) = \frac{(XY - 1)|C^M(d)|}{2dN(N - 1)}.$$

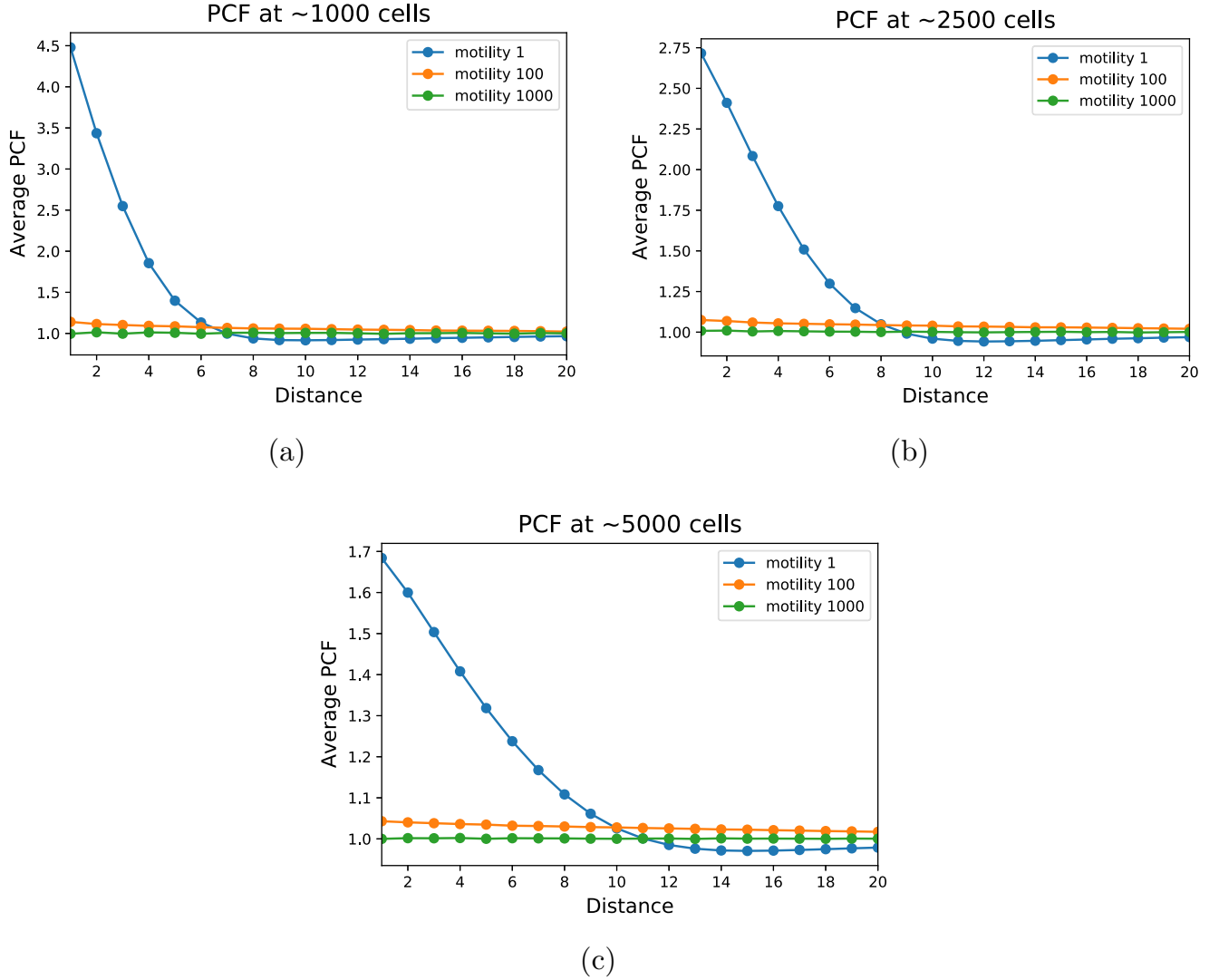


FIG. 8. Plotting the PCFs $\hat{f}(d)$ for $d = 1, 2, \dots, 20$ averaged over 20 realizations of the Myopic Remain ABM for motility rates $r_m = 1, 100, 1000$ with a fixed proliferation rate $r_p = 1$ and number of stages $K = 10$. Panels (a), (b), and (c) give the average PCF at a time where the average number of cells over the 20 realizations (for the different motility rates) first exceeds 1000, 2500, and 5000, respectively.

Given lattices from multiple realizations of a cell proliferation ABM, we can calculate the PCF for each individual lattice and find the average value of the PCF, $\hat{f}(d)$ over all realizations. If $\hat{f}(d) = 1$ for all distances d , this means that the occupied sites of the lattice under the ABM are uniformly distributed which implies there are no spatial correlations. If $\hat{f}(d) > 1$ for small distances d , then there are more pairs of occupied sites which are close together than we would expect under a uniformly distributed lattice. This would imply the ABM gives rise to positive spatial correlations at short distances. Similarly, an average PCF $\hat{f}(d) < 1$ for small d suggests the ABM causes negative spatial correlations at short distances. More generally, larger quantities $|\hat{f}(d) - 1|$ suggests a stronger spatial correlation.

We see in the plots in Fig. 8 that for motility rate 1, the average PCFs for $d = 1, \dots, 7$ are clearly above 1, which suggests the ABM causes strong spatial correlations. Indeed, in Fig. 9(a), we see that an example realization of a Myopic Remain ABM with motility rate 1 exhibits strong positive correlations at short distances when we reach around 2500 cells.

Increasing the motility rate to $r_m = 100$ lessens the degree of spatial correlation when compared to motility rate $r_m = 1$; however, the PCF plots in Fig. 8 show that $\hat{f}(d) > 1$ for small distances d . This suggests there is still some spatial correlation at motility $r_m = 100$, which can be seen in Fig. 9(b) where the spatial correlation is present but clearly much weaker than in Fig. 9(a). When we increase the motility rate to $r_m = 1000$, then we can see in Fig. 8 that the average PCF satisfies $\hat{f}(d) \approx 1$ for all distances $d = 1, \dots, 20$. This suggests that motility rate $r_m = 1000$ is sufficiently high to remove almost all spatial correlations within the lattice for this proliferation rate. We can observe that this holds in the example lattice in Fig. 9(c), where there appears (by eye) to be no spatial correlation beyond what would be expected on a lattice of uniformly distributed cells. In Appendix B we demonstrate that the PCF behaves similarly for every other ABM.

Now that we have a method to measure the degree of spatial correlation of an ABM, we are able to predict how well the solution to a continuum PDE approximates the average behavior over many realizations of its corresponding ABM.

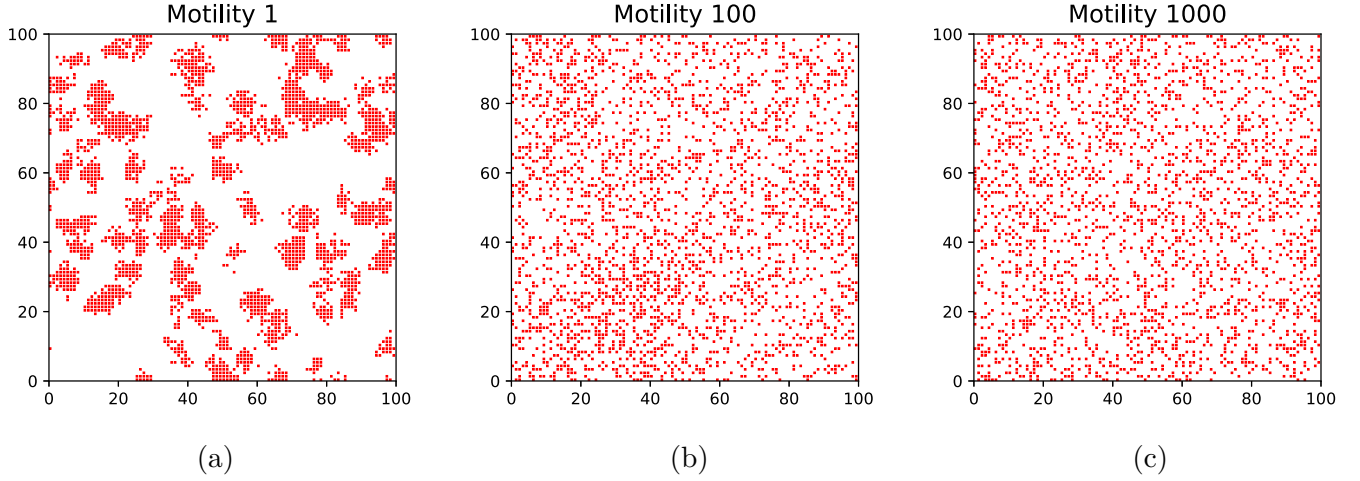


FIG. 9. Example occupancy matrices, M , from the simulations in Fig. 8(b). Each lattice is taken at the first time step where the cell count first exceeds 2500 cells. Panels (a), (b), and (c), correspond to motility rates $r_m = 1$, 100, and 1000, respectively.

We will demonstrate that the PCF allows us to successfully predict which motility rates give good agreement between the ABM and continuum PDE.

B. Comparing ABMs to continuum PDEs

Throughout this section we will only consider the Myopic Remain model as this is the most complex model. We can perform analysis of every other model using similar methods. In order to determine a sufficiently high motility rate for good agreement between the continuum PDEs and corresponding ABMs for the different cell proliferation models which we have presented, we can calculate the average PCFs under these ABMs over many realizations for different motility rates until $\hat{f}(d) \approx 1$ for all distances d .

In Fig. 10 we plot the average column density in each ABM. Given occupancy matrices at time t $M(t)$ of a realization, we can calculate the density of column i , $\hat{L}_i(t)$, of a lattice at time t as follows:

$$\hat{L}_i(t) = \frac{1}{Y} \sum_{j=1}^Y M_{ij}. \quad (24)$$

We begin by noting that the average column density for ABMs is not constant due to the inherent stochasticity of our cell proliferation models. This is in contrast to the uniform profile over the whole domain when cell populations evolve under the system of PDEs. The variance in column densities is larger for smaller motility rates due to the the larger pairwise spatial correlations that build up at short distances for these smaller motility rates. When we average over a finite number of repeats, the column densities are not spread out as effectively for low motilities as they are for high motilities. Increasing the number of repeats we perform, we would see that the average column densities become increasingly uniform.

We also observe that the agreement between the ABM and continuum PDE improves as we increase the motility rate. This is expected behavior given the PCFs in Fig. 8, which show $|\hat{f}(d) - 1|$ decreases as we increase the motility rate for all distances d . At motility rate $r_m = 1$, the agreement

is poor at $t = 5$ and very poor at $t = 7$, which we expect given that the PCF $|\hat{f}(d)|$ is much larger than 1 under motility rate $r_m = 1$. Increasing the motility rate to $r_m = 100$ in the ABM gives a better agreement to the continuum PDE for $t = 5$, but a poor agreement at $t = 7$ (although still an improvement on motility rate $r_m = 1$). This is because the PCF satisfies $|\hat{f}(d) - 1| > 0$ for $d \leq 20$, which means that there are positive spatial correlations in the ABM at these distances. Eventually, these build up enough to slow down the growth of the ABM when compared to the PDE. We see that motility rate $r_m = 1000$ is sufficient to give good agreement between the ABM and the PDE at both times $t = 5, 7$. As the PCF $\hat{f}(d) \approx 1$ at this motility rate, this suggests that the system is well-mixed, which explains the good agreement between the ABM and PDE.

In Fig. 11 we observe across all models, that the error between the discrete and continuum models decreases when we increase from motility rate $r_m = 1$ to motility rate $r_m = 100$.

In general, it is true that PCFs that demonstrate weaker spatial correlations are correlated with better agreement between the average discrete ABM behavior and the continuum PDE solution. To show this, we considered the average column error between the ABM and PDE, $E(t)$, which can be calculated as

$$E(t) = \frac{1}{X} \sum_{i=1}^X |\hat{L}_i(t) - C_i(t)|, \quad (25)$$

where $L_i(t)$ is the average density of column i at time t of the ABM, and $C_i(t)$ is the solution of the corresponding continuum PDE evaluated at time t and the center of lattice point i . We observe that in Fig. 12 that larger motility rates lead to a smaller error between the discrete and continuum model. Furthermore, we can also see that the PCF acts as a good predictor for the ABM-PDE error.

In this section we have observed through numerical simulations that when the motility rate of cells is low, there is poor quantitative agreement between the average ABM behavior (over many realizations) and the related continuum PDE. This is because under parameter regimes where motility is low, the

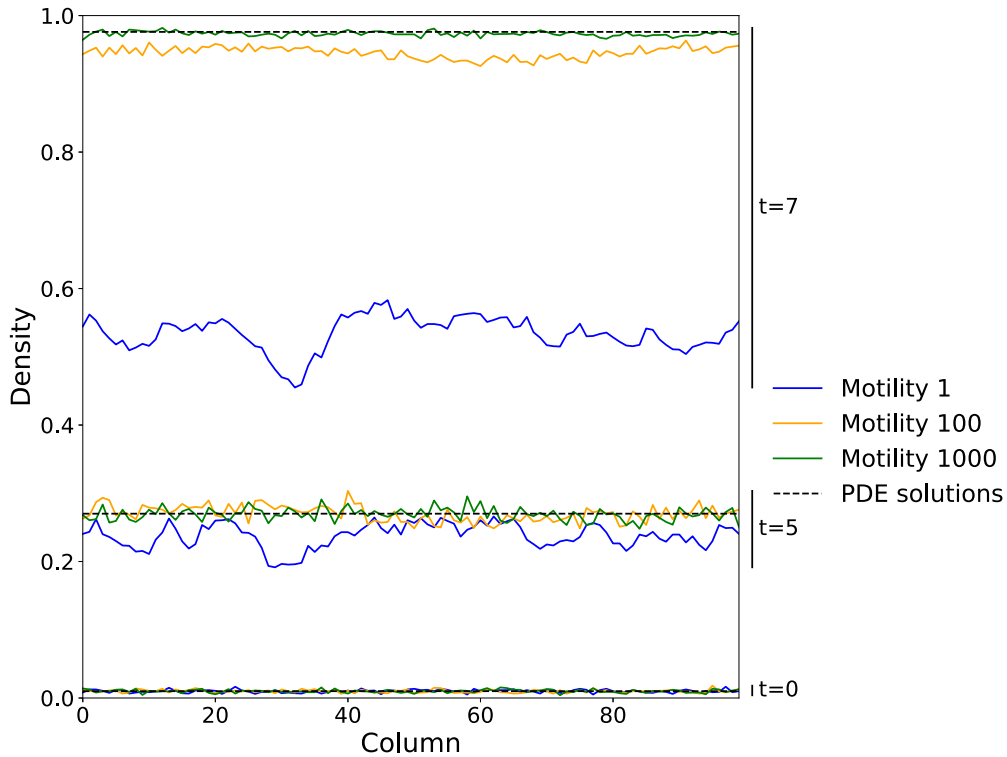


FIG. 10. Comparing the solutions to PDE (18) to the average ABM column densities over 20 realizations at times $t = 0, 5$, and 7 . Parameters for the ABMs are the same as in Fig. 8 and we initialize with a uniform density of 1% as in the simulations for Figs. 6 and 7. We distinguish between motility rates $r_m = 1, 100, 1000$ in ABMs by using different colors and denote the constant solution to PDE (18) over these columns with dashed black line at each time.

assumptions we made in Sec. III to derive continuum models do not hold. In order to quantify the agreement between an ABM and the corresponding continuum PDE, we introduced the PCF, which enabled us to measure the strength of pairwise

correlations in an ABM for a given motility rate. Under a uniform initial condition, we demonstrated that when the PCF differs from 1, the agreement between the continuum PDE and the average ABM behavior is poor, whereas we observe good

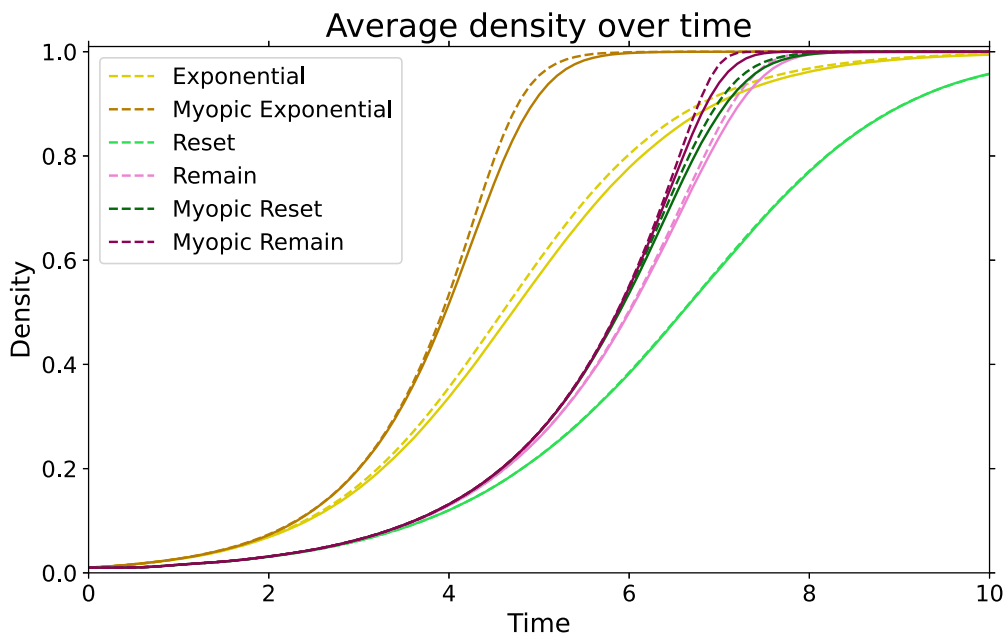


FIG. 11. As in Fig. 7, we compare the average behavior of ABMs to the numerical solutions of the corresponding PDEs (5), (9), (14), and (18). We initialize identically to the simulations in Fig. 7 and increase to motility rate $r_m = 100$ in our ABM simulations.

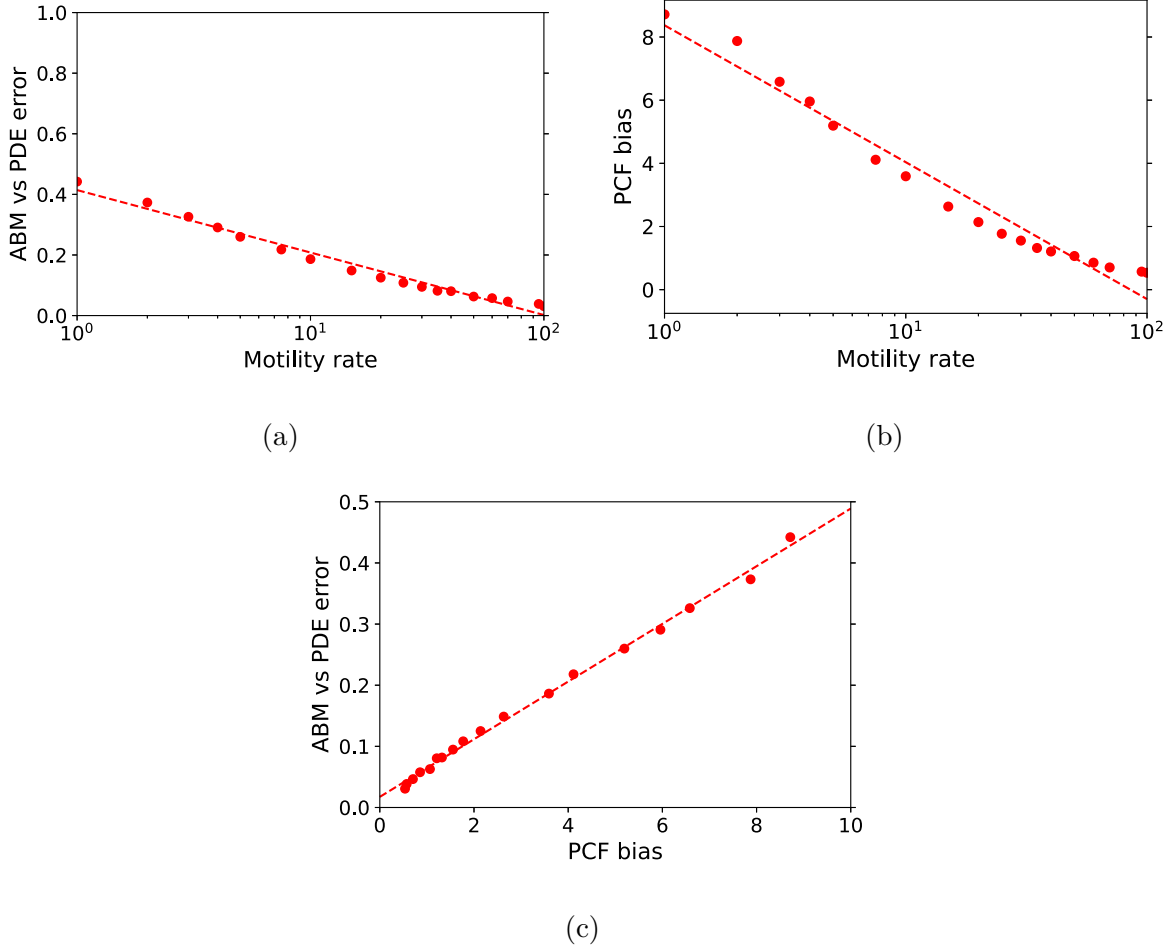


FIG. 12. We simulate the Myopic Remain ABM with $K = 10$ stages with various motility rates r_m between 1 and 100 and fixed proliferation rate $r_p = 1$ for $t = 7$ time units. At each motility rate, we perform 20 repeats, and calculate the average PCF $\hat{f}(d)$ at $t = 1$. For each ABM realization, we calculate the error at the final time of the simulation, $E(7)$, between the ABM solution and the corresponding continuum PDE solution. We considered the average error over all repeats. In panel (a), we compare the motility rate to the error, in panel (b) we compare the motility rate to the PCF bias given by $|\hat{f}(d) - 1|$, and in panel (c) we compare the PCF bias to the error. For each panel, we have plotted the line of best fit through our numerical simulations.

agreement when the PCF is approximately equal to 1 for all pairwise distances. We will now consider nonuniform initial conditions which lead to traveling waves, and study the behavior of our ABMs and corresponding PDEs in this context.

V. TRAVELING WAVES

In tissue invasion barrier assays performed by Streichan *et al.* [35], cell invasion was shown to be density-dependent. At high local densities, cells were shown to arrest and may reactivate when space is available due to cells moving into the free space created when the barrier is removed. Models capturing this behavior have been previously explored. For example, Falcó *et al.* [30] proposed a coupled system of two PDEs which each incorporated density-dependent terms to model the evolution of cell densities labeled by the FUCCI cell-cycle construct [46], with red cells in G_0/G_1 and green cells in $S/G_2/M$. Using this model, they derived wave speeds for the traveling waves. Vittadello *et al.* [29] studied continuum PDEs corresponding to a three-stage Remain model, which included an intermediate labeled stage representing

early S -phase cells. Similarly, they derived wave speeds for this specific system of PDEs.

In this section we will use the systems of PDEs (3), (8), (13), and (17) to model cell invasion for a generic K -stage model. Recall that these equations correspond to the different combinations of reset or remain and nonmyopic or myopic behavior. We first give theory which allows us to calculate the minimum speed of a steady-state wave front for these continuum PDE models. We then run numerical simulations to confirm that the theoretical predictions hold for our systems of PDEs. We also determine under which parameter regimes these theoretical predictions are accurate for the corresponding ABMs. For different ABMs, we compare how traveling waves form and discuss why the behavior in each ABM gives rise to these characteristics.

A. Theoretical wave speeds

We study the theoretical wave speeds of the systems of PDEs (3), (8), (13), and (17). We consider 1D traveling waves by taking the average column density of 2D PDEs initialized

with initial conditions that are translationally invariant in the y direction and with appropriate boundary conditions. Given a continuous 2D concentration over time $C(x, y, t)$, we can write the average column density as

$$\hat{C}(x, t) = \frac{1}{Y} \int_0^Y C(x, y, t) dy. \quad (26)$$

Using Eq. (26) we can rewrite the systems of PDEs (3), (8), (13), and (17) in terms of $\hat{C}$ by integrating over y and

then dividing through by Y . This allows us to study a reduced 1D system of PDEs. If we further assume that the initial condition $C^{(s)}(x, y, 0)$ is translationally invariant in y and we have zero-flux boundary conditions on the horizontal boundaries then $C(x, y, t) = \hat{C}(x, t)$ for any value y . Then, we can explicitly write the 1D systems of PDEs by replacing every C by $\hat{C}$ and $\nabla^2 C$ by $\frac{\partial^2 \hat{C}}{\partial x^2}$. For example, the system of PDEs (17) for the Myopic Remain model can be written

$$\begin{aligned} \frac{\partial \hat{C}^{(1)}}{\partial t} &= D \left[(1 - \hat{C}) \frac{\partial^2 \hat{C}^{(1)}}{\partial x^2} + \hat{C}^{(1)} \frac{\partial^2 \hat{C}}{\partial x^2} \right] + 2\lambda \hat{C}^{(K)} (1 - \hat{C}^4) - \lambda \hat{C}^{(1)}, \\ \frac{\partial \hat{C}^{(s)}}{\partial t} &= D \left[(1 - \hat{C}) \frac{\partial^2 \hat{C}^{(s)}}{\partial x^2} + \hat{C}^{(s)} \frac{\partial^2 \hat{C}}{\partial x^2} \right] + \lambda \hat{C}^{(s-1)} - \lambda \hat{C}^{(s)}, \quad \text{if } s = 2, \dots, K-1, \\ \frac{\partial \hat{C}^{(K)}}{\partial t} &= D \left[(1 - \hat{C}) \frac{\partial^2 \hat{C}^{(K)}}{\partial x^2} + \hat{C}^{(K)} \frac{\partial^2 \hat{C}}{\partial x^2} \right] + \lambda \hat{C}^{(K-1)} - \lambda \hat{C}^{(K)} (1 - \hat{C}^4). \end{aligned} \quad (27)$$

For convenience, we will simply write $\hat{C} = C$ throughout the rest of this section.

If we linearize the reduced 1D system of PDEs (27) about the unstable steady state $C^{(s)} = 0$, we obtain the following linear system of PDEs:

$$\begin{aligned} \frac{\partial C^{(1)}}{\partial t} &= D \frac{\partial^2 C^{(1)}}{\partial x^2} + 2\lambda C^{(K)} - \lambda C^{(1)}, \\ \frac{\partial C^{(s)}}{\partial t} &= D \frac{\partial^2 C^{(s)}}{\partial x^2} + \lambda C^{(s-1)} - \lambda C^{(s)} \quad \text{if } s = 2, \dots, K. \end{aligned} \quad (28)$$

In fact, the linearization of the other 1D reductions of the systems of PDEs (3), (8), and (13) gives the same linear system of PDEs (28). This suggests that, whenever the continuum limits accurately describe the discrete ABMs, the predicted wave speeds will be identical under the different cell proliferation models with the same number of stages K . In particular, different methods of density-dependent regulation does not alter the theoretical wave speed estimate. This is because the wave speed is determined by low-density dynamics, where crowding effects are less impactful.

We assume that solutions take the form $C(x, t) = C(x - vt)$ and satisfy $\lim_{x \rightarrow -\infty} C(x) = 1$, $\lim_{x \rightarrow \infty} C(x) = 0$. This means we have traveling wave front which move rightwards with wave speed v . Using the front propagation method by Van Saarloos [47], Gavagnin *et al.* [23] show that the minimum wave speed in a K -stage model is

$$v = 2\sqrt{D\lambda(2^{1/K} - 1)}. \quad (29)$$

Taking the number of stages to be $K = 1$, we recover the classical invasion speed in the FKPP model, $2\sqrt{D\lambda}$. Note that increasing the number of stages K decreases the minimum wave speed v . Assuming the continuum limit of an ABM is a good approximation to the average of the agent-based dynamics, this suggests that cells proliferating under a multistage model invade more slowly than cells proliferating under an exponential model.

From Sec. IV we know that the motility rate must be sufficiently high relative to the proliferation rate for the continuum descriptions to give good approximations to the average behavior of the ABMs. In the next section, we will perform numerical simulations to investigate the traveling wave fronts which form when the continuum PDEs provide poor agreements to the ABMs. We also verify that the theoretical minimum wave speed (29) is valid under the assumption of sufficiently high motility rates.

B. Comparing ABM wave speeds to theoretical predictions

1. Low motility

We begin by considering a regime in which the motility rate, r_m , is of the same order as the proliferation rate, r_p (previously the low motility regime). In Fig. 13 we can see that the wave fronts which form in each ABM have similar shapes and are very steep. This is because cell proliferation is the major factor driving the traveling wave when the proliferation rate is comparable to the motility rate. This phenomenon has been observed in previous studies of traveling waves [47]. From Fig. 14 we observe that most ABMs exhibit different wave speeds to each other. This is in contrast to the theoretical predictions of the wave speed in Eq. (29) for high motility rates, r_m , where we expect models with the same number of stages, K , to have identical invasion speeds. We can see that the wave speed in Myopic models are all quicker than wave speeds in nonmyopic models with the same number of stages.

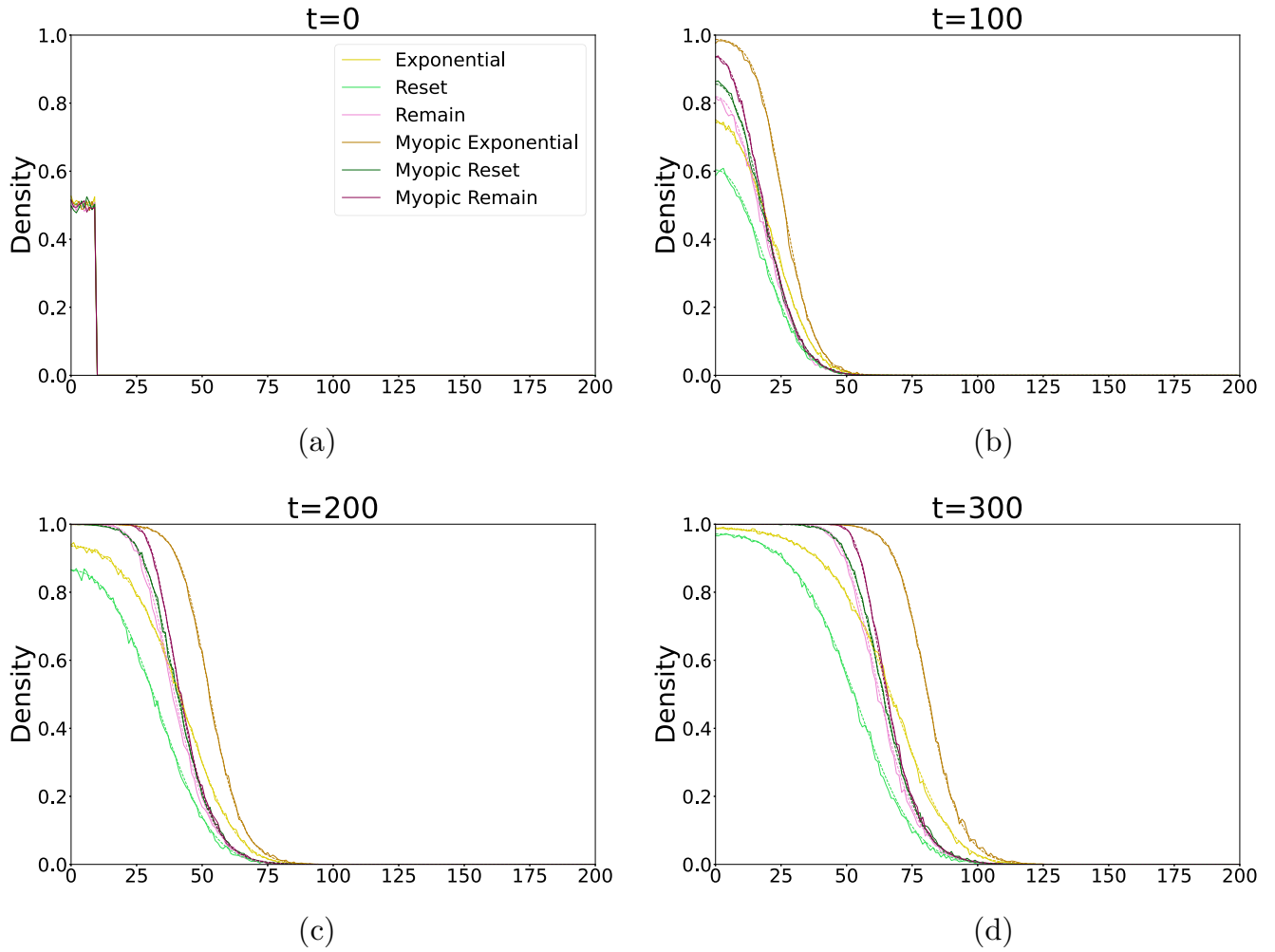


FIG. 13. ABM simulations of traveling wave fronts on a 200×100 lattice. We set the proliferation and motility rates to be equal $r_m = r_p = \sqrt{0.0233 \times 4}$, and the multistage models have $K = 10$ stages. We initialize our ABMs so that the first 10 columns are uniformly populated with stage 1 cells at 50% density [see (a)]. At the horizontal boundaries, we impose periodic boundary conditions, and at the vertical boundaries we impose reflecting boundary conditions. We simulate over $t = 300$ time units, and take the average column density over 20 realizations for each model. In each panel, we plot the column density for each ABM, distinguishing between different models using different colors. Panel (a) displays the initial conditions. Panels (b), (c), and (d) display the wave fronts at $t = 100$, $t = 200$, and $t = 300$ respectively.

We also observe that the speed of invasion in Remain models is quicker than Reset models due to cells at intermediate column densities likely having free space in the direction the traveling wave is moving. This means that cells which can attempt to proliferate more quickly after a failed division event, as is the case in the Remain model, have a proliferative advantage over cells proliferating under the Reset model. However, the Myopic Remain and Myopic Reset models have similar wave speeds. This is because cells exhibiting myopic behavior only fail to proliferate when they are completely surrounded by neighboring cells. Therefore, proliferation, and as a consequence the wave speed, is more limited by the rate at which cells move to create free space, i.e., the motility rate r_m . Recall further that the predicted minimum wave speed in Eq. (29) decreases as we increase the number of stages K . However, in Fig. 14 the Exponential model exhibits a lower wave speed than every model except the Reset model.

The different speeds of invasion in our different models, shown in Fig. 14, qualitatively match the rates of growth in

Fig. 6 at low to intermediate densities. As the wave front is driven by proliferation, the wave speed will correspond to the rate of growth of a population at a low density. Hence, the behavior of traveling waves at relatively low motility rates can be explained by the characteristics of the proliferation models which we described in Sec. II B.

As the motility rate is equal to the proliferation rate in the simulations we performed in Fig. 13 and 14, we do not expect good agreement between the average ABM behavior and the continuum PDE limit. Indeed, in Fig. 15 we can observe that the wave front in the ABMs featuring remain or myopic behavior agree poorly with their continuum PDE wave fronts. We can also see that the shape of the wave fronts for the Exponential and Reset ABMs do not match the shape of the wave fronts their continuum PDEs. However, there is some overlap between the profiles of the wave fronts. This is because the Reset and the nonmyopic Exponential ABMs travel at speeds similar to their corresponding continuum limits, as we saw in Fig. 14.

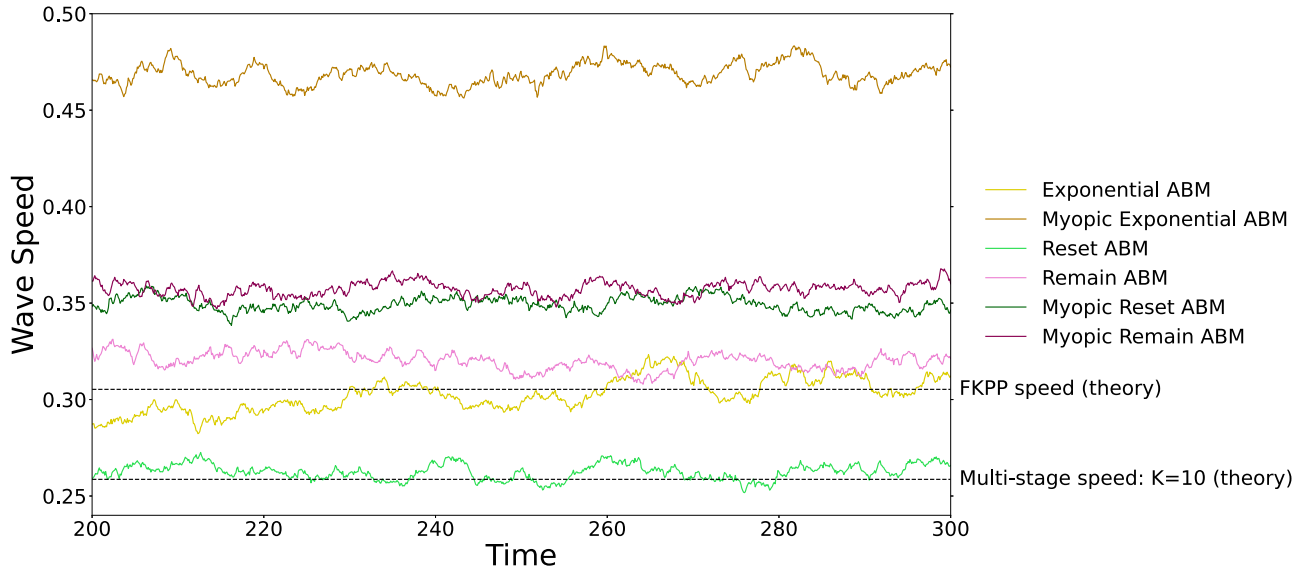


FIG. 14. Wave speeds for the ABM simulations in Fig. 13 between $t = 200$ and $t = 300$. We calculate the wave speed by finding the average mass added per unit time as in Harrison and Yates [48], and plot a moving average wave speed for each ABM with time window 5. We denote the wave speeds of different models using the same colors as Fig. 6, and plot the theoretical wave speeds for the FKPP and 10-stage model in black dashed lines.

In contrast, the ABMs evolve to a wave speed greater than the minimum theoretical wave speed of the continuum model, which results in poor agreement with the corresponding continuum limit. This is because the proliferation methods in these ABMs provide an advantage at the wave front which is not fully captured under the mean-field continuum PDEs. In low motility regimes in our ABMs, cells at the wave front are likely to be surrounded by three cells, even at low to intermediate densities. The empty neighboring space is likely

to be located forwards of the wave front. Cells which remain in the final stage of the cell cycle after a failed proliferation attempt can quickly reattempt to proliferate into the empty space. Cells which exhibit myopic behavior will always proliferate into the empty space. As the empty space is in the direction of the traveling wave front, this has the effect of pushing the wave forwards more quickly in the ABM than the continuum PDE where the direction of proliferation does not have this impact.

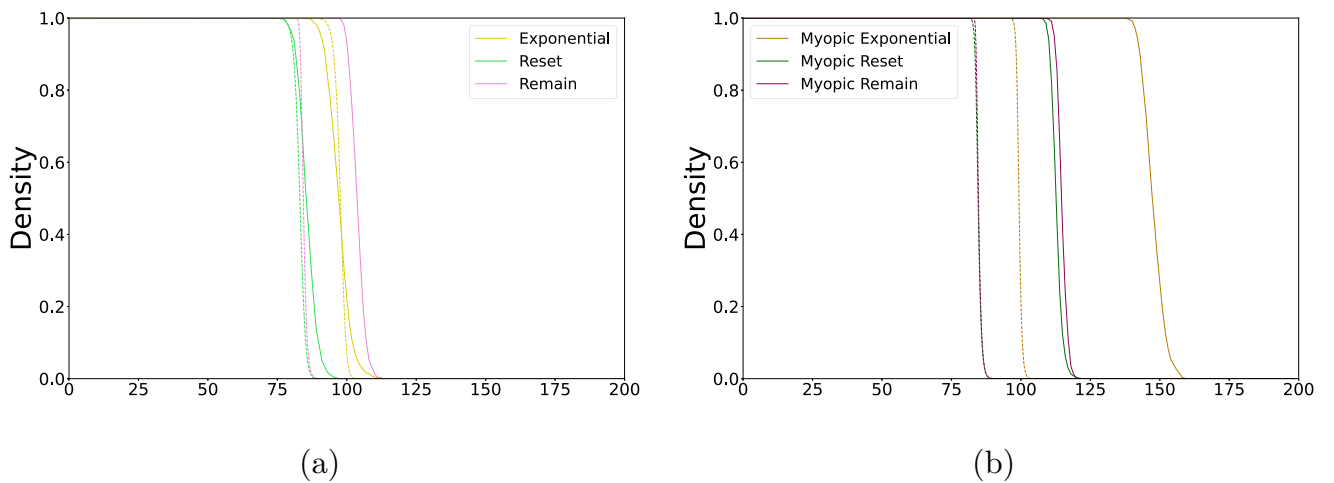


FIG. 15. Comparing the wave fronts at time $t = 300$ for each ABM to the continuum PDE solution at time $t = 300$. We initialize our PDE solutions with $C^{(1)}(x, 0) = 0.5$ for $x \leq 10$, $C^{(1)}(x, 0) = 0$ otherwise, and $C^{(s)}(x, 0) = 0$ for $s > 1$, so that the initial conditions of the PDEs match the corresponding ABM as in Fig. 13(a). We impose periodic boundary conditions on the horizontal axes, and zero-flux boundary conditions on the vertical axes. To solve the system of PDEs, we use the method of lines, discretizing the spatial component with a second-order finite-difference scheme using spatial grid size $\delta x = 0.25$, and integrating forward in time using the RK45 method from the `solve_ivp` function in the `scipy` package in Python. Panels (a) and (b) compare the nonmyopic and myopic ABM wave fronts, respectively, with their corresponding continuum PDE wave fronts. In each plot, we distinguish PDE and ABM wave fronts using dashed and solid lines respectively.

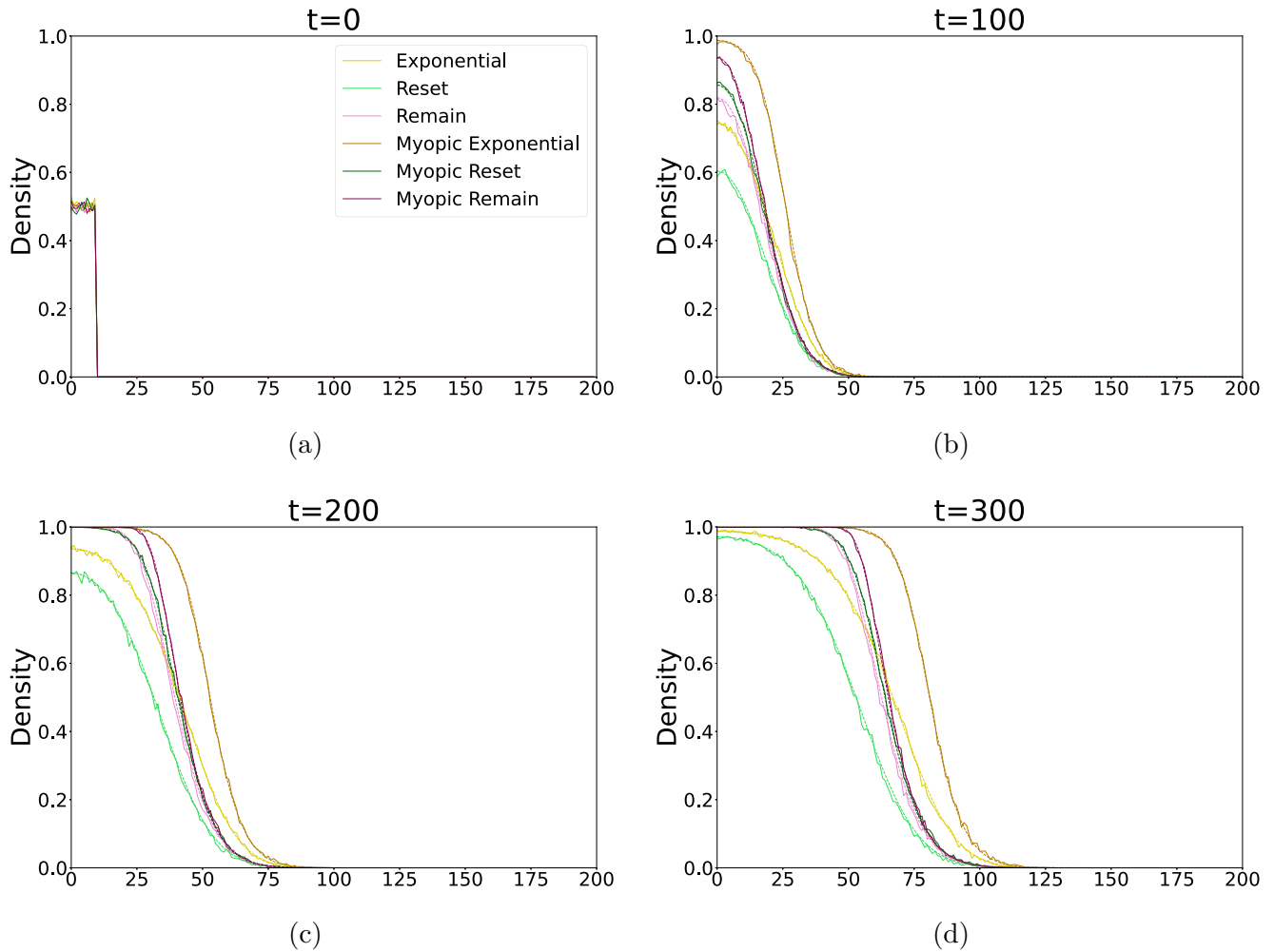


FIG. 16. Comparing ABM simulations to PDE solutions over 300 time units. ABM parameters are identical to the simulations from Fig. 13 but with motility rate $r_m = 4$ and proliferation rate $r_p = 0.0233$. We initialize and solve the PDEs identically to the simulations in Fig. 15, but with parameters updated to be consistent with the different motility and proliferation rates. We distinguish between different cell-cycle models using different colors and denote PDE and ABM wave fronts using dashed and solid lines respectively.

2. High motility

We now consider a parameter regime in which the motility rate is high relative to the proliferation rate. In Fig. 16 we see that even after 300 time units, the Exponential and Reset models have not yet reached a steady state wave front. We also note that the wave fronts across all models under this high motility scheme are shallower than in the wave fronts under a low motility scheme in Fig. 13. This is because the lower proliferation rate means that cells are less quick to fully occupy columns. We also observe that the agreement between the PDE and ABM wave fronts is very good across all models at every time point plotted. The improved agreement is due to the large motility rate relative to the proliferation rate which means that the continuum limits provide a better description of the discrete ABMs, as discussed in Sec. III.

In Fig. 17 we display asymptotic wave fronts after a long period of time. We can see that models with no proliferative advantage (nonmyopic Reset and nonmyopic Exponential) have a shallower wave-profile than models with a proliferative advantage. This is because these models are less quick to com-

pletely fill a column than cells proliferating under different models. In Fig. 18 we can see that the wave speeds of every multistage model are close. Similarly, the wave speeds for both Exponential models are close to identical. This is because the wave speed of the continuum PDEs are determined by the linearized system (28), which are identical for models with the same number of stages K . As the continuum PDEs provide almost exact agreements to their discrete ABMs, we would expect that traveling waves in ABMs under high motility regimes are characterized purely by the number of stages K in the cell cycle.

We also note that the average wave speeds for every model under the PDE are the minimum theoretical wave speed up to numerical error. The relative error between the theoretical and numerically solved PDE wave speeds can be well approximated using a result from Brunet and Derrida [49]. The numerical error is caused by floating point errors in Python, which causes the wave front to be cut off at small values.

In this section we have presented theoretical predictions of the wave speed of traveling wave fronts for our different models. Using numerical simulations, we confirmed that the

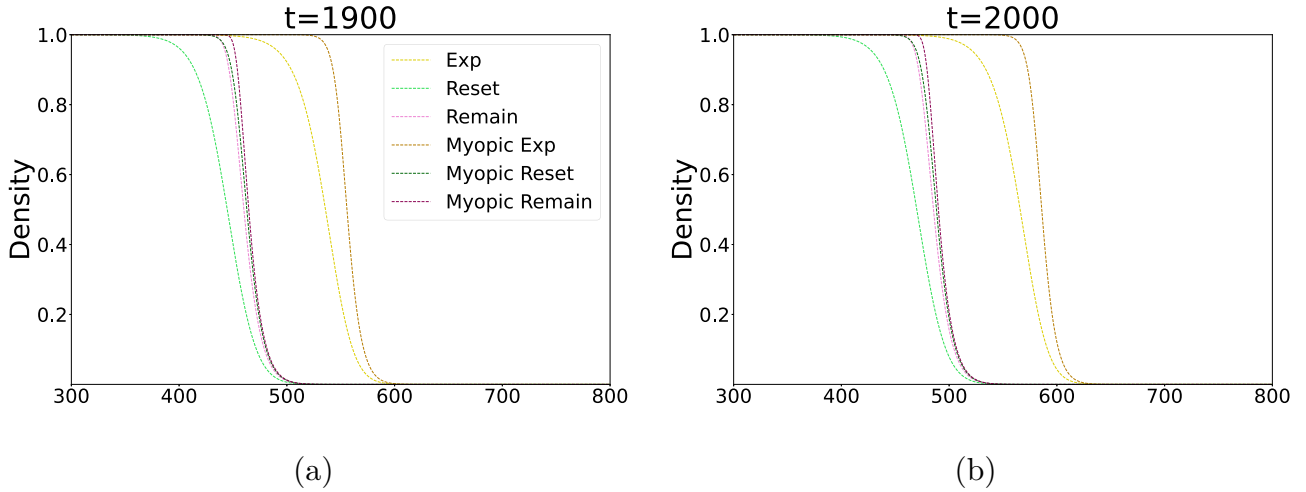


FIG. 17. PDE solutions at $t = 1900$ and $t = 2000$ with parameters and profiles initialized identically to Fig. 16. Panels (a) and (b) show PDE solutions at $t = 1900$ and $t = 2000$, respectively. As in previous figures, different models are represented using different colors.

theory holds well for parameter regimes in which ABMs and their corresponding continuum limits agree well. We have also demonstrated that even under low motility rates relative to the proliferation rate, the theoretical minimum PDE wave speed provides a lower bound for the corresponding ABM wave speeds.

VI. DISCUSSION

In this work we developed a framework for modeling the cell cycle which is capable of capturing realistic cell-cycle time distributions and accommodating a range of density-dependent cellular behaviors. Through discrete agent-based and continuum PDE modeling approaches (the PDEs being derived directly from the ABMs), we investigated how different cellular behaviors affect population-level growth. We incorporated several distinct ways in which cells arrest their

growth in high-density environments consistent with experimental observations showing that cell-cycle progression is constrained by the overall cell population [35].

We compared the growth dynamics of a cell population under these different modeling frameworks. Under the Reset model, a cell that attempts, but fails to proliferate is sent back to the first stage of the multistage cell cycle and must undergo the entire cell cycle again before reattempting to proliferate. In contrast, under the Remain model, cells enter a quiescent state in high local density environments and retain the ability to repeatedly attempt proliferation without the requirement to revisit earlier stages in the cell cycle. Through numerical experiments, we demonstrated that this leads to faster growth in the Remain model than in the Reset model. This result is consistently observed across both ABM and corresponding PDE simulations. Moreover, the capacity for cells to arrest in

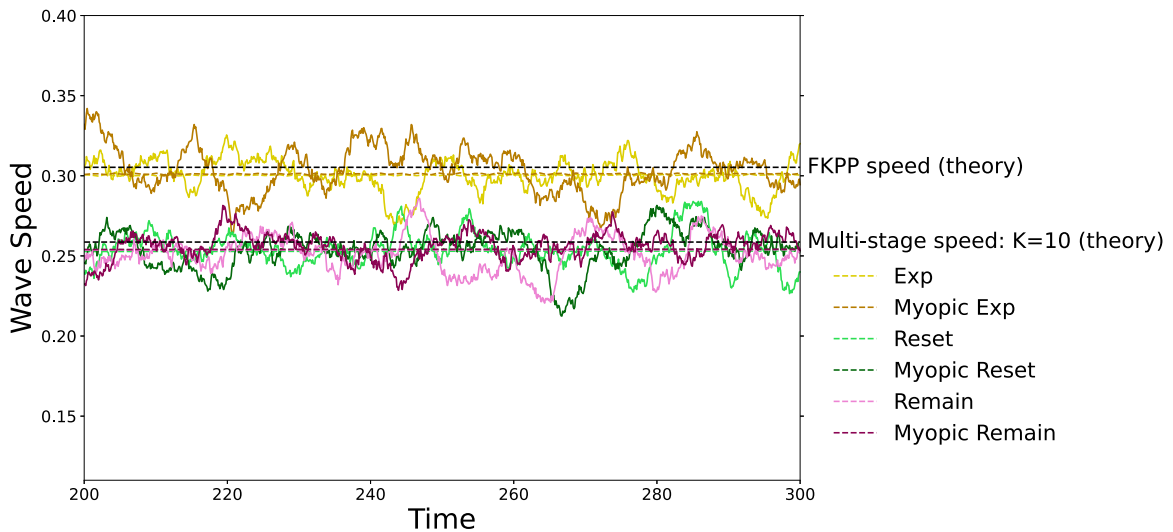


FIG. 18. Comparing ABM and PDE wave speeds. We calculated the PDE wave speeds using the simulations from Fig. 17. We used the steady-state wave profiles from Fig. 17 to initialize ABM simulations on a 800×20 lattice and simulated over 100 time units. We plot the moving average wave speeds over 20 realizations for each of our models. We distinguish between PDE and ABM wave speeds using dashed and solid lines respectively. As before, we differentiate between different models using different colors.

the final stage of the cell cycle allows for the Remain model to increase its density more quickly than the Exponential model at higher densities. This behavior contrasts with nonspatial theory, in which multistage models grow at a slower rate than the Exponential model [16].

We also introduced myopic behavior into our models, a novel adaptation which allows cells to sense their local environment for the purposes of proliferation. Numerical simulations demonstrated that the myopic property confers a greater proliferative advantage to cells than even the most rapidly growing nonmyopic model (the Remain model) throughout the entire growth process. Again, this behavior was consistent in both ABM and corresponding PDE simulations. Together, these results highlight that classifying cell-cycle times in volume-constrained proliferating cell populations is highly sensitive to the assumed underlying model of cell proliferation.

Although the ABM and PDE frameworks share similar qualitative behavior in growing cell populations, we demonstrated that, under certain parameter regimes, the PDE exhibits poor quantitative agreement with its corresponding ABM. In particular, the continuum PDE approximation overestimates cell density relative to the discrete ABM. We showed that this discrepancy arises because of the requirement that the cell motility rate must be sufficiently large relative to the proliferation rate for the continuum description of our ABMs to be valid. When the cell motility rate is insufficiently high in our ABMs, spatial correlations form which means that lattice sites are no longer independent of each other. This phenomenon leads to slower growth rates of cells in ABMs than the corresponding continuum limits. Numerical experiments with uniformly distributed initial conditions confirm that increasing cell motility, under a fixed proliferation rate, leads to faster population growth in ABMs. In contrast, solutions of the continuum PDEs are independent of the motility rate in initially uniformly distributed environments.

We further demonstrated that the cell motility rate strongly influences the spatial distribution of cells in the ABM. When motility rates are insufficiently high, pairwise spatial correlations emerge, which we quantify using the PCF. These observations show that on-lattice ABMs of cell proliferation can capture both spatial structure and growth dynamics that are not accessible to continuum approaches. Moreover, the PCF provides a quantitative measure of agreement between an ABM and its corresponding continuum limit. Given experimentally observed parameters, this tool can be used to assess whether a continuum PDE description is appropriate, or whether a finer agent-based framework is required to accurately capture cell dynamics.

Continuum PDEs and discrete ABMs are both common approaches for studying cell invasion. We have shown that under parameter regimes in which the continuum PDE description is valid, the speed of cell invasion under an ABM matches established theory. Under parameter regimes for which the PDE description becomes invalid, the speed of invasion of our ABM wave fronts exceeds the theoretical minimum predicted by the corresponding PDE. We have also shown that the shape of the asymptotic wave-front solution is dependent on the cell proliferation model we choose. Cell proliferation models which confer a proliferative advantage, i.e., the Remain and

Myopic models, form steeper wave fronts than models without such an advantage (nonmyopic Reset/Exponential).

Further work could extend the results presented in this paper in several directions. One natural avenue is to fit average cell-cycle times in our cell-proliferation models to experimental data from density-dependent growing cell populations, and to quantify how different modeling assumptions influence the inferred cell-cycle times. Throughout this paper, we have shown that volume exclusion plays a crucial role in driving differences between cell-cycle models, and that neglecting density-dependent effects is therefore likely to impact estimates of cell-cycle duration.

Direct calibration of parameters within our full modeling framework is challenging due to the complexity of the resulting PDE systems. In contrast, parameter fitting is easier when we neglect volume exclusion. Given a specified underlying cell-cycle model, a potential area of future work might be to determine how to infer the true cell-cycle time of a cell when experimental data are fitted using models that assume no density dependence. There is also capacity to improve the models we have presented to more accurately reflect the true biology of a cell. Typically, cells arrest prior to entering S-phase [36]. However, under the Remain model, cells arrest in the final stage of the cell-cycle model. This means that cells which temporarily arrest within the Remain model and go on to divide will do so more quickly than a true cell. Instead, we could incorporate a checkpoint at an intermediate stage, from which cells may only progress if there is available space. We have also made the basic assumption that cell movement is diffusive. However, there are many examples of cells which demonstrate persistence in their movement and so are not exhibiting classical diffusion on short time scales [50–52]. There are also cells which move nonrandomly in response to environmental cues [53,54]. These nondiffusive movement modes may have an impact on the PDE agreement but also allow behaviors such as collective cell migration, cell aggregation, and anisotropic spreading, which cannot be captured by the assumption of purely diffusive movement, to be modeled. It is also possible to extend our modeling framework to incorporate cell-cell adhesion or cell swapping [55]. Including these extensions to our model could provide a highly realistic representation of both cell proliferation and movement.

In this work, we have demonstrated that different cell-cycle modeling assumptions can lead to markedly different behavior at the population level. Within the ABM framework, we showed that the Remain and Myopic models produce distinct growth dynamics, and that multistage models can, under certain density-dependent conditions, exhibit faster population expansion than simpler exponential models. These findings may have important experimental implications, particularly when inferring cell-cycle time distributions from data collected in crowded environments. While the continuum framework performed well when the underlying mean-field assumptions are satisfied, we showed that parameter choices in the ABMs strongly influence the cellular spatial structure, which in turn affects the agreement between discrete and continuum descriptions. Our results can be used to determine whether continuum or individual-based approaches are appropriate for modeling cell populations.

ACKNOWLEDGMENTS

J.C.D. is supported by a scholarship from the EPSRC Centre for Doctoral Training in Statistical Applied Mathematics at Bath (SAMBa), under the project EP/Y034716/1. C.A.Y. and C.A.S. are supported by a Leverhulme Grant Ref. BLS7239 11342.

DATA AVAILABILITY

The code that generates the data within this paper is publicly available from Ref. [56], embargo periods may apply.

APPENDIX A: DERIVING THE CONTINUUM MODELS FOR MYOPIC ABMS

In this appendix we give more details on the derivations of the continuum equations for the Myopic Reset (13) and Myopic Remain (17) models. In the master equations (10) and (15), we can separate out the contributions due to cell movement, cell stage progression, and cell proliferation. As the contributions arising from cell movement and cell-cycle progression are independent of the proliferative mechanism, they are identical to those in the nonmyopic models. Therefore, we need only consider the contribution due to proliferation. Let $p(t + \tau) - p(t)$ denote the contribution to the change in cell density at lattice site (i, j) over a small time step τ due solely to proliferation events under a Myopic model. Then $p(t + \tau) - p(t)$ takes the form

$$\begin{aligned} p(t + \tau) - p(t) = & \lambda\tau (C_{i-1,j}^{(K)} P((i, j) | (i-1, j)) \\ & + C_{i+1,j}^{(K)} P((i, j) | (i+1, j)) \\ & + \lambda\tau (C_{i,j-1}^{(K)} P((i, j) | (i, j-1)) \\ & + C_{i,j+1}^{(K)} P((i, j) | (i, j+1))), \end{aligned} \quad (\text{A1})$$

where $P((i, j) | (i', j'))$ denotes the probability that a myopic cell at site (i', j') chooses to attempt to proliferate into site (i, j) . We will show that the proliferative term evolves according to the following ODE:

$$\frac{dp}{dt} = \lambda C^{(K)} (1 - C^4), \quad (\text{A2})$$

under the assumption that as $\tau, \Delta \rightarrow 0$, we hold λ constant.

If we expand the terms $C_{i\pm 1,j}, C_{i,j\pm 1}$ around site (i, j) in Eq. (A1), then we obtain the following equation:

$$\begin{aligned} p(t + \tau) - p(t) = & \tau \lambda C_{i,j}^{(K)} \sum_{i',j'} P((i, j) | (i', j')) \\ & + \tau O(\Delta), \end{aligned} \quad (\text{A3})$$

where $\sum_{i',j'} P((i, j) | (i', j'))$ is the sum over all possible neighboring sites so we have that $(i', j') \in \{(i \pm 1, j), (i, j \pm 1)\}$. We will now simplify Eq. (A3), by expanding the terms $P((i, j) | (i', j'))$, and finding a simpler expression for the sum of the probabilities.

The probabilities $P((i, j) | (i \pm 1, j)), P((i, j) | (i, j \pm 1))$ in Eq. (A1) can be written

$$\begin{aligned} P((i, j) | (i \pm 1, j)) &= (1 - C_{i,j}) \times \sum_{l=0}^3 \mathcal{N}_l(i \pm 1, j), \\ P((i, j) | (i, j \pm 1)) &= (1 - C_{i,j}) \times \sum_{l=0}^3 \mathcal{N}_l(i, j \pm 1), \end{aligned} \quad (\text{A4})$$

where each $\mathcal{N}_l(i', j')$ represents the probability that a cell at site (i', j') with l occupied neighbors chooses to proliferate into site (i, j) . For the horizontal neighbors of site (i, j) , recall that we can write $\mathcal{N}_l = \mathcal{N}_l(i \pm 1, j)$ as

$$\begin{aligned} \mathcal{N}_0 &= \frac{1}{4} [(1 - C_{i\pm 2,j})(1 - C_{i\pm 1,j-1})(1 - C_{i\pm 1,j+1})] \\ \mathcal{N}_1 &= \frac{1}{3} C_{i\pm 2,j} (1 - C_{i\pm 1,j+1})(1 - C_{i\pm 1,j-1}) \\ &\quad + \frac{1}{3} C_{i\pm 1,j-1} (1 - C_{i\pm 2,j})(1 - C_{i\pm 1,j+1}) \\ &\quad + \frac{1}{3} C_{i\pm 1,j+1} (1 - C_{i\pm 2,j})(1 - C_{i\pm 1,j-1}) \\ \mathcal{N}_2 &= \frac{1}{2} C_{i\pm 2,j} C_{i\pm 1,j-1} (1 - C_{i\pm 1,j+1}) \\ &\quad + \frac{1}{2} C_{i\pm 2,j} C_{i\pm 1,j+1} (1 - C_{i\pm 1,j-1}) \\ &\quad + \frac{1}{2} C_{i\pm 1,j+1} C_{i\pm 1,j-1} (1 - C_{i\pm 2,j}) \\ \mathcal{N}_3 &= C_{i\pm 2,j} C_{i\pm 1,j-1} C_{i\pm 1,j+1}, \end{aligned} \quad (\text{A5})$$

and we have similar forms for the terms $\mathcal{N}_l(i, j \pm 1)$. We now give the Taylor expansion of the terms $\mathcal{N}_l$ for each $l = 0, 1, 2, 3$ about site (i, j) . This gives the following expansions:

$$\begin{aligned} \mathcal{N}_0 &= \frac{1}{4} (1 - C_{i,j})^3 + O(\Delta) \\ \mathcal{N}_1 &= C_{i,j} (1 - C_{i,j})^2 + O(\Delta) \\ \mathcal{N}_2 &= \frac{3}{2} C_{i,j}^2 (1 - C_{i,j}) + O(\Delta) \\ \mathcal{N}_3 &= C_{i,j}^3 + O(\Delta). \end{aligned} \quad (\text{A6})$$

Using Eq. (A4), this gives explicit forms for $P((i, j) | (i \pm 1, j))$ and $P((i, j) | (i, j \pm 1))$. Summing these four probabilities gives the following equation:

$$\begin{aligned} \sum_{i',j'} P((i, j) | (i', j')) &= (1 - C_{i,j})^4 + 4C_{i,j} (1 - C_{i,j})^3 \\ &\quad + 6C_{i,j}^2 (1 - C_{i,j})^2 + 4C_{i,j}^3 (1 - C_{i,j}), \\ &= (1 - C_{i,j}^4). \end{aligned} \quad (\text{A7})$$

Dividing Eq. (A1) by τ gives the following [when expanding to $O(\Delta)$]:

$$\begin{aligned} \frac{p(t + \tau) - p(t)}{\tau} &= \lambda (C_{i-1,j}^{(K)} P((i, j) | (i-1, j)) \\ &\quad + C_{i+1,j}^{(K)} P((i, j) | (i+1, j)) \\ &\quad + \lambda (C_{i,j-1}^{(K)} P((i, j) | (i, j-1)) \\ &\quad + C_{i,j+1}^{(K)} P((i, j) | (i, j+1))) \\ &= \lambda C_{i,j}^{(K)} \sum_{i',j'} P((i, j) | (i', j')) + O(\Delta) \\ &= \lambda C_{i,j}^{(K)} (1 - C_{i,j}^4) + O(\Delta), \end{aligned} \quad (\text{A8})$$

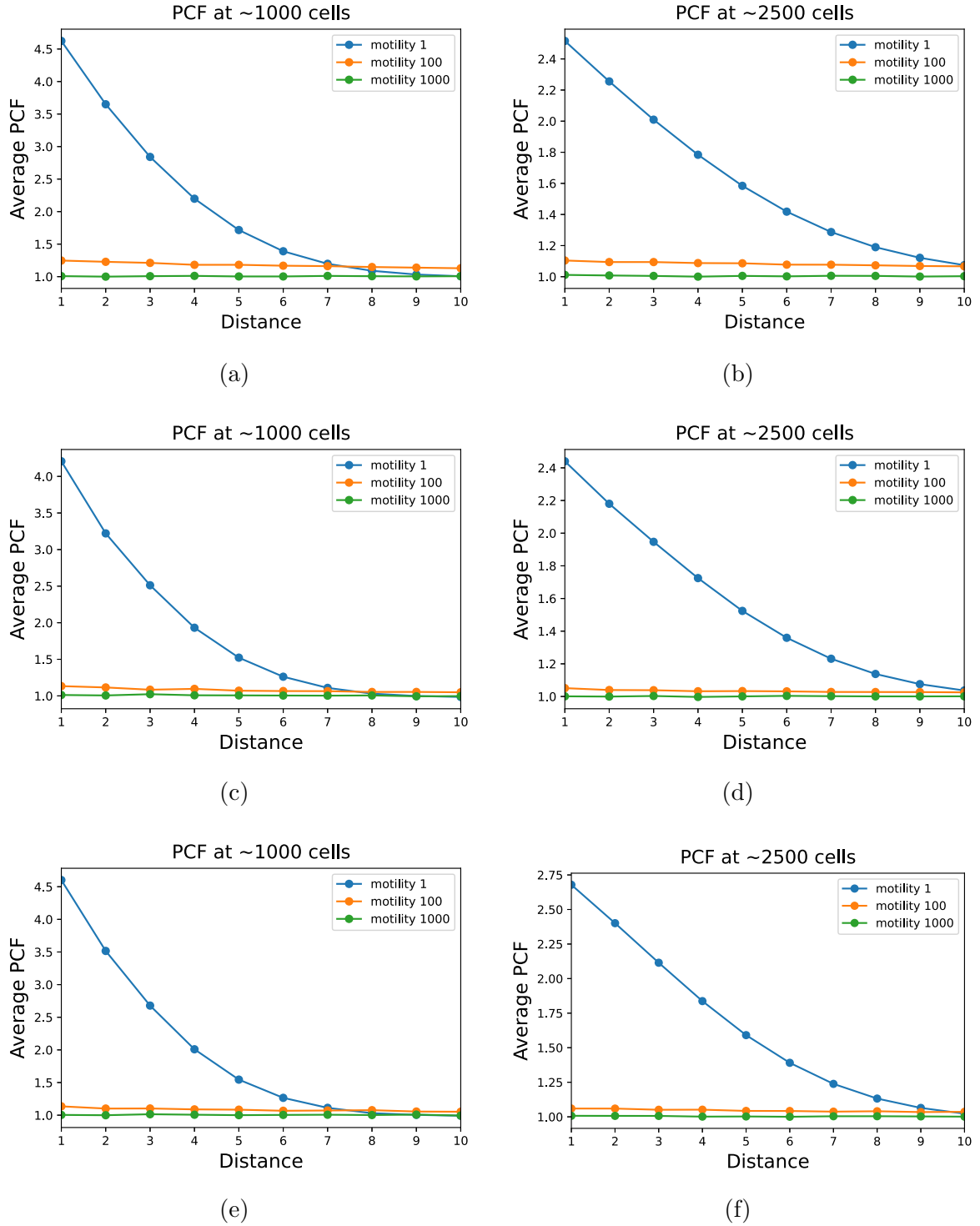


FIG. 19. Plotting the PCFs $\hat{f}(d)$ over distances $d = 1, 2, \dots, 10$ for nonmyopic models averaged over 20 realizations of different ABMs for motility rates $r_m = 1, 100, 1000$ with a fixed proliferation rate $r_p = 1$ and number of stages $K = 10$. For each model, in each realization, we record the time point where the ABM first exceeds 1000 and 2500 cells, and use this to calculate the average PCF. Panels (a) and (b) show the results of the Exponential model, panels (c) and (d) show the results of the Reset model, and panels (e) and (f) show the results of the Remain model.

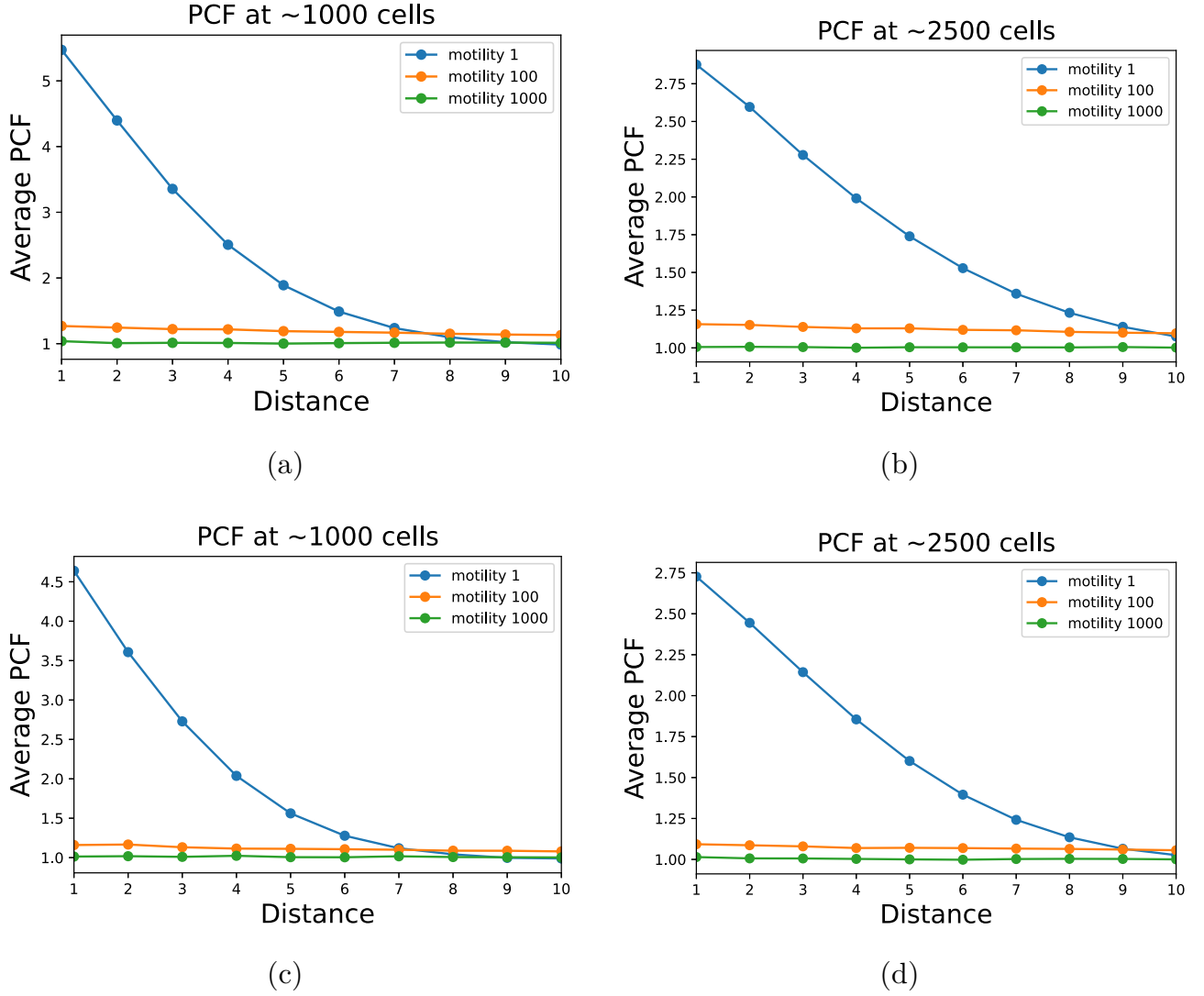


FIG. 20. Plotting average PCFs for Myopic models, as in Fig. 19. Panels (a) and (b) show the results of the Myopic Exponential model; panels (c) and (d) show the results of the Myopic Reset model.

where we used Eqs. (A3) and (A7) to simplify the resulting expressions.

If we take the limit of Eq. (A8) as $\tau, \Delta \rightarrow 0$, then we obtain the following ODE:

$$\frac{dp}{dt} = \lambda C^{(K)}(1 - C^4).$$

This is precisely the ODE (A2).

Recall that the other contributions to the PDE for stage 1 cells under the Myopic Reset and Myopic Remain models from the master Eqs. [(10) and (15)] involve cell movement and cell progression terms which are not impeded by the proliferative mechanism and so are the same as those of the nonmyopic models. This allows us to use the diffusive and proliferative terms in PDEs (3) and (8) to obtain the following PDE for stage 1 cells under the Myopic Reset model:

$$\begin{aligned} \frac{\partial C^{(1)}}{\partial t} = & D[(1 - C)\nabla^2 C^{(1)} + C^{(1)}\nabla^2 C] \\ & + \lambda C^{(K)}(2 - C^4) - \lambda C^{(1)}. \end{aligned}$$

Similarly, for the Myopic Remain model, adapting Eq. (8) we have the resultant PDE

$$\begin{aligned} \frac{\partial C^{(1)}}{\partial t} = & D[(1 - C)\nabla^2 C^{(1)} + C^{(1)}\nabla^2 C] \\ & + 2\lambda C^{(K)}(1 - C^4) - \lambda C^{(1)}. \end{aligned}$$

APPENDIX B: PCF PLOTS FOR ABMS OTHER THAN MYOPIC REMAIN

In Figs. 19 and 20, we plot the pair correlation functions (PCFs) for ABMs other than the Myopic Remain model. Recall that given an $(X \times Y)$ lattice, we can calculate the PCF at distance d in repeat m using the following equation:

$$f^m(d) = \frac{(XY - 1)|C^M(d)|}{2dN(N - 1)}, \quad (\text{B1})$$

where N is the number of cells in the lattice, and $C^M(d)$ is as written in Eq. (22) so that $|C^M(d)|$ is the number of pairs of cells separated by distance d . This definition is still valid

for our different ABMs, as each simulation uses the same boundary conditions (periodic). We considered the average PCF $\hat{f}(d)$ over multiple realizations of each ABM.

As in the Myopic Remain model, we see that for motility rate $r_m = 1$, each ABM demonstrates strong pairwise correlations for small distances d . At motility rate $r_m = 100$, we see the strength of correlation decreases, but is still present

in every model as in the Myopic Remain model. When we increase the motility rate to $r_m = 1000$, this effectively removes the pairwise correlations in every model, as $\hat{f}(d) \approx 1$ for all distances d . There are slight differences in the value $|\hat{f}(d) - 1|$, in each model, but it is clear that the strength of correlations between lattice sites is approximately the same under every ABM.

- [1] B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. Watson, *Molecular Biology of the Cell*, 4th ed. (Garland, New York, 2002).
- [2] G. Gurtner, S. Werner, Y. Barrandon, and M. Longaker, Wound repair and regeneration, *Nature (London)* **453**, 314 (2008).
- [3] D. Hanahan and R. Weinberg, Hallmarks of cancer: The next generation, *Cell* **144**, 646 (2011).
- [4] K. Schafer, The cell cycle: A review, *Vet. Pathol.* **35**, 461 (1998).
- [5] J. Pietenpol and Z. Stewart, Cell cycle checkpoint signaling: Cell cycle arrest versus apoptosis, *Toxicology* **181-182**, 475 (2002).
- [6] D. Johnson and C. Walker, Cyclins and cell cycle checkpoints, *Annu. Rev. Pharmacol. Toxicol.* **39**, 295 (1999).
- [7] A. Meerson, M. Milyavsky, and V. Rotter, p53 mediates density-dependent growth arrest, *FEBS Lett.* **559**, 152 (2004).
- [8] O. Leontieva, Z. Demidenko, and M. Blagosklonny, Contact inhibition and high cell density deactivate the mammalian target of rapamycin pathway, thus suppressing the senescence program, *Proc. Natl. Acad. Sci. USA* **111**, 8832 (2014).
- [9] C. Glen, M. Kemp, and E. Voit, Agent-based modeling of morphogenetic systems: Advantages and challenges, *PLoS Comput. Biol.* **15**, e1006577 (2019).
- [10] G. An, Q. Mi, J. Dutta-Moscato, and Y. Vodovotz, Agent-based models in translational systems biology, *WIREs Mech. Dis.* **1**, 159 (2009).
- [11] M. Baar, L. Coquille, H. Mayer, M. Hölzel, M. Rogava, T. Tüting, and A. Bovier, A stochastic model for immunotherapy of cancer, *Sci. Rep.* **6**, 24169 (2016).
- [12] R. Mort, R. Ross, K. Hainey, O. Harrison, M. Keighren, G. Landini, R. Baker, K. Painter, I. Jackson, and C. Yates, Reconciling diverse mammalian pigmentation patterns with a fundamental mathematical model, *Nat. Commun.* **7**, 10288 (2016).
- [13] C. Turner, A. Stinchcombe, M. Kohandel, S. Singh, and S. Sivaloganathan, Characterization of brain cancer stem cells: A mathematical approach, *Cell Prolif.* **42**, 529 (2009).
- [14] D. Gillespie, Exact stochastic simulation of coupled chemical reactions, *J. Phys. Chem.* **81**, 2340 (1977).
- [15] A. Golubev, Applications and implications of the exponentially modified gamma distribution as a model for time variabilities related to cell proliferation and gene expression, *J. Theor. Biol.* **393**, 203 (2016).
- [16] C. Yates, M. Ford, and R. Mort, A multi-stage representation of cell proliferation as a Markov process, *Bull. Math. Biol.* **79**, 2905 (2017).
- [17] H. X. Chao, R. I. Fakhreddin, H. K. Shimerov, K. M. Kedziora, R. J. Kumar, J. Perez, J. C. Limas, G. D. Grant, J. G. Cook, G. P. Gupta, and J. E. Purvis, Evidence that the human cell cycle is a series of uncoupled, memoryless phases, *Mol. Syst. Biol.* **15**, MSB188604 (2019).
- [18] J. Wang and P. Levin, Metabolism, cell growth and the bacterial cell cycle, *Nat. Rev. Microbiol.* **7**, 822 (2009).
- [19] R. Baker and M. Simpson, Correcting mean-field approximations for birth-death-movement processes, *Phys. Rev. E* **82**, 041905 (2010).
- [20] M. Simpson, K. Landman, and B. Hughes, Multi-species simple exclusion processes, *Physica A* **388**, 399 (2009).
- [21] M. Simpson, K. Landman, and B. Hughes, Cell invasion with proliferation mechanisms motivated by time-lapse data, *Physica A* **389**, 3779 (2010).
- [22] M. Simpson, W. Jin, S. T. Vittadello, T. Tambyah, J. Ryan, G. Gunasingh, N. Haass, and S. McCue, Stochastic models of cell invasion with fluorescent cell cycle indicators, *Physica A* **510**, 375 (2018).
- [23] E. Gavagnin, M. Ford, R. Mort, T. Rogers, and C. Yates, The invasion speed of cell migration models with realistic cell cycle time distributions, *J. Theor. Biol.* **481**, 91 (2019).
- [24] K. Landman and A. Fernando, Myopic random walkers and exclusion processes: Single and multispecies, *Physica A* **390**, 3742 (2011).
- [25] S. Vittadello, S. McCue, G. Gunasingh, N. Haass, and M. Simpson, Mathematical models incorporating a multi-stage cell cycle replicate normally-hidden inherent synchronization in cell proliferation, *J. R. Soc. Interface* **16**, 20190382 (2019).
- [26] E. Gavagnin, S. T. Vittadello, G. Gunasingh, N. K. Haass, M. J. Simpson, T. Rogers, and C. A. Yates, Synchronized oscillations in growing cell populations are explained by demographic noise, *Biophys. J.* **120**, 1314 (2021).
- [27] S. Vittadello, S. McCue, G. Gunasingh, N. Haass, and M. Simpson, A novel mathematical model of heterogeneous cell proliferation, *J. Math. Biol.* **82**, 34 (2021).
- [28] J. Murray, *Mathematical Biology: I. An Introduction*, Interdisciplinary Applied Mathematics, Vol. 17 (Springer Science & Business Media, 2007).
- [29] S. Vittadello, S. McCue, G. Gunasingh, N. Haass, and M. Simpson, Mathematical models for cell migration with real-time cell cycle dynamics, *Biophys. J.* **114**, 1241 (2018).
- [30] C. Falcó, D. J. Cohen, J. A. Carrillo, and R. E. Baker, Quantifying cell cycle regulation by tissue crowding, *Biophys. J.* **124**, 923 (2025).
- [31] J. Wang, T. Broer, T. Chavez, C. Zhou, S. Tran, Y. Xiang, A. Khodabukus, Y. Diao, and N. Bursac, Myoblast deactivation within engineered human skeletal muscle creates a transcriptionally heterogeneous population of quiescent satellite-like cells, *Biomaterials* **284**, 121508 (2022).
- [32] G. Yao, Modelling mammalian cellular quiescence, *Interface Focus* **4**, 20130074 (2014).
- [33] M. Simpson, B. Hughes, and K. Landman, Diffusing populations: Ghosts or folks? *Austral. J. Eng. Ed.* **15**, 59 (2009).

- [34] D. Kendall, On the role of variable generation time in the development of a stochastic birth process, *Biometrika* **35**, 316 (1948).
- [35] S. Streichan, C. Hoerner, T. Schneidt, D. Holzer, and L. Hufnagel, Spatial constraints control cell proliferation in tissues, *Proc. Natl. Acad. Sci. USA* **111**, 5586 (2014).
- [36] L. Donker, R. Houtekamer, M. Vliem, F. Sipieter, H. Canever, M. Gómez-González, M. Bosch-Padrós, W.-J. Pannekoek, X. Trepát, N. Borghi, and M. Gloerich, A mechanical G2 checkpoint controls epithelial cell division through E-cadherin-mediated regulation of Wee1-Cdk1, *Cell Rep.* **41**, 111475 (2022).
- [37] M. Uroz, S. Wistorf, X. Serra-Picamal, V. Conte, M. Sales-Pardo, P. Roca-Cusachs, R. Guimerà, and X. Trepát, Regulation of cell cycle progression by cell–cell and cell–matrix forces, *Nat. Cell Biol.* **20**, 646 (2018).
- [38] K. Höllring, L. Nuić, L. Rogić, S. Kaliman, S. Gehrler, C. Wollnik, F. Rehfeldt, M. Hubert, and A.-S. Smith, Capturing the mechanosensitivity of cell proliferation in models of epithelium, *Proc. Natl. Acad. Sci. USA* **121**, e2308126121 (2024).
- [39] P. Kulesa, C. Bailey, J. Kasemeier-Kulesa, and R. McLennan, Cranial neural crest migration: New rules for an old road, *Dev. Biol.* **344**, 543 (2010).
- [40] P. Maini, D. McElwain, and D. Leavesley, Travelling waves in a wound healing assay, *Appl. Math. Lett.* **17**, 575 (2004).
- [41] G. C. Paul, Tauhida, and D. Kumar, Revisiting Fisher-KPP model to interpret the spatial spreading of invasive cell population in biology, *Heliyon* **8**, e10773 (2022).
- [42] D. Markham, M. Simpson, P. Maini, E. Gaffney, and R. Baker, Comparing methods for modelling spreading cell fronts, *J. Theor. Biol.* **353**, 95 (2014).
- [43] A. Cai, K. Landman, and B. Hughes, Multi-scale modeling of a wound-healing cell migration assay, *J. Theor. Biol.* **245**, 576 (2007).
- [44] A. Parker, M. Simpson, and R. Baker, The impact of experimental design choices on parameter inference for models of growing cell colonies, *R. Soc. Open Sci.* **5**, 180384 (2018).
- [45] E. Gavagnin, J. Owen, and C. Yates, Pair correlation functions for identifying spatial correlation in discrete domains, *Phys. Rev. E* **97**, 062104 (2018).
- [46] A. Sakaue-Sawano, H. Kurokawa, T. Morimura, A. Hanyu, H. Hama, H. Osawa, S. Kashiwagi, K. Fukami, T. Miyata, H. Miyoshi, T. Imamura, M. Ogawa, H. Masai, and A. Miyawaki, Visualizing spatiotemporal dynamics of multicellular cell-cycle progression, *Cell* **132**, 487 (2008).
- [47] W. Van Saarloos, Front propagation into unstable states, *Phys. Rep.* **386**, 29 (2003).
- [48] J. Harrison and C. Yates, A hybrid algorithm for coupling partial differential equation and compartment-based dynamics, *J. R. Soc. Interface.* **13**, 20160335 (2016).
- [49] E. Brunet and B. Derrida, Shift in the velocity of a front due to a cutoff, *Phys. Rev. E* **56**, 2597 (1997).
- [50] D. Selmeczi, S. Mosler, P. Hagedorn, N. Larsen, and H. Flyvbjerg, Cell motility as persistent random motion: Theories from experiments, *Biophys. J.* **89**, 912 (2005).
- [51] P. Wu, A. Giri, S. Sun, and D. Wirtz, Three-dimensional cell migration does not follow a random walk, *Proc. Natl. Acad. Sci. USA* **111**, 3949 (2014).
- [52] J. C. J. Heyn, J. O. Rädler, and M. Falcke, Mesenchymal cell migration on one-dimensional micropatterns, *Front. Cell Dev. Biol.* **12**, 1352279 (2024).
- [53] Y. Artemenko, T. Lampert, and P. Devreotes, Moving towards a paradigm: Common mechanisms of chemotactic signaling in Dictyostelium and mammalian leukocytes, *Cell. Mol. Life Sci.* **71**, 3711 (2014).
- [54] J. Espina, C. Marchant, and E. Barriga, Durotaxis: The mechanical control of directed cell migration, *FEBS J.* **289**, 2736 (2022).
- [55] S. Noureen, R. Mort, and C. Yates, Modeling adhesion in stochastic and mean-field models of cell migration, *Phys. Rev. E* **111**, 014419 (2025).
- [56] J. C. Dimaculangan, jcd65/msm-volume-exclusion-cell-models: msm-volume-exclusion-cell-models_v1.0.0 (Version final) [Computer software], Zenodo, 2026, <https://doi.org/10.5281/zenodo.23099835>.