Macroscopic limits of individual-based models for motile cell populations with volume exclusion

Louise Dyson, Philip K. Maini, and Ruth E. Baker

Centre for Mathematical Biology, Mathematical Institute, University of Oxford, 24-29 St. Giles', Oxford OX1 3LB, United Kingdom (Received 29 February 2012; published 5 September 2012)

Partial differential equation models are ubiquitous in studies of motile cell populations, giving a phenomenological description of events which can be analyzed and simulated using a wide range of existing tools. However, these models are seldom derived from individual cell behaviors and so it is difficult to accurately include biological hypotheses on this spatial scale. Moreover, studies which do attempt to link individual- and population-level behavior generally employ lattice-based frameworks in which the artifacts of lattice choice at the population level are unclear. In this work we derive limiting population-level descriptions of a motile cell population from an off-lattice, individual-based model (IBM) and investigate the effects of volume exclusion on the population-level dynamics. While motility with excluded volume in on-lattice IBMs can be accurately described by Fickian diffusion, we demonstrate that this is not the case off lattice. We show that the balance between two key parameters in the IBM (the distance moved in one step and the radius of an individual) determines whether volume exclusion results in enhanced or slowed diffusion. The magnitude of this effect is shown to increase with the number of cells and the rate of their movement. The method we describe is extendable to higher-dimensional and more complex systems and thereby provides a framework for deriving biologically realistic, continuum descriptions of motile populations.

DOI: 10.1103/PhysRevE.86.031903

PACS number(s): 87.10.-e, 87.17.Aa, 87.17.Jj, 87.18.Gh

I. INTRODUCTION

The movement of discrete individuals plays an important role in many different systems, from the rearrangement of cells during embryonic development [1-3] to the emergency evacuation of crowds [4]. Continuum limits of these systems can be derived from lattice-based cellular automaton models, thus validating the widespread use of partial differential equations (PDEs) to model diffusion, chemotaxis, etc. [5–12]. Moreover, exclusion processes [10-17], where at most one cell occupies each lattice site, naturally encode volume exclusion, which may be fundamental, particularly where population pressure is hypothesized as a mechanism for invasion. However, the dangers of lattice artifacts [11,18] and a restricted ability to accurately describe physical processes suggests that off-lattice approaches are more appropriate. Until now their use has been limited, as both analytical and numerical exploration is more involved, particularly when the finite volume of individuals is included [19].

Here we derive a continuum limit of an off-lattice individual-based model (IBM) with volume exclusion. We show that while an on-lattice random walk with exclusion may be well-described by linear diffusion, this is not true off lattice. The distance an individual moves and the space it occupies together determine the deviation from simple Fickian diffusion. We demonstrate the accuracy of our new description, for a broad range of parameter values, using averaged realizations of the underlying IBM. Our method is applicable to a wide range of interaction mechanisms and in higher spatial dimensions, enabling the systematic derivation of PDE models from a description of the underlying individual behaviors.

A. Model description

We consider a one-dimensional domain with boundaries at B_L and B_R . Each individual is defined by a position x and occupies an interval (x - R, x + R), where R is the analog of cell

radius in two dimensions. Individuals move by hopping with rate α a distance d > 0 to the left or right. An attempted move is aborted if it requires moving through any point that is already occupied by another individual. For example, a move to the right from position x is aborted if there is another cell center in the interval [x + 2R, x + d + 2R). The abortion of attempted moves is a distinguishing factor from the modeling of inanimate particles, which is also well studied [20]. Here, we make the biologically realistic assumption that individuals are able to "sense" the presence of others by using filopodia, for example. If an individual attempts to move out of the domain, it is reflected back from the boundary, so that the cell center remains in the domain. The model is shown schematically in Fig. 1.

II. DERIVATION OF CONTINUUM EQUATIONS

Let $C_i(x,t)$ be the probability density function (pdf) for the *i*th cell center position at time *t*. For $\Delta t \ll 1$,

$$C_{i}(x,t + \Delta t)$$

$$= C_{i}(x,t) \left\{ 1 - \alpha \Delta t + \frac{\alpha \Delta t}{2} \left[P_{L}^{i}(x,t) + P_{R}^{i}(x,t) \right] \right\}$$

$$+ C_{i}(x - d,t) \frac{\alpha \Delta t}{2} \left[1 - P_{R}^{i}(x - d,t) \right]$$

$$+ C_{i}(x + d,t) \frac{\alpha \Delta t}{2} \left[1 - P_{L}^{i}(x + d,t) \right] + \mathcal{O}(\Delta t^{2}), \quad (1)$$

where $P_L^i(x,t)$ and $P_R^i(x,t)$ are the probabilities of a cell, other than cell *i*, being present in the regions (x - d, x] and [x, x + d), respectively:

$$P_{R}^{i}(x,t) = \sum_{j \neq i} \int_{2R}^{2R+d} C_{j}(x+\bar{x},t)d\bar{x};$$
 (2)

$$P_L^i(x,t) = \sum_{j \neq i} \int_{-2R-d}^{-2R} C_j(x+\bar{x},t) d\bar{x}.$$
 (3)



FIG. 1. Schematic view of the IBM, showing individuals in one spatial dimension. Each individual is represented by its center point, x, and occupies a region around this point, (x - R, x + R). Individuals move with rate α , a distance d to the left or right.

For now we assume d < 4R so that at most one individual may be present in the intervals considered above. In this derivation, we are making the usual moment-closure approximation, that the probability of two cells, *i* and *j*, being at *x* and *y*, respectively, is $C_i(x,t)C_j(y,t)$, the product of the individual probabilities [21–23]. This assumes independence between the two cells and is generally a valid assumption for on-lattice models, particularly in higher spatial dimensions and in the absence of mechanisms, such as proliferation, which enhance correlations between adjacent sites [5]. If 2R + d is small compared to the length scale on which *C* changes then we can expand $P_L^i(x,t)$ and $P_R^i(x,t)$ in a Taylor series to obtain

$$P_R^i(x,t) = \sum_{j \neq i} \left\{ dC_j + \frac{d}{2} (4R+d) \frac{\partial C_j}{\partial x} + \frac{d}{6} (12R^2 + 6Rd+d^3) \frac{\partial^2 C_j}{\partial x^2} + \mathcal{O}[(2R+d)^4] \right\},$$
(4)

$$P_{L}^{i}(x,t) = \sum_{j \neq i} \left\{ dC_{j} - \frac{d}{2}(4R+d)\frac{\partial C_{j}}{\partial x} + \frac{d}{6}(12R^{2} + 6Rd+d^{3})\frac{\partial^{2}C_{j}}{\partial x^{2}} + \mathcal{O}[(2R+d)^{4}] \right\}.$$
 (5)

Taking the limit as $\Delta t \rightarrow 0$, we find that

$$\frac{\partial C_i}{\partial t} = \frac{\alpha d^2}{2} \frac{\partial^2 C_i}{\partial x^2} + \frac{\alpha d^2}{2} (4R - d) \frac{\partial}{\partial x} \left(C_i \sum_{j \neq i} \frac{\partial C_j}{\partial x} \right) + \frac{\alpha d^2}{2} \mathcal{O}[(R + d)^2].$$
(6)

If the initial conditions are the same for all cells (i.e., each cell has an initial position chosen at random from the same distribution), $C_i(x,t) = C_j(x,t) \forall i, j$. Hence, $C = \sum_{i=1}^N C_i$, the average total cell density, satisfies

$$\frac{\partial C}{\partial t} = \frac{\alpha d^2}{2} \frac{\partial}{\partial x} \left\{ \left[1 + (4R - d) \frac{(N-1)}{N} C \right] \frac{\partial C}{\partial x} \right\} + \frac{\alpha d^2}{2} \frac{(N-1)}{N} \mathcal{O}[(R+d)^2],$$
(7)

where *N* is the number of cells in the domain. Taking the limit as $d \rightarrow 0$,

$$\frac{\partial C}{\partial t} = \hat{\alpha} \frac{\partial}{\partial x} \left\{ \left[1 + 4R \frac{(N-1)}{N} C \right] \frac{\partial C}{\partial x} \right\} + \frac{(N-1)}{N} \mathcal{O}(R^2),$$
(8)

with $\hat{\alpha} = \lim_{d\to 0} \alpha d^2/2$ held constant [24]. Note that in Eqs. (7) and (8) 4*RC* may be of order 1, since *C* is in the range

[0,1/(2R)]. Hence, the second term in the expansion may be highly important in correctly modeling cell movement.

A. Initial and boundary conditions

The condition that individuals must begin with the same initial distribution may seem restrictive but this is not the case; we only consider the average distribution of cells over multiple realizations, in a similar way to [10,26]. Appropriate boundary conditions must be considered carefully, since the center of a cell cannot be within a distance R of the boundary. In essence, we have no-flux conditions at $B_L + R$ and $B_R - R$, where B_L and B_R are the left and right boundaries, respectively. This neglects effects close to the boundary, where P_L^i and P_R^i may be different compared to the interior of the domain, and for now we simply choose a domain large enough to ensure cells do not reach the boundaries.

B. Limiting equations

Equation (8) gives the limiting continuum model for finitesized individuals moving with diffusivity $\hat{\alpha}$. To examine the accuracy of our derived equations under different parameter regimes, however, simulations of the IBM must be carried out with d > 0, and so it may be more appropriate to use Eq. (7). In this case, if [1 + (4R - d)(N - 1)/N C(x,t)] < 0initially for any x, then Eq. (7) is ill posed, whereas Eq. (8) is well posed for all possible initial conditions. This reflects the assumption in Eqs. (2) and (3) that there can be at most one cell in the interval [x + 2R, x + 2R + d), which is not satisfied if $d \ge 4R$. However, when $d \ge 4R$, Eqs. (2) and (3) are still valid first-order approximations for $C_i \ll 1$ since, using



FIG. 2. (Color online) Comparison between the IBM and the continuum models with d > 0 [Eq. (7)] and as $d \rightarrow 0$ [Eq. (8)], plotted at t = 0, 200, and 600 for $x \in [0,65]$. The solution to the diffusion equation is also shown for comparison. Realizations of the IBM were performed using the Gillespie algorithm [25]. Sixty cells were placed at regular intervals in the middle of the domain, with the leftmost cell given an initial position drawn from the normal distribution $\mathcal{N}(25, 1)$. For the PDEs we used a finite difference method with initial conditions determined by the average initial distribution from simulations, linearly interpolated onto a mesh with spacing dx = 0.05. Parameters: $\alpha = 2.222$, R = 0.1, and d = 0.15, giving $\hat{\alpha} = 0.025$.



FIG. 3. (Color online) Comparison between the IBM and the continuum model with d > 0 [Eq. (7)], plotted at t = 0,300,600 for $x \in [0,65]$. The solution to the diffusion equation is also shown for comparison. Realizations of the IBM were performed using the Gillespie algorithm [25]. Two groups of 30 cells were placed at regular intervals giving a density of 80%. The leftmost cells in the first and second groups are given an initial position drawn from $\mathcal{N}(16,1)$ and $\mathcal{N}(38,1)$, respectively. For the PDEs we used a finite difference method with initial conditions determined by the average initial distribution from simulations, linearly interpolated onto a mesh with spacing dx = 0.1. Parameters: $\alpha = 0.8$, R = 0.15, and d = 0.25, giving $\hat{\alpha} = 0.025$.

moment-closure assumptions:

$$\begin{split} P_{R}^{i}(x,t) &= \int_{2R}^{2R+d} \left\{ 1 - \prod_{j \neq i} [1 - C_{j}(x + \bar{x}, t)] \right\} d\bar{x}, \\ &= \int_{2R}^{2R+d} \left[\sum_{j \neq i} C_{j}(x + \bar{x}, t) - \sum_{k \neq j \neq i} C_{j}(x + \bar{x}, t) C_{k}(x + \bar{x}, t) + \mathcal{O}\left(C_{j}^{3}\right) \right] d\bar{x}. \end{split}$$

It is interesting to consider the case where [1 + (4R - d)(N - 1)/N C(x,t)] < 0, since a diffusion equation with a diffusion coefficient that can become negative for sufficiently high cell densities is often used to model cellular aggregation. Aggregation then gives rise to cell clustering [27,28], and there is evidence that this may indeed occur in some biological situations, such as in glioma cell invasion [17,29–32]. However, aggregation is difficult to assess in one dimension: Cells cannot move past each other in our model so that clusters tend to form regardless of the sign of [1 + (4R - d)(N - 1)/N C(x,t)]. Hence, a full consideration of clustering requires extension of the model to two or three dimensions and, as such, is beyond the scope of this paper.

III. RESULTS

Figures 2 and 3 demonstrate the very good agreement between averaged realizations of the IBM and our derived models. In particular, there is a much better correspondence with either Eqs. (7) or (8) than with the linear diffusion equation. Thus, we emphasize the dangers of naively using linear diffusion to model motility in crowded situations.

The differences between Eqs. (7) and (8) become more apparent when we consider the case where R = 0, so that the cells are point particles; that is, they have no volume. In the case where $d \rightarrow 0$, taking R = 0 reduces Eq. (8) to the diffusion equation. In Eq. (7), however, we do not recover the diffusion equation since cells cannot move past each other, a form of contact inhibition.

A. Exploring parameter space

As the number of cells, N, or the space occupied by an individual cell, 2R, increases, the domain becomes more crowded. In this case C(x,t) will also, in general, increase, since $\int C dx = N$ [note that, while C_i is a probability density function, C(x,t) is not]. Hence, the excluded volume term in Eqs. (7) and (8) exerts more influence relative to the linear diffusion term. For fixed d, however, the difference between Eq. (7) and the linear diffusion equation is minimal when



FIG. 4. Comparison of the varying importance of the exclusion term in Eq. (7). (a) Plot of $\int_{B_L}^{B_R} |C - C_D| dx/(B_R - B_L)$, where C_D is the solution to the diffusion equation, as R and N change, with d = 0.15. (b) Plot of $\int_{B_L}^{B_R} |S - C_D| dx/(B_R - B_L)$, where C_D is the solution to the diffusion equation and S is the averaged simulation results, as R and d change. In each case, $\alpha = 2.222$, ten cells were placed on the interval [-40,40], with initial positions drawn from the normal distribution $\mathcal{N}(0,3)$.



FIG. 5. (Color online) Comparison between the IBM and the continuum models with $d \sim \mathcal{N}(0, \sigma^2)$ [Eq. (11)] and as $\sigma \to 0$ [Eq. (8)], plotted at t = 0, 200, and 600 for $x \in [0,65]$. The solution to the diffusion equation is also shown for comparison. Realizations of the IBM were performed using the Gillespie algorithm [25]. Sixty cells were placed at regular intervals in the middle of the domain, with the leftmost cell given an initial condition drawn from the normal distribution $\mathcal{N}(25, 1)$. For the PDEs we used a finite difference method with initial positions determined by the average initial distribution from simulations, linearly interpolated onto a mesh with spacing dx = 0.05. Parameters: $\alpha = 2.222$, R = 0.1, and $\sigma = 0.1$.

R = d/4, so that as R decreases through this value, the importance of the exclusion effects increases again [Fig. 4(b)].

We note also that the error of our continuum approximations depends sensitively, and nonlinearly, on R, d, and N. For example, as N increases from a few individuals to many, crowding effects become more important due to the factor of (N - 1)/N, but at the same time the error also increases. Note that as $N \to \infty$, we may use $(N - 1)/N \approx 1$. Counterintuitively, decreasing the jump distance, d, can increase the effects of volume exclusion if R > d/4, since the correction to the diffusion equation is multiplied by (4R - d). This represents the two opposing effects of exclusion: Individuals move less often, but a higher proportion of these moves will be into unoccupied areas. As the distance moved increases past d = 4R, the higher bias towards empty spaces is counteracted by the cessation of movement. This effect can be seen when we compare the difference between results from averaged realizations and the simple diffusion equation [Fig. 4(a)]. Different movement mechanisms may not give rise to this effect. For example, if instead of aborting the attempted movement the cell moved next to the obstructing cell, then this effect may be reduced.

B. Extending to normally distributed distance moved

We may easily extend our method to consider variable jump distances. With the pdf of d given by $f(u_d)$, for $u_d \in \mathbb{R}$, we have

$$\begin{split} \frac{\partial C_i}{\partial t} &= \alpha \int_{-\infty}^{\infty} f(u_d) \bigg[-u_d \frac{\partial C_i}{\partial x} + \frac{u_d^2}{2} \frac{\partial^2 C_i}{\partial x^2} \bigg] du_d \\ &+ \alpha \sum_{j \neq i} \int_0^{\infty} \bigg[f(u_d) \bigg(C_i(x,t) \int_{2R}^{2R+u_d} C_j(x+\bar{x},t) d\bar{x} \\ &- C_i(x-u_d,t) \int_{2R-u_d}^{2R} C_j(x+\bar{x},t) d\bar{x} \bigg) \\ &+ f(-u_d) \bigg(C_i(x,t) \int_{-2R-u_d}^{-2R} C_j(x+\bar{x},t) d\bar{x} \\ &- C_i(x+u_d,t) \int_{-2R}^{-2R+u_d} C_j(x+\bar{x},t) d\bar{x} \bigg) \bigg] du_d. \end{split}$$

If the jump direction is unbiased, $f(u_d) = f(-u_d)$,

$$\frac{\partial C_i}{\partial t} = \frac{\alpha \langle d^2 \rangle}{2} \frac{\partial^2 C_i}{\partial x^2} + \alpha (N-1) \bigg(2R \langle d^2 \rangle - \int_0^\infty f(u_d) u_d^3 du_d \bigg) \frac{\partial}{\partial x} \bigg(C_i \sum_{j \neq i} \frac{\partial C_j}{\partial x} \bigg).$$
(9)



FIG. 6. Comparison of the error between the IBM and the continuum model with (a) fixed jump distance, d, [Eq. (8)] and (b) $d \sim \mathcal{N}(0,\sigma)$, [Eq. (11)]. In each case, $\alpha = 2.222$, ten cells were placed in the interval [-40,40], with initial positions drawn from the normal distribution $\mathcal{N}(0,3)$.

For example, if $d \sim \mathcal{N}(0,\sigma^2)$, $\int_0^\infty f(u_d) u_d^3 du_d = \sigma^3 \sqrt{2/\pi}$ and $\langle d^2 \rangle = \sigma^2$, so that

$$\frac{\partial C_i}{\partial t} = \frac{\alpha \sigma^2}{2} \frac{\partial^2 C_i}{\partial x^2} + \alpha \left(2R\sigma^2 - \sigma^3 \sqrt{\frac{2}{\pi}} \right) \frac{\partial}{\partial x} \left(C_i \sum_{j \neq i} \frac{\partial C_j}{\partial x} \right).$$
(10)

As before, if the initial distribution is the same for all C_i , then $C = \sum_{i=1}^{N} C_i$ satisfies

$$\frac{\partial C}{\partial t} = \frac{\alpha \sigma^2}{2} \frac{\partial^2 C}{\partial x^2} + \frac{\alpha \sigma^2}{2} \frac{N-1}{N} \left(4R - 2\sigma \sqrt{\frac{2}{\pi}} \right) \frac{\partial}{\partial x} \left(C \frac{\partial C}{\partial x} \right).$$
(11)

Numerical solutions of this model also compare well with averaged realizations of the IBM (Fig. 5) for a range of σ , *R*, and α . Since the assumption that individuals do not try to move further than 4*R* in a jump is less often violated, even for large σ and small *R*, the error between averaged realizations and Eq. (11) is smaller than with constant *d* (Fig. 6). This model is closer to the classic depiction of Brownian dynamics, with [19,33] or without [24] hard core exclusion effects. Note that taking $\sigma \rightarrow 0$ with $\alpha \sigma^2/2$ held constant gives Eq. (8).

- [1] P. M. Kulesa et al., Dev. Biol. 316, 275 (2008).
- [2] B. J. Binder et al., Bull. Math. Biol. 74, 474 (2012).
- [3] J. Tanimoto, A. Hagishima, and Y. Tanaka, Physica A 389, 5611 (2010).
- [4] D. Helbing, I. Farkas, and T. Vicsek, Nature (London) 407, 487 (2000).
- [5] R. E. Baker and M. J. Simpson, Phys. Rev. E 82, 041905 (2010).
- [6] B. J. Binder, K. A. Landman, M. J. Simpson, M. Mariani, and D. F. Newgreen, Phys. Rev. E 78, 031912 (2008).
- [7] S. Muhuri, L. Shagolsem, and M. Rao, Phys. Rev. E 84, 031921 (2011).
- [8] M. Alber, N. Chen, P. M. Lushnikov, and S. A. Newman, Phys. Rev. Lett. 99, 168102 (2007).
- [9] H. G. Othmer and A. Stevens, SIAM J. Appl. Math. 57, 1044 (1997).
- [10] M. J. Simpson, R. E. Baker, and S. W. McCue, Phys. Rev. E 83, 021901 (2011).
- [11] C. J. Penington, B. D. Hughes, and K. A. Landman, Phys. Rev. E 84, 041120 (2011).
- [12] P. M. Lushnikov, N. Chen, and M. Alber, Phys. Rev. E 78, 061904 (2008).
- [13] B. J. Binder and K. A. Landman, J. Theor. Biol. 259, 541 (2009).
- [14] B. J. Binder and K. A. Landman, Phys. Rev. E 83, 041914 (2011).
- [15] S. Cannas, Math. Biosci. 183, 93 (2003).
- [16] A. E. Fernando, K. A. Landman, and M. J. Simpson, Phys. Rev. E 81, 011903 (2010).

IV. CONCLUSIONS

In summary, we have developed a simple, yet effective, method for deriving macroscopic equations from off-lattice IBMs with excluded volume effects. This method may be extended naturally to higher spatial dimensions and more complicated mechanisms of movement and, as such, provides a natural framework for multiscale modeling of motile populations. Our models accurately predict the average behavior of our IBMs for arbitrary initial conditions and a range of parameter values. We have demonstrated the errors arising from a naive choice of linear diffusion, and shown how an informed choice can be made about whether a motile population requires detailed consideration beyond Fickian diffusion. Counterintuitively, we have shown that increasing the size of cells does not necessarily lead to a larger deviation from diffusion due to exclusion effects. Instead, the magnitude of the effect of exclusion depends on the relative sizes of cell jumps and radius. Our results suggest that modeling of a large number of finite-sized individuals requires consideration of the exclusion effects in order to accurately capture populationlevel behavior.

ACKNOWLEDGMENT

L.D. was funded by the EPSRC through the Oxford Systems Biology Doctoral Training Centre.

- [17] C. Deroulers, M. Aubert, M. Badoual, and B. Grammaticos, Phys. Rev. E 79, 031917 (2009).
- [18] A. Flache and R. Hegselmann, J. Artif. Soc. Soc. Simul. 4, 4 (2001); http://jasss.soc.surrey.ac.uk/4/4/6.html.
- [19] M. Bruna and S. J. Chapman, Phys. Rev. E 85, 011103 (2012).
- [20] P. Résibois and M. De Leener, *Classical Kinetic Theory of Fluids* (Wiley, New York, 1977).
- [21] C. S. Gillespie, IET Syst. Biol. 3, 52 (2009).
- [22] R. E. Baker, C. A. Yates, and R. Erban, Bull. Math. Biol. 72, 719 (2010).
- [23] R. Erban and H. G. Othmer, SIAM J. Appl. Math. 65, 361 (2004).
- [24] E. A. Codling, M. J. Plank, and S. Benhamou, J. R. Soc., Interface 5, 813 (2008).
- [25] D. T. Gillespie, J. Phys. Chem. 81, 2340 (1977).
- [26] M. J. Simpson et al., Phys. A 389, 1412 (2010).
- [27] M. J. Simpson, K. A. Landman, and B. D. Hughes, Physica A 389, 3779 (2010).
- [28] E. Khain and L. M. Sander, Phys. Rev. E 77, 051129 (2008).
- [29] E. Khain, L. M. Sander, and C. M. Schneider-Mizell, J. Stat. Phys. 128, 209 (2006).
- [30] E. Khain, C. M. Schneider-Mizell, M. O. Nowicki, E. a. Chiocca, S. E. Lawler, and L. M. Sander, Europhys. Lett. 88, 28006 (2009).
- [31] E. Khain, M. Katakowski, S. Hopkins, A. Szalad, X. Zheng, F. Jiang, and M. Chopp, Phys. Rev. E 83, 031920 (2011).
- [32] E. Khain, M. Katakowski, N. Charteris, F. Jiang, and M. Chopp, Phys. Rev. E 86, 011904 (2012).
- [33] E. Barkai and R. Silbey, Phys. Rev. E 81, 041129 (2010).