

Simulating invasion with cellular automata: Connecting cell-scale and population-scale properties

Matthew J. Simpson,^{*} Alistair Merrifield, Kerry A. Landman, and Barry D. Hughes
Department of Mathematics and Statistics, The University of Melbourne, Victoria 3010, Australia
 (Received 30 March 2007; revised manuscript received 20 June 2007; published 17 August 2007)

Interpretive and predictive tools are needed to assist in the understanding of cell invasion processes. Cell invasion involves cell motility and proliferation, and is central to many biological processes including developmental morphogenesis and tumor invasion. Experimental data can be collected across a wide range of scales, from the population scale to the individual cell scale. Standard continuum or discrete models used in isolation are insufficient to capture this wide range of data. We develop a discrete cellular automata model of invasion with experimentally motivated rules. The cellular automata algorithm is applied to a narrow two-dimensional lattice and simulations reveal the formation of invasion waves moving with constant speed. The simulation results are averaged in one dimension—these data are used to identify the time history of the leading edge to characterize the population-scale wave speed. This allows the relationship between the population-scale wave speed and the cell-scale parameters to be determined. This relationship is analogous to well-known continuum results for Fisher's equation. The cellular automata algorithm also produces individual cell trajectories within the invasion wave that are analogous to cell trajectories obtained with new experimental techniques. Our approach allows both the cell-scale and population-scale properties of invasion to be predicted in a way that is consistent with multiscale experimental data. Furthermore we suggest that the cellular automata algorithm can be used in conjunction with individual data to overcome limitations associated with identifying cell motility mechanisms using continuum models alone.

DOI: [10.1103/PhysRevE.76.021918](https://doi.org/10.1103/PhysRevE.76.021918)

PACS number(s): 87.17.Jj, 87.15.Vv, 87.18.Bb

I. INTRODUCTION

Cell migration models have been typically constructed using either continuum [1–5] or discrete [1,6–13] approaches. Continuum models describe population-scale properties while discrete models describe cell-scale properties. We will refer to a population-scale property as a property of a sufficiently large group of cells. For example, the speed at which a population of cells invades into unoccupied regions is a population-scale property. We will refer to a cell-scale property as a property attributed to a single cell. An example of a cell-scale property is the trajectory of an individual cell within a population of motile cells.

While some experimental observations of cell motility and cell invasion are made exclusively at the population scale [5], recent advances in microscopy such as the use of confocal microscopy, time-lapse imaging, and magnetic resonance imaging, mean that biological observations and measurements can be made over a range of scales [14] including individual cell-scale behavior [15–20]. Predictive tools relevant to these kinds of multiscale observations cannot be based on continuum or discrete models alone. What is required is an ability to simulate processes occurring over a range of scales [14]. One way to achieve this is to understand how individual interactions in a discrete model give rise to emergent population-scale patterns [21].

Cell invasion is ubiquitous in many areas of cell biology, including epidermal wound healing [2], tumor invasion [22], and key events during developmental morphogenesis [4,5]. Cell invasion also occurs *in vitro* during a cell migration wound healing assay [7] and in tissue engineering applica-

tions [23]. Cell invasion involves a population of cells moving into previously unoccupied substrates or tissues and proliferating to occupy these regions at a capacity density. This requires two key cell functions: (i) cells are motile, and move due to some mechanism such as an unbiased random walk, and (ii) cells proliferate according to some carrying capacity-limited proliferation mechanism, such as logistic proliferation. Carrying capacity-limited proliferation is sometimes called contact-inhibited proliferation. The carrying capacity density reflects cell crowding effects—cells proliferate when the cell density is below carrying capacity but proliferation ceases when the cell density reaches capacity density. The domain on which the cells move and proliferate may itself grow, for example, if the domain is a developing embryonic tissue [24]. Carrying capacity-limited proliferation mechanisms can also be used to simulate proliferation in this growing case. Here we focus on nongrowing domains. Continuum models of cell invasion in nongrowing domains typically support traveling wave solutions characterized by constant shape, constant speed invasion waves [3].

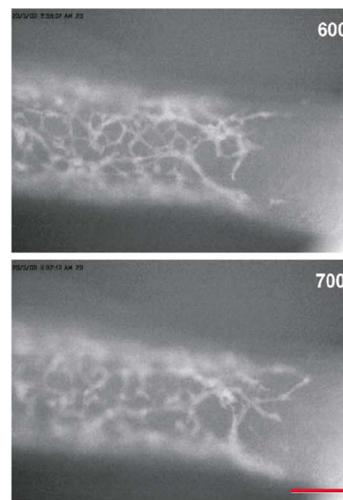
During the last decade there has been increasing interest in the use of discrete models for simulating cell motility in a variety of contexts. Discrete models are formulated using a number of approaches including the cellular Potts model [10,12,13,25–29], individual or agent-based cellular automata (CA) [6,9,11,21,30–33], as well as transition probability rules that are related to, or extracted from, discretized partial differential equations [1,8,34–36]. These approaches have successfully replicated various cell motility phenomena and some studies have included carrying capacity-limited proliferation [7,37,38]. Less progress has been made toward formulating and understanding discrete models of traveling wave-type invasive systems and relating such discrete systems to their continuum analogs or relevant experimental data.

^{*}m.simpson@ms.unimelb.edu.au

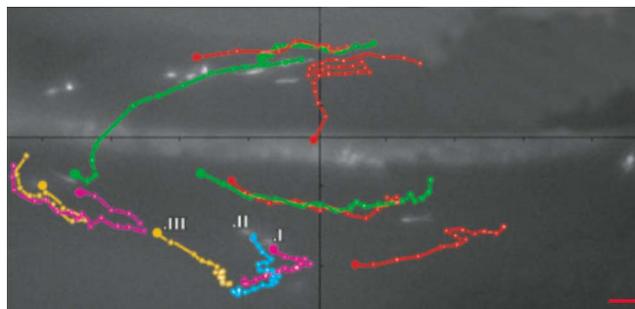
Our aim is to design, implement, and analyze a discrete tool capable of replicating invasion phenomena. We will also show how the discrete model can be used to overcome certain limitations associated with continuum invasion models, and thereby illustrate how the discrete tool can be used to contribute to biological knowledge. An archetypal example of cell invasion occurs during the formation of the enteric nervous system (ENS) during vertebrate embryogenesis. The ENS formation is particularly attractive because of the large amount of experimental data that has been collected for this system. ENS development involves neural crest (NC) cells invading the developing gut tissue as a unidirectional invasion wave moving in a rostral-to-caudal (stomach-to-anus) direction [17,18,20]. NC cells are motile and proliferate to a carrying capacity density [39]. The cells differentiate into neurons and glia to form the ENS which gives rise to normal gut function and peristaltic contraction. Failure of the invasion process results in Hirschsprung's Disease, a relatively common and potentially fatal human birth defect, where the terminal intestine fails to produce peristalsis due to the absence of the ENS.

An important aspect of the approach taken in this work is that experimentally motivated cell-based rules are identified and implemented in the CA algorithm. Although it is possible to formulate individual transition rules from a discretized partial differential equation [1,7,8,34–36], this approach may not always be feasible, especially in situations when it is unclear what kind of partial differential equation is relevant. We anticipate that, with the increasing use of time-lapse imaging, analysts must become more reliant on interpreting experimental data to formulate discrete rules. These rules may not correspond to any known partial differential equation. Therefore it is of interest to develop experimentally motivated discrete rules and to investigate whether or not the emergent population-scale behavior corresponds to any known continuum results.

In this paper we present a CA algorithm which replicates traveling wavelike invasion profiles. Two stochastic cell-scale rules are introduced: (i) cell motility, governed by a probability of motility P_m and a carrying capacity density κ ; and (ii) carrying capacity-limited proliferation, governed by a probability of proliferation P_p and a carrying capacity density κ . The speed of the invasion wave, denoted by c , is measured and related to the CA parameters. The cell-scale parameters are also related to equivalent population-scale parameters relevant at the continuum scale. In particular, we identify estimates for a diffusivity D and mitotic index λ . The diffusivity characterizes the cell motility and the mitotic index characterizes the proliferation rate which is related to the doubling time of cells [4]. CA simulations are performed on a square lattice with unit lattice spacings and time is discretized into unit increments. For an arbitrary lattice with spacing Δx and time steps Δt , the appropriate dimensional diffusivity and mitotic index are given by $D\Delta x^2/\Delta t$ and $\lambda/\Delta t$, where D and λ are the quantities established for the unit lattice system. This allows us to connect the individual-scale CA parameters with population-scale observations through an expression for the CA wave speed that is similar to established expressions for continuum models [40]. Observations of individual cell-scale properties are made by



(a)



(b)

FIG. 1. (Color online) A suite of multiscale experimental observations of NC cell invasion. Time-lapse imaging in (a) shows the population-scale rostral to caudal (left to right) progression of the NC cell population in a mouse embryo. Two panels are shown with time increasing down the page, the wave speed is approximately $35 \mu\text{m}/\text{h}$. Scale $100 \mu\text{m}$. (Reprinted from Ref. [20], with permission.) (b) Individual NC cell trajectories at different locations within the invasion wave of a chick embryo are traced and shown. Scale $20 \mu\text{m}$. (Reprinted from Ref. [18], with permission.)

tracking cell trajectories within the wave. Various properties of these trajectories are statistically quantified.

Recent experimental results associated with NC cell invasion of the developing intestine are shown in Fig. 1. The population-scale data [Fig. 1(a)] show a large population of cells invading from left to right in an embryonic mouse gut. Detailed cell-scale data [Fig. 1(b)] show individual cell trajectories located in different positions along the invasion wave in an embryonic chick gut. The CA algorithm will be used to replicate both the cell-scale and population-scale experimental observations.

II. CONTINUUM INVASION MODEL

The archetypal continuum invasion model is Fisher's equation [3,40]. This model assumes cells are diffusively motile with diffusivity D and proliferate logistically with mi-

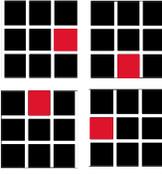
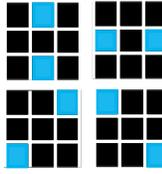
CA RULES	MOTILITY RULE	PROLIFERATION RULE
Current configuration	If $R > P_m$ no movement 	If $R > P_p$ no proliferation 
Potential new configuration	If $R \leq P_m$ & $M \leq \kappa$ select one of four new locations 	If $R \leq P_p$ & $M \leq \kappa$ select one of four new locations 

FIG. 2. (Color online) Diagram showing the two CA rules. Empty lattice sites are black while cells are either red (dark gray) or blue (light gray). If a cell is motile it can transition to any of the nearest four neighboring sites with equal probability provided that the site is empty and that the local density around the cell in the new location is not above carrying capacity. If an uncrowded cell with $M \leq \kappa$ is proliferative, the daughter cells move into one of four configurations so that the two daughter cells move in opposing directions. Proliferation is terminated if either of the selected sites is occupied.

otic index λ to reach a carrying capacity density K . Fisher's equation, with a generalized logistic source term, can be written as

$$\frac{\partial u}{\partial t} = D \frac{\partial^2 u}{\partial x^2} + \lambda u \left(1 - \left[\frac{u}{K} \right]^m \right), \quad (1)$$

where $u(x, t)$ is the cell density, x is the spatial coordinate, t is time, and m is a positive parameter ($m=1$ in Fisher's equation). Here we consider space and time to be scaled with Δx and Δt , respectively, so that the continuum parameters D and λ match the discrete CA parameters. Solutions of Eq. (1) on $-\infty < x < \infty$, starting with $u(x, 0)$ having compact support, evolve into invasion waves moving with constant speed c , where

$$c = 2\sqrt{\lambda D}. \quad (2)$$

Other wave speeds with $c > 2\sqrt{\lambda D}$ are possible for different $u(x, 0)$ [3].

The dependent variable $u(x, t)$ in Fisher's equation represents the number of cells per unit length (cell density). This continuum description is valid only for large numbers of cells so that an averaged property of the population, such as cell density, can be defined (see Bear [41] for a comprehensive discussion). Fisher's equation can replicate the population-scale properties of invasion, such as the invasion speed [4,5], as well as cell-scale data by discretizing the equation and extracting cell-scale transition rules [1,7,36].

III. CELLULAR AUTOMATA ALGORITHM

The CA algorithm is developed on a rectangular lattice in the x - y plane. Each lattice site can be occupied by only one cell. We take unit lattice spacing so the coordinates (x, y) of a site are integers. The lattice has the dimensions $L_x \times L_y$ so that $0 \leq x \leq L_x$ and $0 \leq y \leq L_y$ and the site corresponding to $(0, 0)$ is located in the lower left-hand corner of the lattice. Periodic boundary conditions are imposed at $y=0$ and $y=L_y$

so that a cell located at $(x, 0)$ appears at (x, L_y) after moving in the negative y direction, and a cell located at (x, L_y) appears at $(x, 0)$ after moving in the positive y direction. We impose these periodic boundary conditions since NC cells move on the closed surface of cylindrical gut tissues. Reflecting (zero-flux) boundary conditions are imposed at $x=0$ and $x=L_x$ to represent the physical boundary at the longitudinal ends of the substrate or tissue domain. For our initial conditions (discussed in Sec. IV), the formation of an invasion wave is insensitive to the details of the boundary condition at $x=0$. We also choose L_x to be sufficiently large so that the invasion wave does not interact with the boundary at the far end during the time of the simulation.

The relationship between the computational time step and the real time step is governed by the proliferation parameter P_p . We assume, in the absence of any further detailed information, that the proliferative properties of all cells simulated in the CA algorithm are uniform. If the real dimensional time taken for a cell to proliferate (the cell cycle time) is t_p , then $P_p = \Delta t / t_p$ [23], where Δt is the real dimensional step size. Since the proliferation parameter must satisfy $0 \leq P_p \leq 1$, we require $0 \leq \Delta t \leq t_p$. This means that the real time step cannot be larger than the cell cycle time.

At any particular time t , the system contains $N(t)$ cells. During each time step, $N(t)$ random selections of cells are made and these cells are given the opportunity to move according to a motility rule. Other $N(t)$ random selections are made and the selected cells are given the opportunity to proliferate. Once the $2 \times N(t)$ selections have been made, the system has advanced through one time step to time $t+1$. The number of cells $N(t+1)$ is adjusted to reflect birth processes during that time step. This algorithm can allow any particular cell to undergo more than one movement or proliferation in any particular time step. The details of both the motility and proliferation rules, summarized in Fig. 2, will now be discussed.

A. Motility rule

Two parameters are introduced to describe cell motility: P_m and κ . Here, P_m is the probability that a cell will be

motile with cell movements restricted to the four nearest-neighbor lattice sites. We assume that cells are able to detect which of their nearest eight neighbor sites are occupied by other cells. We introduce κ as the (integer) carrying capacity density of the lattice with $1 \leq \kappa \leq 8$.

Cell motility is based on an unbiased nearest-neighbor random walk. To be consistent with experimental observations, cell motility will be modulated by neighbor exclusion and crowding effects. A cell at (x, y) is selected and a random number R is chosen from a uniform distribution on $[0, 1]$. If $R > P_m$, the cell remains at (x, y) . If $R \leq P_m$, the cell can move, with equal probability, to $(x \pm 1, y)$ or $(x, y \pm 1)$. If the selected site is occupied the movement is aborted. To limit overcrowding, the number of neighbors M surrounding the cell in the new location is determined. If $M > \kappa$ the movement is aborted, otherwise the movement is allowed.

B. Proliferation rule

A cell is selected at random. If this cell has already moved during the current time step then proliferation is aborted. This rule is chosen to be consistent with time-lapse data which suggest that NC cells cease moving just prior to mitosis [20]. If a cell is chosen that has not moved during that time step, a random number R is chosen from a uniform distribution on $[0, 1]$. If $R > P_p$ the cell does not proliferate, whereas if $R \leq P_p$ the cell may proliferate depending upon the local cell density. The number of neighbors M occupying the eight neighboring sites surrounding the cell is counted; if $R \leq P_p$ and $M \leq \kappa$ the cell divides into two daughter cells.

Once a proliferative cell is identified, the location of the daughter cells must be specified. Time-lapse imaging suggests that when cells divide both daughter cells move in opposing directions away from the site of the mother cell [20]. If a cell at (x, y) proliferates, the daughter cells are allowed to occupy one of four configurations with equal probability: (i) $(x \pm 1, y)$; (ii) $(x, y \pm 1)$; (iii) $(x - 1, y - 1)$, $(x + 1, y + 1)$; or (iv) $(x - 1, y + 1)$, $(x + 1, y - 1)$. If the chosen configuration contains an occupied site then proliferation is aborted.

C. Crowding, cell removal, and cell-cell interactions

The cell crowding and cell-cell interaction principles described in the motility and proliferation rules have been selected to be consistent with properties thought to be relevant to ENS development. A standard discretization of Fisher’s equation implies that cells will be removed from the system whenever the local density rises above the carrying capacity density [7]. Removal of cells implies that cell death is present [42]. Experimental observations of ENS development reveals no evidence of cell death [43]. Therefore care must be taken to ensure cell removal is not simulated with the discrete rules. We note that the proliferation rule can generate a local vicinity where the local density is greater than κ if $\kappa < 8$, but in this case further proliferation and motility into this localized region is suppressed.

If a carrying capacity density is not incorporated into the motility rule, all simulations lead to invasion waves where

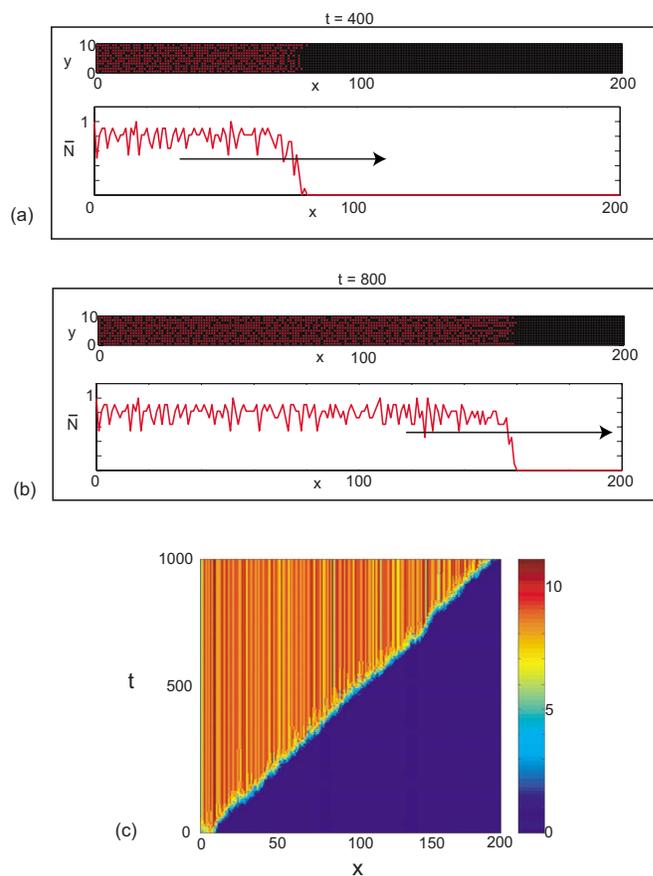


FIG. 3. (Color online) A single realization of the population-scale dynamics of an invasion wave simulated with the CA algorithm for $P_p=0.25$, $P_m=0.15$, and $\kappa=4$. Results in (a) and (b) show two snapshots of the wave at $t=400$ and 800 , respectively. These diagrams show both the position of the individual cells on the lattice and the distribution of \bar{N} , the normalized number of cells per column, plotted against the distance along the longitudinal axis of invasion x . The direction of invasion is indicated by the arrows. A space-time diagram is given in (c) for $0 \leq t \leq 1000$ with the color bar representing the number of cells in each vertical column of the CA lattice.

the invaded region is almost entirely occupied by cells. To circumvent this, the cell motility rule only allows cell movement where the new location of a motile cell is less than or equal to a specified carrying capacity density. This rule allows the CA algorithm to generate invasion waves of varying density without the need for simulating cell removal. The robustness of the CA algorithm to these details will be discussed in Sec. VI.

IV. WAVE SPEED ANALYSIS

The CA rules were realized on a narrow two-dimensional lattice with $L_x=500$ and $L_y=10$ with the net invasion direction in the positive x direction. A wide range of parameters were considered with all parameter combinations resulting in constant speed invasion waves such as the example shown in Fig. 3. Simulations are initiated by randomly seeding the region $0 \leq x \leq 10$ and $0 \leq y \leq 10$ at carrying capacity density.

The CA rules were allowed to evolve over 1500 time steps. This amount of time was chosen because it was sufficient to enable the formation of constant speed invasion waves. If a specific biological example is considered, then the time steps and lattice dimensions must be chosen according to the context. The location of the wave front was determined by counting the number of cells in each vertical grid column. The wave front was defined as the x coordinate of the first column, counting from the right, where the number of cells per column equaled or exceeded H . Results are insensitive to H ; all results presented here correspond to $H=3$. Results are also relatively insensitive to the height of the domain L_y .

The time evolution of the wave front was measured 40 times for each parameter combination to generate a suitable set of wave speed data. The mean wave speed was determined in two ways. First, the wave speed was estimated for each of the 40 realizations, and these estimates were averaged to give the mean wave speed. Second, at each time point, the average position of the leading edge of the 40 simulations was evaluated and these data were used to determine the mean wave speed. Both approaches gave similar results. All results reported here were obtained using the first method. The variability of the wave speed estimates among the 40 simulations was measured using the standard deviation, σ , of the sample. Results are presented as a mean wave speed c only. Values of the ratio σ/c were small, with typical values $<15\%$. This variability is qualitatively consistent with experimental observations of invasion speed data which also show some variability [20].

General trends and quantitative relationships between the mean wave speed c and the CA parameters were investigated using 125 different combinations of parameters ($\kappa=4, 5, 6, 7, \text{ and } 8$; $P_p=0.05, 0.10, 0.15, 0.20, \text{ and } 0.25$; and $P_m=0.05, 0.10, 0.15, 0.20, \text{ and } 0.25$). In all, 5000 simulations were performed as each parameter combination was simulated 40 times. A striking feature of the results is that the wave speed is insensitive to κ . Figure 4(a) shows the mean wave speed plotted against P_p . Three groups of data corresponding to three different values of P_m are shown. Within each of these groups, five profiles are shown corresponding to different values κ . Similar results are shown in Fig. 4(b) where the data are plotted against P_m and the data grouped by P_p and κ . These data show that the wave speed is insensitive to variations in κ while depending strongly on P_m and P_p . Comparing the slope of the profiles in Figs. 4(a) and 4(b) shows that c is more sensitive to variations in P_p than P_m . These outcomes are consistent with some properties of the analogous continuum result for Fisher's equation [Eq. (2)], namely the wave speed is independent of the capacity density and increases with the motility (D) and proliferation (λ). To confirm the apparent insensitivity to κ the entire data set was grouped by κ and the observed distributions are shown in Fig. 4(c) as box plots. The box plots indicate that the median and interquartile range of the observed distributions of c are similar regardless of κ . These data suggest that the CA wave speed may be given by

$$c = \alpha P_p^\beta P_m^\delta. \quad (3)$$

The coefficients are determined using nonlinear regression

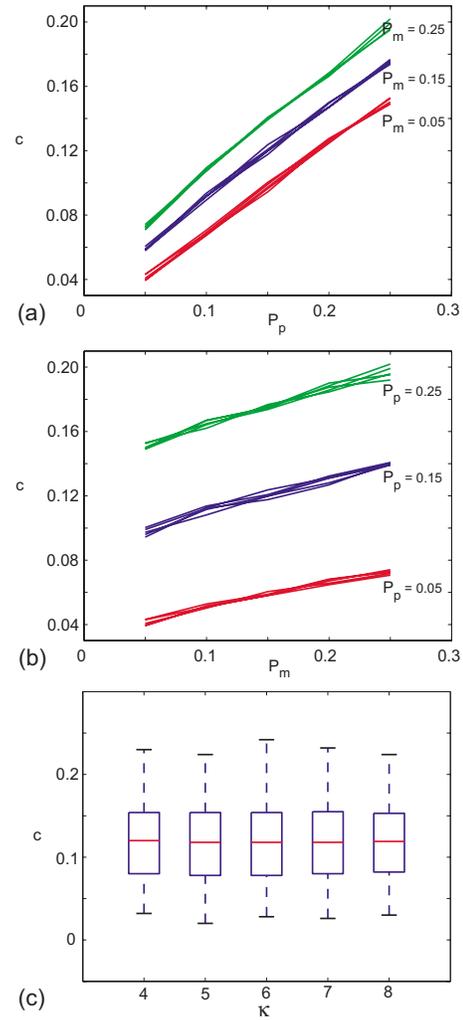


FIG. 4. (Color online) A summary of wave speed results from the CA algorithm showing how the mean wave speed c depends on the three CA parameters P_p , P_m , and κ . Results in (a) show the mean wave speed plotted against P_p . The data are presented in three groups of varying P_m with different colors corresponding to different P_m . Within each of these groups, five profiles are shown for different values of carrying capacity density $\kappa=4, 5, 6, 7, \text{ and } 8$. Similarly, the results in (b) show the mean wave speed plotted against P_m . The data are presented in three groups of varying P_p with different colors corresponding to different P_p . Within each of these groups, five profiles are shown for different values of carrying capacity density $\kappa=4, 5, 6, 7, \text{ and } 8$. Box plots in (c) show that the distribution of mean wave speed data is insensitive to κ .

algorithms in MATHEMATICA giving $\alpha=0.68$, $\beta=0.70$, and $\delta=0.20$ ($R^2=0.98$).

In summary these results show that the motility and proliferation rules in the CA algorithm give rise to constant speed invasion profiles. This is consistent with known population-scale patterns of invasion, such as the invasion wave shown in Fig. 1(a). This relationship is Fisher-like in the sense that the wave speed is independent of the carrying capacity density. The CA wave speed is more sensitive to the proliferation parameter than the motility parameter, whereas the wave speed for Fisher's equation is equally sensitive to the motility and proliferation parameters.

V. CONNECTING THE CELL-SCALE AND POPULATION-SCALE PARAMETERS

Our results have shown that the population-scale mean wave speed is reliably predicted by Eq. (3). This relationship has similarities to the analogous result [Eq. (2)] for continuum invasion models. Further relationships between (i) the individual motility rule and a nondimensional population-scale diffusivity D , and (ii) the individual proliferation rule and a nondimensional population-scale mitotic index λ will now be deduced.

A. Mitotic index

To relate the individual proliferation parameters to equivalent population-scale parameters, a series of computational experiments was performed on a square lattice with 2500 lattice sites. All simulations are performed in the absence of motility ($P_m=0$) with a range of parameters ($\kappa=4, 5, 6, 7$, and 8 and $P_p=0.05, 0.10, 0.15, 0.20$, and 0.25). The initial condition is a randomly seeded lattice with 1% of sites occupied. The CA rules were allowed to evolve until the total number of cells reached a steady value. The CA population density is $U(t)=N(t)/2500$, where $N(t)$ is the number of cells at time step t .

In continuum invasion models, the most common form of carrying capacity-limited proliferation is the logistic model. A generalized logistic model is given by

$$du/dt = \lambda u(1 - [u/K]^m), \quad u(0) = u_0, \quad (4)$$

which has the solution

$$t\lambda = \ln(u/u_0) + (1/m)\ln([K^m - u_0^m]/[K^m - u^m]). \quad (5)$$

The discrete proliferation rule will now be related to the logistic model. CA simulations were performed 40 times for each parameter combination and the results were used to construct various density evolution profiles $U(t)$. For each parameter combination, the $U(t)$ data were averaged to create an ensemble density profile $\hat{U}(t)$. The population-scale capacity density K was estimated as $K=\hat{U}(\infty)$. The $\hat{U}(t)$ profile was fitted to Eq. (5); an example ensemble profile is given in Fig. 5(a), showing that the CA data can be described by the logistic equation. Four logistic curves are shown for varying values of the power m . The shape was best captured with $m < 1$; in particular we found that $m=1/5$ provided a good fit for all simulations. The population-scale parameter λ was estimated using the regression algorithms in MATHEMATICA. Data for all parameters are summarized in Fig. 5(b) showing a clear linear relationship between the individual proliferation rule and the equivalent population-scale parameter,

$$\lambda = \Lambda(m)P_p, \quad (6)$$

where the proportionality constant $\Lambda(m)$ is dependent on the power m . The CA data show that $\Lambda(1/5)=1.28$, $\Lambda(1/3)=0.84$, $\Lambda(1/2)=0.63$, and $\Lambda(1)=0.52$ with $m=1/5$ providing the best fit to the data. This expression shows that the equivalent population-scale description of the CA proliferation rule is independent of the capacity density since results for vari-

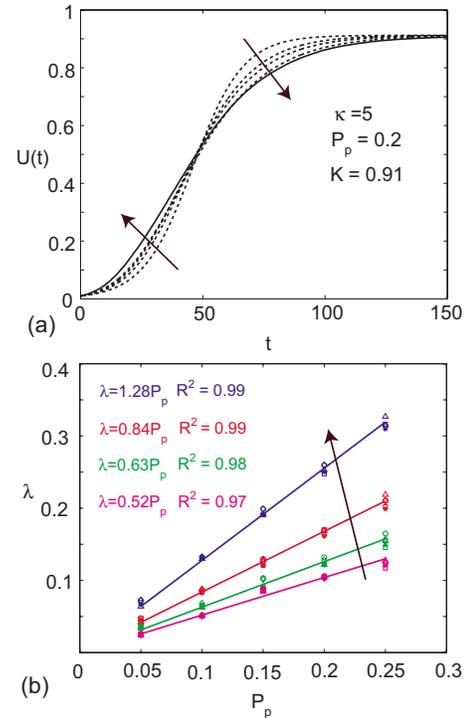


FIG. 5. (Color online) Establishing a relationship between the cell-scale carrying capacity-limited proliferation rule and the population-scale logistic model. An ensemble of CA results with $P_m=0$ is given in (a) showing the time evolution of the averaged CA density $\hat{U}(t)$ (solid line) and several fitted logistic models (dotted lines). Four fits are given for $m=1, 1/2, 1/3$, and $1/5$ with the direction of decreasing m indicated by the arrows. A summary of all results over a wide range of CA parameters is given in (b) showing various relationships between P_p and λ for different values of m . Four straight line fits are given for $m=1, 1/2, 1/3$, and $1/5$ with the direction of decreasing m indicated by the arrow. In each case various results are superimposed for various CA carrying capacities; $\kappa=4$ (circles), $\kappa=5$ (diamonds), $\kappa=6$ (squares), $\kappa=7$ (triangles), and $\kappa=8$ (stars). All results show a linear relationship between the mitotic index λ and the CA proliferation parameter P_p . This relationship is independent of κ .

ous κ follow Eq. (6), as illustrated in Fig. 5(b). We note that a small value of m in Eq. (5) is required to obtain a good fit to the CA data; however, certain difficulties were observed using very small values $m \leq 1/10$ since the values of du/dt in Eq. (5) change very rapidly in time requiring high-quality data. Nonetheless, our results with $m=1/5$ provide a good match.

B. Diffusivity

The invasion simulations involve an unbiased random walk of cells modulated by cell-cell interactions. We anticipate that this rule can be described and parametrized with a nonlinear diffusivity. To estimate the diffusivity, simulations were performed on a lattice with $L_x=L_y=100$ in the absence of proliferation ($P_p=0$). The lattice was randomly seeded with various background densities of $U=0, 0.2, 0.4, 0.6$, and 0.8 and in each case the diffusivity was estimated for three

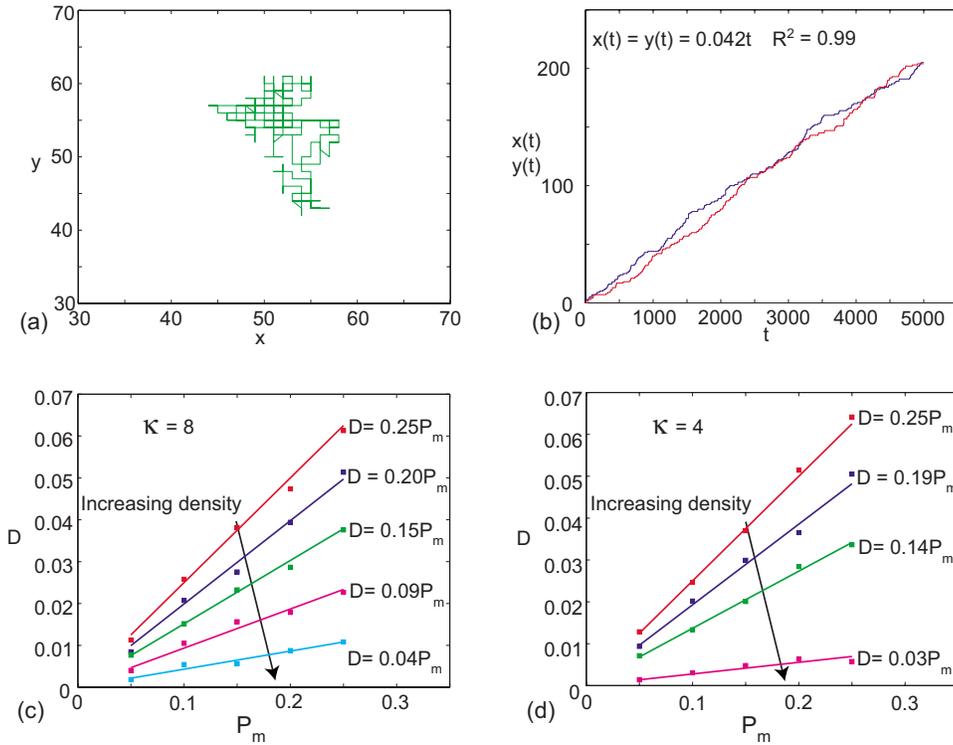


FIG. 6. (Color online) Establishing a relationship between the cell-scale nearest neighbor random walk rule and a diffusivity. An example trajectory is shown in (a) with a background density of $U=0.4$, carrying capacity $\kappa=6$, $P_m=0.15$, and $P_p=0$. The trajectory is shown over 5000 time steps with initial location (50, 50). The time evolution of $x(t)$ (blue or dark gray) and $y(t)$ (red or light gray) is given in (b). These data correspond to the trajectory in (a). These trajectories behave as $x(t)=y(t)=0.042t$ ($R^2=0.99$), giving $D=0.021$. An ensemble of simulations for varying background densities and P_m for $\kappa=8$ and $P_p=0$ are given in (c) with five profiles corresponding to different background densities of 0, 0.2, 0.4, 0.6, and 0.8. The arrow shows the direction of increasing U . An ensemble of simulations for varying background densities and P_m for $\kappa=4$ are given in (d) with four profiles corresponding to different background densities of 0, 0.2, 0.4, and 0.6. The arrow shows the direction of increasing U .

carrying capacity densities $\kappa=4, 6$, and 8. For each simulation, a cell was placed at (50,50) and the CA rules were evolved over 5000 time steps. This procedure was repeated for a range of P_m ($P_m=0.05, 0.10, 0.15, 0.20$, and 0.25). The evolution of the square of the displacements of the tagged cell was recorded:

$$x(t) = \sum_{k=1}^t (x_k - x_{k-1})^2, \quad y(t) = \sum_{k=1}^t (y_k - y_{k-1})^2, \quad (7)$$

where x_k and y_k are the coordinates of the tagged cell at time k . In all simulations $x(t)$ and $y(t)$ evolved linearly in time. The slopes of $x(t)$ versus t and $y(t)$ versus t correspond to $2D$, where D is the diffusivity [7,44].

An example of a random walk used to extract the diffusivity data is given in Fig. 6(a) where the simulation was performed with a background density of $U=0.40$ with parameters $\kappa=6$ and $P_m=0.15$. The time evolution of $x(t)$ and $y(t)$ is given in Fig. 6(b). Results for a wide range of P_m , κ , and U are summarized in Figs. 6(c) and 6(d) showing that the diffusivity is a decreasing function of the background density. This is expected since the background cells impede cell movement due to cell-cell interactions. Comparing Figs. 6(c) and 6(d) shows that the diffusivity increases with κ for $U>0$. This is physically consistent with the discrete rules

since transitions are less likely to occur for lower κ owing to cell-cell interactions and crowding effects. Note that results are not provided for $\kappa=4$ and $U=0.8$ since the tagged cell did not move in these simulations. Overall these results indicate that cells located at the extreme front of an invasion wave will be more motile than their counterparts well behind the wave front which have a larger number of neighbors and a higher local cell density.

The diffusivity data can be summarized in the form of a density-dependent relationship $D=G(U, \kappa)$ with $\partial G/\partial U < 0$. Continuum invasion models with nonlinear nondegenerate decreasing diffusivity and a general logistic proliferation term [Eq. (4)] give rise to invasion waves moving with a minimum speed of $c=2\sqrt{\lambda D(0)}$ [4], where $D(0)$ is the diffusivity at zero cell density. Accordingly, to describe the CA wave speed we are interested in the estimates of $D(0)$ that are given by the top red curve in Figs. 6(c) and 6(d). Note that the relationship between $D(0)$ and P_m is the same regardless of κ since there are no cell-cell interactions in this case. The relevant relationship between the CA proliferation rule and the diffusivity is given by

$$D(0) = 0.25P_m. \quad (8)$$

This result is intuitive since transitions on a unit lattice in uncrowded regions will occur with one-quarter the total

probability of motility since the motility is unbiased. The functional relationship between the individual parameters P_m and P_p and a population-scale diffusivity and mitotic index (6) and (8) can be substituted into the wave speed relationship [Eq. (3)] to give an equivalent relationship between the population-scale CA wave speed and the relevant population-scale parameters. This procedure illustrates that it is possible to identify how the individual rules in the CA algorithm interact to emerge into a well-defined and predictable pattern at the population level.

In summary, we have shown that the CA algorithm predicts well-known properties of invasion that are associated with continuum invasion models. Namely, it is possible to prescribe individual CA rules which emerge into constant speed waves of invasion. The relationship between the wave speed and the individual-scale parameters can be deduced through computational experimentation. This procedure demonstrates the ability of the CA algorithm to reproduce all forms of data that can be extracted from a continuum invasion model. We will now discuss certain limitations of continuum invasion models and show how we can make further use of the CA algorithm by extracting finer cell-scale data which are not available from continuum models. This additional data can assist in overcoming some of the difficulties in using continuum models to describe invasion processes.

VI. POPULATION-SCALE REDUNDANCY

The CA algorithm is designed to simulate general invasion processes, yet it can be easily modified to account for certain application-specific features. For example, specific crowding and cell-cell interactions (Secs. III A and III B) were introduced to reflect known properties of the NC cell invasion system. Additional simulations under modified conditions, for example, allowing cell removal when the local density rises above capacity density (that is, $M > \kappa$), shows that these details are unimportant with regard to the generation of constant speed invasion waves. We have found that altering the motility mechanism by introducing various forms of biased cell movements also leads to constant speed invasion waves. This was achieved by making transitions in the positive x direction more likely than transitions to the negative x direction. In particular, we experimented by introducing both uniformly biased movements as well as various forms of density-dependent bias rules. The wave speed for these modified conditions can also be analyzed in terms of the population-scale relationship given by Eq. (3) with different parameters α , β , and δ .

A similar redundancy is observed in continuum invasion models. Recent numerical and perturbation analyses of a range of continuum invasion models with different motility mechanisms concluded that the formation of constant speed invasion waves is insensitive to the particular motility mechanism and that it is the carrying capacity-limited proliferation mechanism which largely controls this aspect of the system [4]. Several plausible models with different motility mechanisms can all be parametrized to fit experimental wave speed data. For example, Fisher's equation assumes that cells move according to a linear diffusion motility rule and yield

traveling wave solutions moving with speed $c \propto D^{0.5}$, where D is the diffusivity. A chemotaxis continuum model also yields traveling wave solutions, which move with speed $c \propto \chi^{0.5}$ [45], where χ is the chemotactic sensitivity. Therefore it is possible to apply both models and calibrate values of the motility coefficients by fitting data to the expression for the wave speed c . While it is possible to calibrate a range of continuum models to match experimental wave-speed data [2,4,23], this does not necessarily provide a comprehensive understanding of the details of the motility mechanisms in the invasion system.

Population-scale properties such as the traveling wave speed do not provide sufficient information to identify the details of the motility mechanism in an invasion system. These limitations can be overcome if additional cell-scale data can be extracted from mathematical models. Discretizing continuum partial differential equations introduces multiple cells on lattice sites—this is inappropriate if the lattice spacing corresponds to the dimensions of a cell, or alternatively means that an intermediate-scale and not a cell-scale approach is being used. Furthermore, standard continuum invasion models implicitly permit cell death [7], which may be inappropriate for some applications. Discrete CA models do not have these limitations, since they do not have to correspond to any partial differential equation. CA models can be used to identify additional cell-scale properties of the system which can be sensitive to the choice of cell motility rules. Additional information, such as tracking individual cell trajectories, is available from a cell-based CA model, and allows a more detailed probing of active mechanisms and interactions inside the invasion wave.

VII. USING THE CA ALGORITHM TO LOOK INSIDE AN INVASION WAVE

Recent experimental analyses of NC cell invasion have involved labeling either a subset of NC cells within the wave [17,18] or labeling the entire NC cell population [20] and performing time-lapse imaging of the system as the invasion proceeds. These experiments were used to extract cell-scale data and in particular to track individual NC cell trajectories within the invasion wave. The most recent results, presented by Druckenbrod and Epstein [18], give details of the individual NC cell trajectories as a function of position within the invasion wave. This level of detail cannot be captured by standard continuum models. The ability of the CA algorithm to replicate this kind of cell-scale phenomena will now be demonstrated.

To track individual trajectories within the CA algorithm, a series of computational experiments, demonstrated in Fig. 7, were performed. Results were generated for a dense ($\kappa=8$) and less dense ($\kappa=4$) invasion wave. These waves were given sufficient time to establish a constant speed invasion profile. The algorithm was halted and two cells within the wave were tagged. To be consistent with experimental results, we tagged the most advanced cell (trajectory 1) and a second cell close behind the wave front (trajectory 2), as shown in Fig. 7. The tagged cells follow the same rules as the background cells. The system was allowed to evolve and

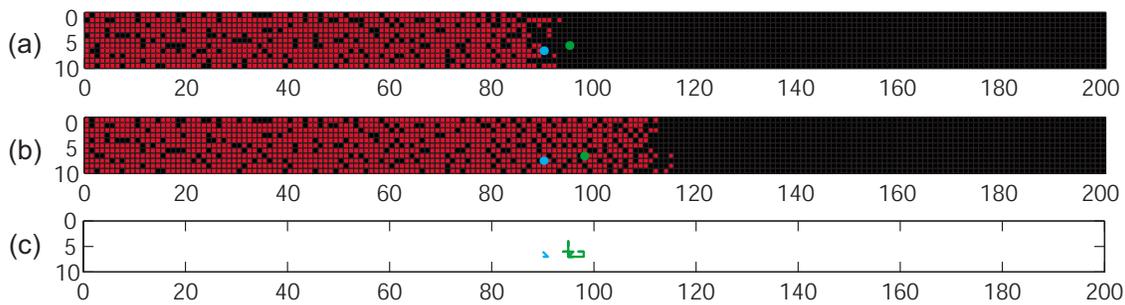


FIG. 7. (Color online) The cell-scale dynamics are demonstrated with a single realization of the CA algorithm with $P_p=0.15$, $P_m=0.15$, and $\kappa=4$. A profile of the invasion wave is given in (a) showing the left to right progression of the population. The simulation was halted at this particular time and two cells were tagged with green (trajectory 1) and blue (trajectory 2) dots which have been enlarged for ease of visualization. Trajectory 1 is associated with the most advanced cell, while trajectory 2 is slightly behind. The CA simulation progressed for 200 time steps to produce the profile in (b). The trajectories of the tagged cells over the 200 time steps are given in (c).

the trajectories of the tagged cells were tracked.

The individual trajectory simulations were repeated 40 times with the same initial condition to generate a suite of trajectory data from which ensemble properties can be measured. Results are presented for $\kappa=4$ only, as there were no major differences between these results and those with $\kappa=8$. Displacement vectors, given in row (a) of Fig. 8, represent the total displacement for each of the 40 simulations over 200 time steps. This time interval was chosen to ensure that the position of the wave front relative to the tagged cells did not change dramatically over the time period. For example, a cell at the leading edge should remain relatively close to the leading edge during the time interval. Total displacement is the difference between the final and initial po-

sitions of the tagged cells. Statistical information about the direction of displacement is given in row (b) as a histogram of directions showing the frequency of movements in a particular direction. The CA data are presented in this way to be consistent with the experimental direction frequency data of Young *et al.* [20] (although this experimental work recorded the position of the leading cell rather than the position of a uniquely tagged cell). Note that there are not 40 unique displacement vectors shown in each diagram in row (a) of Fig. 8. There are two reasons for this: (i) in some realizations there is no net cell movement, and (ii) the displacement vectors for some realizations are identical to the displacement vectors in other realizations meaning that some displacement vectors are indistinguishable.

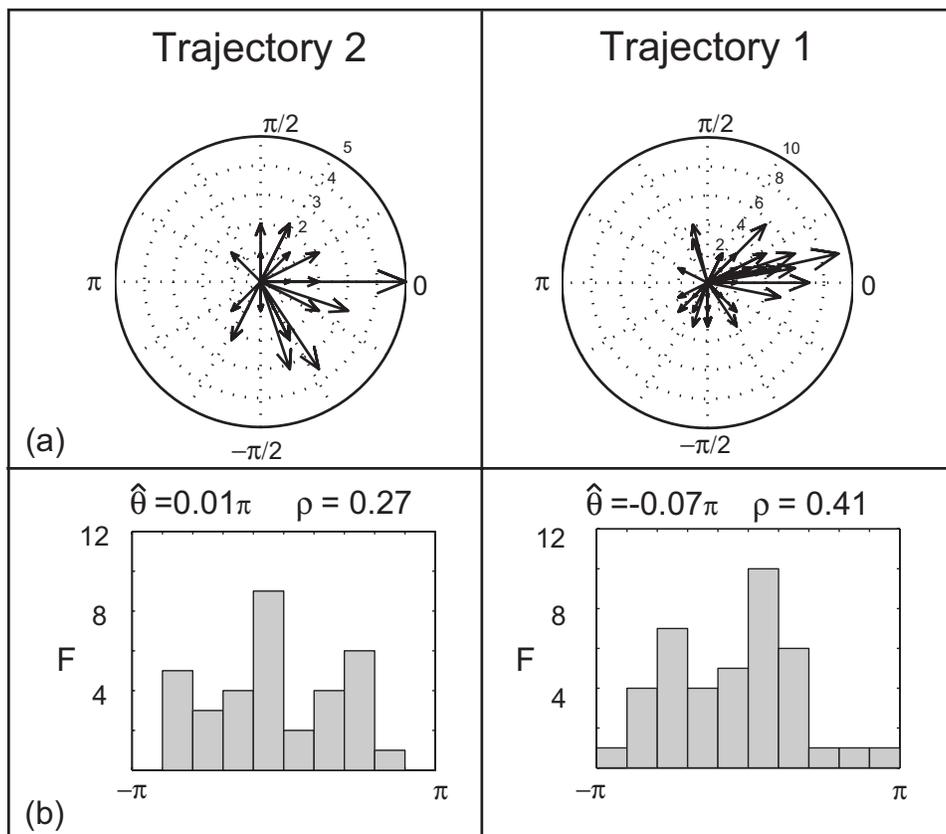


FIG. 8. Trajectories 1 and 2 correspond to the most advanced cell and a cell in the wave front, respectively (see Fig. 7). All results are generated with $P_p=0.15$, $P_m=0.15$, and $\kappa=4$; a suite of 40 simulations was performed. Results in row (a) show the displacement vectors of the tagged cells indicating both the displacement and direction of the cell trajectory over 200 time steps. Note that the polar diagrams have different scales. Results in row (b) show the frequency of movement in various directions for the 40 realizations. The mean direction $\hat{\theta}$ and polarization ρ are given to quantify directional trends.

Certain properties of the trajectory data can be described statistically. The mean angle of the sample $\hat{\theta}$, as defined by Fisher [46], is evaluated for each data set. It is convenient to define the mean angle to be in the interval $-\pi < \hat{\theta} \leq \pi$ with $\hat{\theta}=0$ parallel to the positive x axis. Highly aligned data will have the directions of the individual trajectories clustered around the $\hat{\theta}$ direction. If the data are not highly aligned the directions of the individual trajectories will not be clustered around $\hat{\theta}$. A measure of alignment about the mean angle is given by the polarization ρ , which can be written as [46,47]

$$\rho = \frac{1}{\mathcal{N}} \left| \sum_{i=1}^{\mathcal{N}} \mathbf{x}_i \right|, \quad (9)$$

where \mathbf{x}_i are the individual unit direction vectors, $|\cdot|$ is the vector magnitude, and \mathcal{N} is the sample size. This measure lies in the range $0 \leq \rho \leq 1$ with $\rho=1$ corresponding to a perfectly aligned group with all vectors pointing in the direction of $\hat{\theta}$, and $\rho=0$ is a group with a large distribution of angles. As noted by Fisher [46] the quantity $(1-\rho)$ is similar to the variance of univariate data.

The individual trajectory data show that trajectory 2 undergoes much smaller displacements than trajectory 1. This is because the movement of cells behind the wave front is impeded by cell-cell contact. This is consistent with the previous measurements of diffusivity where our experiments showed that cell motility decreased with increasing cell density. Although individual isolated cells make unbiased transitions, it is not clear whether the cell-cell interactions might bias the movement of cells within the wave or whether the bias might depend on the position of a cell within the wave. The data show that the cell associated with trajectory 1 is more likely to move in the $\hat{\theta}$ direction as the polarization values indicate that the data for trajectory 1 are more aligned with the mean direction compared with the data for trajectory 2. This indicates that the CA rules predict that cells at or near the wave front follow trajectories which are biased in the direction of the net invasion. This outcome is qualitatively consistent with experimental observations made by Druckenbrod and Epstein [18]. The degree of polarization observed in this kind of computational experiment can be modified by using different variations on the motility rule.

VIII. CONCLUSION

We present and investigate a computational tool to replicate both the population-scale and cell-scale properties of invasive cell populations. The CA algorithm combines an individual cell motility rule with a carrying capacity-limited proliferation rule devised according to biologically plausible observations of cell invasion.

The CA algorithm predicts a wave of invasion moving with constant speed, which is consistent with well-known continuum results and experimental observations. The rela-

tionship between the population-scale pattern of invasion and the CA parameters is deduced to give an expression for the wave speed analogous to that associated with continuum models. Both the discrete and continuum relationships for the wave speed increase with the motility and proliferation rate of cells while being independent of the carrying capacity density. Therefore we have made a connection between the cell-scale and population-scale processes replicated by the CA algorithm. Subtle differences between the continuum and discrete wave speed expressions are observed. The analytically derived wave speed for Fisher's equation is equally sensitive to the motility and proliferation parameters since $c \propto \lambda^{1/2} D^{1/2}$, while the discrete fitted wave speed is more sensitive to the proliferation parameter than the motility since $c \propto \lambda^{7/10} D^{1/5}$. The exponents in the latter expression are empirical values derived from our simulations. The values of these exponents depend on the choice of rules in the CA algorithm.

Similar to continuum invasion models, we find that the carrying capacity-limited (contact-inhibited) proliferation rule in the CA algorithm largely controls traveling wavelike behavior of the system since the invasion waves are relatively insensitive to the details of the cell motility mechanism. This redundancy poses a major limitation on the use of population-scale properties alone to deduce mechanisms from experimental data. Continuum models only produce population-scale properties and therefore always suffer from this limitation. The CA algorithm has the advantage that additional cell-scale data can be collected and compared with experimental data to assist in deducing the mechanisms driving invasion. Our CA algorithm permits individual cell trajectories to be tracked and the properties of these trajectories to be quantified. Consequently the CA algorithm replicates both population level and cell-scale experimental data, providing an opportunity to match multiscale experimental data.

Our approach of devising experimentally motivated individual CA rules is advantageous over extracting individual rules from a discretized partial differential equation. We rely solely on experimental observations and avoid making underlying assumptions about a partial differential equation description of the invasion process. This approach is broadly applicable to various applications including wound healing, tissue engineering, and tumor invasion. We anticipate that with the increasing use of time-lapse data in cell biology, theoreticians must embrace the idea of extracting relevant cell-scale rules to build predictive and interpretive models of biological processes.

ACKNOWLEDGMENTS

This work was supported by the Australian Research Council (ARC). We appreciate the support of the Particulate Fluids Processing Centre, an ARC Special Research Centre. We thank Joanne Britto, Don Newgreen, Ray Watson, and Heather Young for their advice and assistance.

- [1] A. R. A. Anderson and M. A. J. Chaplain, *Bull. Math. Biol.* **60**, 857 (1998).
- [2] P. K. Maini, D. L. S. McElwain, and D. Leavesley, *Tissue Eng.* **10**, 475 (2004).
- [3] J. D. Murray, *Mathematical Biology I: An Introduction*, 3rd ed. (Springer-Verlag, Heidelberg, 2002).
- [4] M. J. Simpson, K. A. Landman, B. D. Hughes, and D. F. Newgreen, *J. Theor. Biol.* **243**, 343 (2006).
- [5] M. J. Simpson, D. C. Zhang, M. Mariani, K. A. Landman, and D. F. Newgreen, *Dev. Biol.* **302**, 553 (2007).
- [6] U. Börner, A. Deutsch, H. Reichenbach, and M. Bär, *Phys. Rev. Lett.* **89**, 078101 (2002).
- [7] A. Q. Cai, K. A. Landman, and B. D. Hughes, *J. Theor. Biol.* **245**, 576 (2007).
- [8] R. Erban and H. G. Othmer, *SIAM J. Appl. Math.* **65**, 361 (2004).
- [9] D. Longo, S. M. Peirce, T. C. Skalak, L. Davidson, M. Marsden, B. Dzamba, and D. W. DeSimone, *Dev. Biol.* **271**, 210 (2004).
- [10] R. M. H. Merks and J. A. Glazier, *Physica A* **352**, 113 (2005).
- [11] S. M. Peirce, E. J. Van Gieson, and T. C. Skalak, *FASEB J.* **18**, 731 (2004).
- [12] N. J. Savill and J. A. Sherratt, *Dev. Biol.* **258**, 141 (2003).
- [13] S. Turner and J. A. Sherratt, *J. Theor. Biol.* **216**, 85 (2002).
- [14] D. Noble, *Nat. Rev. Mol. Cell Biol.* **3**, 459 (2002).
- [15] S. L. M. A. Beeres, F. M. Bengel, J. Bartunek, D. E. Atsma, J. M. Hill, M. Vanderheyden, M. Penicka, M. J. Schalijs, W. Wijns, and J. J. Bax, *J. Am. Coll. Cardiol.* **49**, 1137 (2007).
- [16] L. M. Bernas, P. J. Foster, and B. K. Rutt, *J. Neurosurg.* **106**, 306 (2007).
- [17] N. R. Druckenbrod and M. L. Epstein, *Dev. Biol.* **287**, 125 (2005).
- [18] N. R. Druckenbrod and M. L. Epstein, *Dev. Dyn.* **236**, 84 (2007).
- [19] W. Saadi, S.-J. Wang, F. Lin, and N. L. Jeon, *Biomed. Microdevices* **8**, 109 (2006).
- [20] H. M. Young, A. J. Bergner, R. B. Anderson, H. Enomoto, J. Milbrandt, D. F. Newgreen, and P. M. Whittington, *Dev. Biol.* **270**, 455 (2004).
- [21] A. Deutsch and S. Dormann, *Cellular Automaton Modeling of Biological Pattern Formation* (Birkhäuser, Boston, 2005).
- [22] J. A. Sherratt and M. A. J. Chaplain, *J. Math. Biol.* **43**, 291 (2001).
- [23] B. G. Sengers, C. P. Please, and R. O. C. Oreffo, *J. R. Soc. Interface* (to be published).
- [24] M. J. Simpson, K. A. Landman, and D. F. Newgreen, *J. Comput. Appl. Math.* **192**, 282 (2006).
- [25] M. Alber, N. Chen, T. Glimm, and P. M. Lushnikov, *Phys. Rev. E* **73**, 051901 (2006).
- [26] A. F. M. Marée and P. Hogeweg, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3879 (2001).
- [27] R. H. M. Merks and J. A. Glazier, *Nonlinearity* **19**, C1 (2006).
- [28] N. J. Popławski, M. Swat, J. Scott Gens, and J. A. Glazier, *Physica A* **373**, 521 (2007).
- [29] S. Turner, J. A. Sherratt, K. J. Painter, and N. J. Savill, *Phys. Rev. E* **69**, 021910 (2004).
- [30] T. Alarcón, H. M. Byrne, and P. K. Maini, *J. Theor. Biol.* **225**, 257 (2003).
- [31] J. Galle, G. Aust, G. Schaller, T. Beyer, and D. Drasdo, *Cytometry A* **69A**, 704 (2006).
- [32] V. Grimm, E. Revilla, U. Berger, F. Jeltsch, W. M. Mooij, S. F. Railsback, H.-H. Thulke, J. Weiner, T. Wiegand, and D. L. DeAngelis, *Science* **310**, 987 (2005).
- [33] J. K. Parrish and L. Edelstein-Keshet, *Science* **284**, 99 (1999).
- [34] A. Q. Cai, K. A. Landman, and B. D. Hughes, *Bull. Math. Biol.* **68**, 25 (2006).
- [35] K. J. Painter and J. A. Sherratt, *J. Theor. Biol.* **225**, 327 (2003).
- [36] A. Stéphanou, S. R. McDougall, A. R. A. Anderson, and M. A. J. Chaplain, *Math. Comput. Model.* **44**, 96 (2006).
- [37] T. Callaghan, E. Khain, L. M. Sander, and R. M. Ziff, *J. Stat. Phys.* **122**(5), 909 (2006).
- [38] D. C. Walker, J. Southgate, G. Hill, M. Holcombe, D. R. Hosea, S. M. Wood, S. Mac Neil, and R. H. Smallwood, *Bio-systems* **76**, 89 (2004).
- [39] D. F. Newgreen, I. Jahnke, I. J. Allan, and I. Lewis Gibbins, *Cell Tissue Res.* **208**, 1 (1980).
- [40] R. A. Fisher, *Ann. Eugen.* **7**, 353 (1937).
- [41] J. Bear, *Dynamics of Fluids in Porous Media* (Elsevier, New York, 1972).
- [42] M. J. Simpson, K. A. Landman, and K. Bhaganagarapu, *J. Theor. Biol.* **247**, 525 (2007).
- [43] S. Gianino, J. R. Grider, J. Cresswell, H. Enomoto, and R. O. Heuckeroth, *Development* **130**, 2187 (2003).
- [44] B. D. Hughes, *Random Walks and Random Environments* (Oxford University Press, Oxford, UK, 1995), Vol. 1.
- [45] K. A. Landman, M. J. Simpson, J. L. Slater, and D. F. Newgreen, *SIAM J. Appl. Math.* **65**, 1420 (2005).
- [46] N. I. Fisher, *Statistical Analysis of Circular Data* (Cambridge University Press, Cambridge, 1993).
- [47] A. Merrifield, M. R. Myerscough, and N. Weber, *Math. Biosci.* **203**, 64 (2006).