Pathlines in exclusion processes

Matthew J. Simpson,* Kerry A. Landman, and Barry D. Hughes
Department of Mathematics and Statistics, The University of Melbourne, Victoria 3010, Australia
(Received 27 October 2008; revised manuscript received 23 February 2009; published 27 March 2009)

Trajectory data from observations of a random-walk process are often used to characterize macroscopic transport coefficients and to make inferences about motility mechanisms. Continuum equations describing the average moments of the position of an agent in an exclusion process are derived and validated with simulation data. Unlike standard noninteracting random walks, the moment equations for the exclusion process explicitly represent the interaction of agents since they depend on the averaged macroscopic agent density. Key issues associated with the validity of the continuum equations and interpretation of experimental data are discussed.

DOI: 10.1103/PhysRevE.79.031920 PACS number(s): 87.10.Ed, 87.10.Hk

I. INTRODUCTION

Several kinds of interacting particle models have been applied to biological phenomena. These include the following: contact processes that model infection spreading [1]; voter processes that model ecological interactions [2]; the Eden model that represents the growth of cell colonies [3]; and exclusion processes that model the motility of individuals within a population [4]. Discrete random-walk models based on a simple exclusion process are often used to represent cell motility [5–9]. Simple exclusion processes are a natural choice to represent cell motility as they incorporate realistic cell-cell interactions by allowing at most one motile agent per lattice site.

Experimental cell trajectory data are routinely analyzed to improve our understanding of collective behavior in cell populations [10]. Trajectory data can be analyzed by quantifying various properties such as the total trajectory length, velocity, mean-square displacement, persistence time, trajectory orientation, and turning frequency [5,11–17]. In this work we develop equations governing average agent trajectories that are relevant to random-walk models based on a simple exclusion process.

II. MODELING PATHLINES IN AN EXCLUSION PROCESS

A. Population-level density information

An asymmetric simple exclusion process [1,18], with a random sequential update method [4], is realized on a two-dimensional square lattice with spacing Δ. In each time step, of duration τ, all agents have the opportunity to move with probability P. A motile agent at (x,y) steps to $(x,y\pm\Delta)$ with probability $(1\pm\rho_y)/4$, or to $(x\pm\Delta,y)$ with probability $(1\pm\rho_x)/4$, where $|\rho_x|\leq 1$ and $|\rho_y|\leq 1$. If an agent attempts to move into an occupied site then that motility event is aborted.

The simple exclusion process is related to a continuum partial differential equation in the appropriate limit as $\Delta\to 0$ and $\tau\to 0$ [7,19–22]. If the average occupancy of site (i,j), averaged over many realizations, is $\langle C_{i,j}\rangle$, the spatial and temporal evolutions of the corresponding continuous density $C(x,y,t)$ is governed by

$$\frac{\partial C}{\partial t}=D\nabla^2 C-\nabla\cdot[\mathbf{v}C(1-C)], \tag{1}$$

where $\mathbf{v}=(v_x,v_y)$ is the advective velocity with components

$$v_x=\lim_{\Delta,\tau\to 0}\left(\frac{P\rho_x\Delta}{2\tau}\right),\quad v_y=\lim_{\Delta,\tau\to 0}\left(\frac{P\rho_y\Delta}{2\tau}\right),$$

and the diffusivity is given by

$$D=\lim_{\Delta,\tau\to 0}\left(\frac{P\Delta^2}{4\tau}\right).$$

To obtain a well-defined continuum limit, we require that if ρ_x and ρ_y are nonzero (biased motility), they decrease to zero as $\rho_x=O(\Delta)$ and $\rho_y=O(\Delta)$ [23]. This nonlinear advection diffusion model is related to traffic flow models and Burgers' equation [24]. All continuum equations developed in this work assume that adjacent lattice sites have independent occupancy status. This neglect of spatial correlations at each instant is questionable for individual realizations but proves much more satisfactory for averaged properties as evidenced by simulation data reported here.

B. Individual-level pathline information

Instead of focusing on the evolution of the averaged agent macroscopic density, we take a different point of view and model the average evolution behavior of a tagged agent within the population. The expected displacement of a tagged agent at site (i,j) during the next time step is

$$\delta p_x=\Delta P\frac{(1+\rho_x)}{4}(1-\langle C_{i+1,j}\rangle)-\Delta P\frac{(1-\rho_x)}{4}(1-\langle C_{i-1,j}\rangle), \tag{2}$$

$$\delta p_y=\Delta P\frac{(1+\rho_y)}{4}(1-\langle C_{i,j+1}\rangle)-\Delta P\frac{(1-\rho_y)}{4}(1-\langle C_{i,j-1}\rangle). \tag{3}$$

The factors $(1-\langle C_{i\pm 1,j}\rangle)$ and $(1-\langle C_{i,j\pm 1}\rangle)$ in Eqs. (2) and (3) account for attempted moves that are aborted due to the tar-

*m.simpson@ms.unimelb.edu.au

get site being occupied. Dividing these expressions by τ and holding Δ^2/τ constant, we let $\Delta \to 0$ and $\tau \to 0$ simultaneously giving

$$\frac{dp_x}{dt} = -2D\frac{\partial C}{\partial x} + v_x(1-C), \tag{4}$$

$$\frac{dp_y}{dt} = -2D\frac{\partial C}{\partial y} + v_y(1-C). \tag{5}$$

The solution of these differential equations gives $p_x(t)$ and $p_y(t)$, which are the coordinates of the average trajectory of a tagged agent in an exclusion process initially at position $[p_x(0), p_y(0)]$. The curve given by $[p_x(t), p_y(t)]$ is analogous to a pathline in potential flow [25].

Two important differences between the equations governing the evolution of the average pathlines in an exclusion process compared a standard noninteracting random walk [10,23] are as follows: (i) an unbiased exclusion process ($\rho_x = \rho_y = 0$) can give rise to a net drift due to the average agent density gradient, whereas the net drift of agents undergoing a standard noninteracting unbiased random walk is always zero, and (ii) the average pathlines in an exclusion process depend on an interaction with the average agent density distribution, whereas there is no such interaction for a standard random-walk model. These differences have several important and practical consequences which we will explore by comparing the solution of Eq. (4) with simulation data.

Continuum equations for the evolution of the pathline variance (the variance of the tagged agent's position) are now developed. The discrete equations governing the variance are taken as

$$\delta(\sigma_x^2) = P(\Delta - \delta p_x)^2\frac{(1+\rho_x)}{4}(1-\langle C_{i+1,j}\rangle) + (0-\delta p_x)^2 \times\left(1-\frac{P}{2}\right) + P(-\Delta - \delta p_x)^2\frac{(1-\rho_x)}{4}(1-\langle C_{i-1,j}\rangle), \tag{6}$$

$$\delta(\sigma_y^2) = P(\Delta - \delta p_y)^2\frac{(1+\rho_y)}{4}(1-\langle C_{i,j+1}\rangle) + (0-\delta p_y)^2 \times\left(1-\frac{P}{2}\right) + P(-\Delta - \delta p_y)^2\frac{(1-\rho_y)}{4}(1-\langle C_{i,j-1}\rangle). \tag{7}$$

These equations neglect spatial occupancy correlations and treat the abortion of moves to occupied sites exactly as in Eqs. (2) and (3). They also carry an additional approximation, the neglect of temporal correlations in the motion (effectively successive steps are taken as independent). To see the effect of this, let X denote the x coordinate of the position of the tagged particle at a given time in one realization, and δX denote the x displacement increment in the next time step. If temporal correlations are retained, the average of $2(X-p_x)(\delta X - \delta p_x)$ must be added to the right-hand side of Eq. (6). A way to do this using only averages of tractable quantities is not easily apparent.

Dividing Eqs. (6) and (7) by τ and letting $\Delta \to 0$ and $\tau \to 0$ in the usual way gives

$$\frac{d(\sigma_x^2)}{dt} = \frac{d(\sigma_y^2)}{dt} = 2D(1-C). \tag{8}$$

The solution of these differential equations gives $\sigma_x^2(t)$ and $\sigma_y^2(t)$, which are the predictions, based on the stated approximations, of the variances of the tagged agent's coordinates.

We note that as well as the issue of temporal correlations, there are other differences between the discrete and continuum equations that describe the evolution of the average pathlines and the pathline variance. For both we assume the macroscopic density is governed by Eq. (1), which is obtained in the limit where $\Delta, \tau \to 0$, $\rho_x = O(\Delta)$, and $\rho_y = O(\Delta)$ such that Δ^2/τ is finite. After dividing by τ, each term in the discrete expressions for the agent displacement [Eqs. (2) and (3)] gives rise to terms which are proportional to $O(\Delta^2/\tau)$ and $O(\rho_x\Delta/\tau)$ or $O(\rho_y\Delta/\tau)$ and therefore contributes a term dependent on the diffusivity and the advective velocity in the continuum model [Eqs. (4) and (5)]. In contrast, using the same procedure for the discrete variance expressions [Eqs. (6) and (7)], there is only one term proportional to $O(\Delta^2/\tau)$, while the terms involving the bias are $O(\rho_x\Delta^2/\tau)$ or $O(\rho_y\Delta^2/\tau)$ or smaller. Therefore, in the continuum limit, the variance is not explicitly dependent on the advective velocity.

To solve the continuum models presented here, we approximate Eq. (1) numerically with a centered-in-space finite difference approximation and implicit Crank-Nicolson time integration with time step h [26]. Within each time step used to solve Eq. (1), we also integrate Eqs. (4), (5), and (8) over the same time interval with a Crank-Nicolson method [26]. A pathline with initial position $[p_x(0), p_y(0)]$ is determined by approximating $\partial C/\partial x$, $\partial C/\partial y$, and C at $[p_x(0), p_y(0)]$ using central differences and linear interpolation. With this information Eqs. (4), (5), and (8) are integrated giving $[p_x(h), p_y(h)]$ and $[\sigma_x^2(h), \sigma_y^2(h)]$. This procedure is repeated over many time steps giving $[p_x(t), p_y(t)]$ and $[\sigma_x^2(t), \sigma_y^2(t)]$.

To examine how the solutions of Eqs. (4), (5), and (8) compare with averaged simulation results, we perform a series of computational experiments as follows. A particular agent within a population undergoing an exclusion process is tagged. The trajectory of the tagged agent is averaged over M identically prepared realizations to give

$$\langle x(t)\rangle = \frac{1}{M}\sum_{m=1}^{M} x_m(t), \quad \langle y(t)\rangle = \frac{1}{M}\sum_{m=1}^{M} y_m(t). \tag{9}$$

The solution of Eq. (8), with $\sigma_x^2(0) = \sigma_y^2(0) = 0$, will be compared with the evolution of the sample variance given by

$$s_x^2(t) = \frac{1}{(M-1)}\sum_{m=1}^{M}[x_m(t) - \langle x(t)\rangle]^2, \tag{10}$$

$$s_y^2(t) = \frac{1}{(M-1)}\sum_{m=1}^{M}[y_m(t) - \langle y(t)\rangle]^2. \tag{11}$$

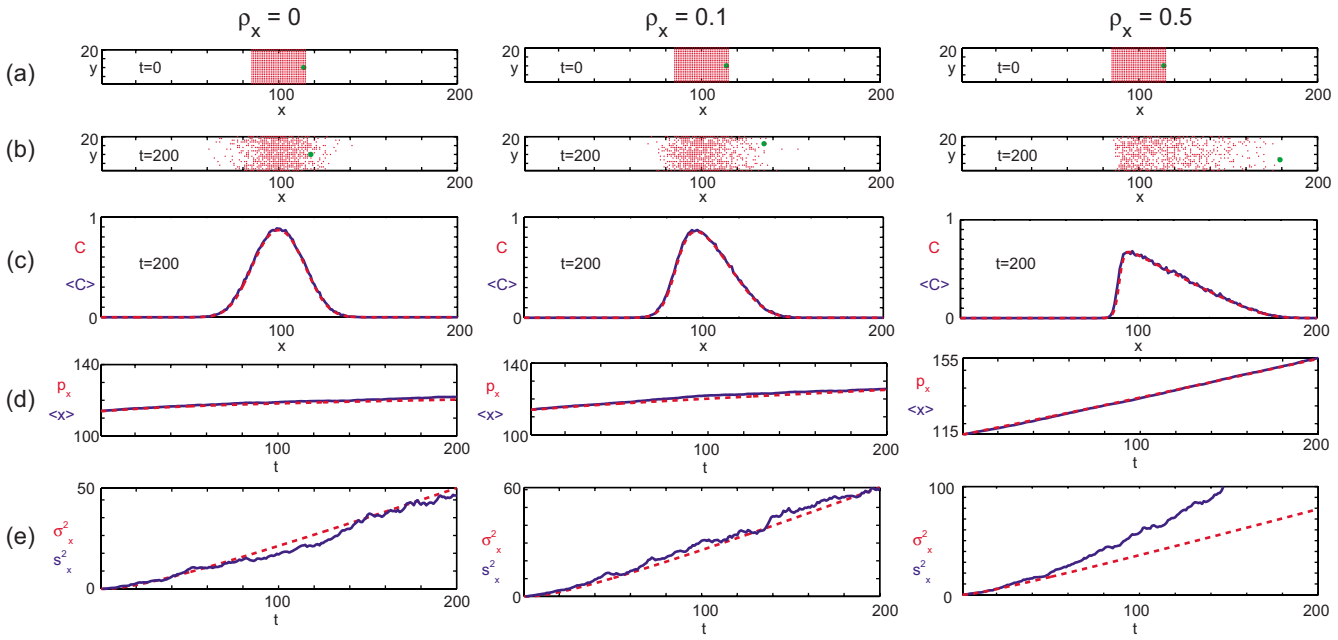

FIG. 1. (Color online) Averaged agent density and pathline results with $P=1$ and $\rho_y=0$. Results are given for different amounts of motility bias in the x direction with $\rho_x=0$ (left column), $\rho_x=0.1$ (middle column), and $\rho_x=0.5$ (right column). Snapshots of a single realization at $t=0$ and 200 are given in rows (a) and (b), respectively, showing the evolution of a population of agents and the position of a tagged agent initially at (114,10) is shown with an enlarged green bullet. Row (c) shows the averaged density profiles $\langle C\rangle$ (solid blue) extracted from the simulation data compared with the continuum solution of Eq. (1) C (dotted red). Rows (d)–(e) compare the evolution of the mean position and variance for the tagged agent. Continuum results (dotted red) are compared with simulation data (solid blue). All simulation results are averaged over 100 Monte Carlo realizations with $\Delta=\tau=1$. All continuum models are solved numerically with a spatial discretization $\delta x=0.1$ and temporal discretization $h=0.25$.

III. COMPARING CONTINUUM RESULTS AND SIMULATION DATA

We compare discrete and continuum results where the initial average occupancy of all sites within every column is a constant and $\rho_y=0$, so that there is bias present at most in the x direction. This configuration reduces the system to a one-dimensional problem where we need only consider the x component of the pathline data [22]. A one-dimensional form of Eq. (1) is considered, and we will compare (i) $p_x(t)$ with $\langle x(t)\rangle$ and (ii) $\sigma_x^2(t)$ with $s_x^2(t)$.

Simulations are performed on a lattice with $L_x=200$, $L_y=20$, and periodic boundary conditions at $y=1$ and $y=L_y$. Initially all sites between $85\leq x\leq 115$ are occupied and a tagged agent is placed at (114,10). The three columns in Fig. 1 contain results for different values of the bias parameter, $\rho_x=0$, 0.1, and 0.5. Snapshots of the agents at $t=0$ and 200 [Fig. 1 row (a)–(b)] illustrate the influence of the motility bias as the unbiased population ($\rho_x=0$) spreads symmetrically about $x=100$, whereas the biased populations ($\rho_x>0$) move in the positive x direction as expected. The influence of the motility bias is also apparent in Fig. 1 row (c), where the continuum solution of Eq. (1) is compared with averaged one-dimensional density profiles from the simulation data. The density profiles from the simulation data are obtained by averaging the occupancy across the height of the lattice L_y for M identically prepared realizations [22]. In all cases the continuum and discrete profiles compare very well and the influence of the nonlinear advection term is apparent as the density profile becomes triangular and moves in the positive x direction as ρ_x increases.

Results in Fig. 1 row (d) show that the evolution of the average pathline for our tagged agent predicted by Eq. (4) accurately matches the simulation data in all cases. Regardless of the bias parameter the average pathline drifts in the positive x direction. This result for the unbiased case is very interesting as Eq. (1) reduces to a linear diffusion equation, which is traditionally associated with a noninteracting unbiased random walk and no net drift [10]. For the unbiased exclusion process the average pathline drifts from $x=114$ to $x=120.4$ during the interval $0\leq t\leq 200$. This observation highlights the importance of interpreting pathline data appropriately. The details of the pathline depend strongly on the initial location of the tagged agent within the population. For the initial condition in Fig. 1, the unbiased pathline starting from $(x,y)=(86,10)$ drifts to the left (not shown), whereas the unbiased pathline starting from (114,10) drifts to the right. Furthermore, if the tagged agent is placed symmetrically within a distribution of unbiased agents, then its averaged pathline would not drift at all due to symmetry. This highlights the critical importance of interpreting pathline data relative to the initial distribution of the population and the initial point on the pathline.

Results in Fig. 1 row (e), compare the evolution of the predicted pathline variance [Eq. (8)] in the continuum model and the corresponding sample variance [Eq. (10)]. Continuum and discrete results compare very well for $\rho_x=0$ and 0.1. However, the two do not match satisfactorily when the

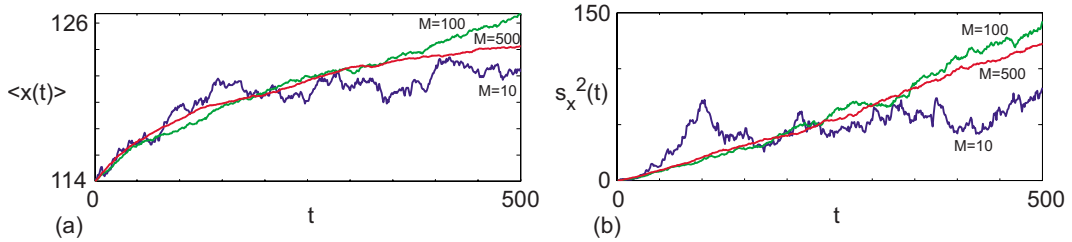

FIG. 2. (Color online) Unbiased pathline simulation data corresponding to the results in the left column of Fig. 1 show the convergence behavior with respect to M, the number of Monte Carlo realizations. The evolution of (a) the agent position and (b) the variance are given for averages over $M=10$ (blue), $M=100$ (green), and $M=500$ (red) simulations. All results correspond to $\Delta=\tau=P=1$, $\rho_y=\rho_x=0$, and $[p_x(0),p_y(0)]=(114,10)$.

horizontal bias is increased to $\rho_x=0.5$. The two match reasonably well for the initial part of the simulation where $t\leq 50$, but at later times the continuum result underestimates the simulation data. We found that increasing the number of Monte Carlo realizations did not improve this comparison. We note again that the bias parameter (involved in the advective velocity) does not explicitly appear in the evolution of $\sigma_x^2(t)$ [Eq. (8)]. However, there is an implicit influence of the bias parameter through the population density C in Eq. (8). This is an intriguing result given that the continuum-discrete comparison of the density profiles [Fig. 1 row (c)] and the first moment [Fig. 1 row (d)] matches very well regardless of the value of the bias parameter. Further results (not shown) indicate that Eq. (8) is valid for the problem reported in Fig. 1 for $|\rho_x|\leq 0.1$ and $t\leq 200$.

Additional continuum and discrete pathline data with the same initial distribution of agents but different values of $[p_x(0),p_y(0)]$ to Fig. 1 were also generated. Similarly, further pathline data for different initial distributions of agents also gave comparable results. Comparing the continuum and discrete results for the averaged density profiles and mean position gave qualitatively similar results to those in Figs. 1(c) and 1(d). The same trends with regard to the divergence between the continuum and the sample variance were also observed for these additional simulations.

We now discuss some practical issues associated with the interpretation of averaged simulation data. Simulation results in Fig. 1 were averaged over 100 Monte Carlo realizations for $0\leq t\leq 200$. These parameters were chosen to reflect the key properties of the simulation data.

Several additional issues are highlighted in Fig. 2. The trajectory data in Fig. 2 shows that the fluctuations in the averaged pathline $\langle x(t)\rangle$, and the sample variance $s_x^2(t)$ become smaller as the number of Monte Carlo simulations, M, increases. However, we see that the fluctuations diminish at different rates for the two types of data. In all cases considered in this work, the fluctuations in the averaged pathline data diminished faster than those for the sample variance as the number of Monte Carlo simulations increased. Since it is impractical to collect a very large number of experimental trajectories from identically prepared systems, it is reassuring that the fluctuations in the pathline data diminish faster than those for the variance data, as we are more likely to use pathline data to match with experimental results than variance data.

In all cases considered here, the quality of the match between the averaged simulation data and the continuum models decreased with time. This is why we chose to present results in Fig. 1 for 100 Monte Carlo realizations and $0\leq t\leq 200$. For larger intervals of time, we would need to increase the number of realizations in order to obtain a good match. This indicates that experimental pathline data may be best collected over several short periods of time rather than a single trajectory over a longer period of time.

IV. DISCUSSION AND CONCLUSION

The analysis of pathline data is a standard way to estimate macroscopic transport coefficients that describe the motility of populations. In cell biology, for example, pathline data is collected and analyzed to make inferences about whether cell motility is directed or undirected. This can be achieved by observing the behavior of a population of cells placed in a gradient of a signaling chemical [13,27].

In this work we present continuum equations describing the average pathlines in a population of agents undergoing a simple exclusion process. Unlike standard noninteracting random-walk models, the equations governing the pathlines in an exclusion process explicitly reflect the influence of agent-agent interactions since the evolution of the moments depends on an interaction with the average macroscopic behavior of the system. We note that the pathline equations presented here are consistent with previous simulation data for the simple case of a uniformly occupied lattice with no temporal or spatial gradients present in the system [6].

The theoretical pathline models developed here capture some key features that have been observed in the experimental cell biology literature. For example, both Ward *et al.* [13] and Cai *et al.* [28] studied pathline data *in vitro* and observed that pathlines were biased away from the bulk population. Similarly, Druckenbrod and Epstein [29] reported pathline data collected within an intact tissue culture. These data also described pathlines that were directed down the macroscopic cell density gradient. These observations are consistent with the pathline models developed here and point to the importance of capturing the correct cell-cell interactions that are ignored using a standard random-walk model [10,23].

In addition to presenting and validating continuum models for the position and variance of pathline data in an exclusion process, a key result of this work shows that the simulation data and the continuum model for the pathline variance diverge as the bias parameter, ρ_x, increases. There are a number of potential explanations for this. First, it is

possible that the deviation reflects the fact that the continuum equations are relevant in the limit where $\Delta, \tau \to 0$ simultaneously, with the requirement that the bias parameters must decrease to zero, with $\rho_x = O(\Delta)$ and $\rho_y = O(\Delta)$ [23]. Therefore, we may expect the continuum-discrete comparison to breakdown when ρ_x is no longer sufficiently small. Second, it is possible that the assumption of ignoring instantaneous correlations between occupancy states of adjacent sites becomes less appropriate as the bias parameters increase and that the continuum variance model is sensitive to the breakdown of this assumption. Third, the pathline variance predictions involve the neglect of temporal correlations which may become more important when the bias parameter increases. Of course, the divergence between the continuum and discrete variance could be caused by a combination of these issues.

ACKNOWLEDGMENTS

This work was supported by the Australian Research Council (ARC). We thank Ben Binder, Alana Moore, and Heather Verkade.

[1] T. M. Liggett, *Stochastic Interacting Systems: Contact, Voter, and Exclusion Processes* (Springer-Verlag, Berlin, 1999).
[2] P. Clifford and A. Sudbury, Biometrika **60**, 581 (1973).
[3] M. Eden, *Proceedings of the Fourth Berkeley Symposium*, 1961, Vol. 4, pp. 223–239.
[4] D. Chowdhury, A. Schadschneider, and K. Nishinari, Phys. Life Rev. **2**, 318 (2005).
[5] C. Rosello, P. Ballet, E. Planus, and P. Tracqui, Acta Biotheor. **52**, 343 (2004).
[6] M. J. Simpson, A. Merrifield, K. A. Landman, and B. D. Hughes, Phys. Rev. E **76**, 021918 (2007).
[7] R. Lipowsky, S. Klumpp, and T. M. Nieuwenhuizen, Phys. Rev. Lett. **87**, 108101 (2001).
[8] T. Callaghan, E. Khain, L. M. Sander, and R. M. Ziff, J. Stat. Phys. **122**, 909 (2006).
[9] L. M. Sander and T. S. Deisboeck, Phys. Rev. E **66**, 051901 (2002).
[10] H. C. Berg, *Random Walks in Biology*, Expanded ed. (Princeton University Press, Princeton, NJ, 1983).
[11] L. F. Costa, L. C. Cintra, and D. Schubert, Cytometry A **68A**, 92 (2005).
[12] W. Korohoda, J. Golda, J. Sroka, A. Wojnarowicz, P. Jochym, and Z. Madeja, Cell Motil. Cytoskeleton **38**, 38 (1997).
[13] M. Ward, C. McCann, M. DeWulf, J. Y. Wu, and Y. Rao, J. Neurosci. **23**, 5170 (2003).
[14] C. M. Witt, S. Raychaudhuri, B. Schaefer, A. K. Chakraborty, and E. A. Robey, PLoS Biol. **3**, 1062 (2005).
[15] H. G. Othmer, S. R. Dunbar, and W. Alt, J. Math. Biol. **26**, 263 (1988).
[16] M. A. Rivero, R. T. Tranquillo, H. M. Buettner, and D. A. Lauffenburger, Chem. Eng. Sci. **44**, 2881 (1989).
[17] D. H. Tanaka, M. Yanagida, Y. Zhu, S. Mikami, T. Nagasawa, J.-i. Miyazaki, Y. Yanagawa, K. Obata, and F. Murakami, J. Neurosci. **29**, 1300 (2009).
[18] F. Spitzer, Adv. Math. **5**, 246 (1970).
[19] M. R. Evans, R. Juhasz, and L. Santen, Phys. Rev. E **68**, 026117 (2003).
[20] G. M. Schütz and E. Domany, J. Stat. Phys. **72**, 277 (1993).
[21] G. M. Schütz, J. Stat. Phys. **88**, 427 (1997).
[22] M. J. Simpson, K. A. Landman, and B. D. Hughes, Physica A **388**, 399 (2009).
[23] B. D. Hughes, *Random Walks and Random Environments* (Oxford University Press, Oxford, UK, 1995), Vol. 1.
[24] G. B. Whitham, *Linear and Nonlinear Waves* (Wiley, New York, 1974).
[25] J. Bear, *Dynamics of Fluids in Porous Media* (Dover, New York, 1972).
[26] S. C. Chapra and R. P. Canale, *Numerical Methods for Engineers* (McGraw-Hill, Singapore, 1998).
[27] M. H. Young, C. J. Hearn, P. G. Farlie, A. J. Canty, P. Q. Thomas, and D. F. Newgreen, Dev. Biol. **229**, 503 (2001).
[28] A. Q. Cai, K. A. Landman, and B. D. Hughes, J. Theor. Biol. **245**, 576 (2007).
[29] N. R. Druckenbrod and M. L. Epstein, Dev. Dyn. **236**, 84 (2007).