

## Multiscale dynamics of biological cells with chemotactic interactions: From a discrete stochastic model to a continuous description

Mark Alber,<sup>1,\*</sup> Nan Chen,<sup>1</sup> Tilmann Glimm,<sup>2</sup> and Pavel M. Lushnikov<sup>1,3</sup>

<sup>1</sup>*Department of Mathematics, University of Notre Dame, Notre Dame, Indiana 46656, USA*

<sup>2</sup>*Department of Mathematics, Western Washington University, Bellingham, Washington 98225-9063, USA*

<sup>3</sup>*Landau Institute for Theoretical Physics, Kosygin Street 2, Moscow, 119334, Russia*

(Received 30 January 2006; published 1 May 2006)

The cellular Potts model (CPM) has been used for simulating various biological phenomena such as differential adhesion, fruiting body formation of the slime mold *Dictyostelium discoideum*, angiogenesis, cancer invasion, chondrogenesis in embryonic vertebrate limbs, and many others. We derive a continuous limit of a discrete one-dimensional CPM with the chemotactic interactions between cells in the form of a Fokker-Planck equation for the evolution of the cell probability density function. This equation is then reduced to the classical macroscopic Keller-Segel model. In particular, all coefficients of the Keller-Segel model are obtained from parameters of the CPM. Theoretical results are verified numerically by comparing Monte Carlo simulations for the CPM with numerics for the Keller-Segel model.

DOI: [10.1103/PhysRevE.73.051901](https://doi.org/10.1103/PhysRevE.73.051901)

PACS number(s): 87.18.Ed, 05.40.Ca, 05.65.+b, 87.18.Hf

### I. INTRODUCTION

Biological cell dynamics has been studied at two main scales of description. The macroscopic level provides one with a coarse-grained treatment of biological cells through their macroscopically averaged quantities such as local density of cells [1–4]. The macroscopic scale is large in comparison with the typical size of a cell. Macroscopic models are usually continuous and utilize families of differential or integro-differential equations to describe “fields” of interaction. A much more detailed approach is needed at the second, microscopic level which takes into account stochastic fluctuations of the shape of each individual cell.

Discrete models describe individual (microscopic) behaviors of cells. They are often applied to microscale events where a small number of elements can have a large (and stochastic) impact on a system. For example, while many periodic growth patterns can be modeled using continuous methods, patterns which depend sensitively on interaction between cells and media are best modeled with discrete methods. Simplest discrete models describe cells as pointwise objects. Some bacteria are self-propelled and do not change considerably their shape during motion (e.g., *Escherichia coli* [2,5] and *Myxococcus xanthus* [6,7] bacteria). They can be successfully represented as pointwise objects undergoing reorientation while moving [2,8,10]. In contrast, some other bacteria (e.g., *Dictyostelium discoideum* [11]) experience essential random fluctuations of their shapes and need to be treated as extended objects of variable shapes.

One of the microscopic models dealing with differential adhesion and shape fluctuations is a cellular Potts model (CPM) which is an extension of the well-known Potts model from statistical mechanics [12,13]. In this model each biological cell is represented by a cluster of pixels (spins). The CPM has been used to simulate various biological phenom-

ena such as cell sorting [12,13], fruiting body formation of the slime mold *Dictyostelium discoideum* [14,15], angiogenesis [16], cancer invasion [17], chondrogenesis in embryonic vertebrate limbs [18,19], and many others. (Different applications of the CPM have been reviewed in [20].) Recently an alternative model was suggested [21] which represents a cell as collection of subcellular elements which interact with each other through phenomenological intra-cellular and intercellular potentials.

In addition to short-range cell-cell adhesion and interactions between cells and their surrounding extracellular matrix, cells interact at long range through signal transmission and reception mediated by a diffusing chemical field (chemotaxis). The continuous macroscopic Keller-Segel model of the evolution of the density of cells with chemotactic interactions has been extensively studied [1–4] over the years. In particular, it has been successfully applied to the description of *Escherichia coli* bacteria aggregation due to chemotaxis in [2]. The drawback of continuous models is that they have a lower resolution than discrete models. However, their advantage is the availability of a large set of analytical and numerical tools for analyzing solutions of the corresponding nonlinear partial differential equations (PDE’s). By contrast, the analytical study of discrete models is often impossibly complicated and their computational implementation is often much less efficient in comparison with numerical methods available for PDE’s. It is thus important, for numerical, analytical, as well as conceptual reasons to establish connections between various discrete and continuous models of the same biological problem.

There is a vast literature on studying continuous limits of pointwise discrete microscopic models. In particular, the classical Keller-Segel model has been derived from a model with pointwise representation for cells undergoing random walk [8,22,23]. However, much less work has been done on deriving macroscopic limits of microscopic models which treat cells as extended objects. One of the first attempts at combining microscopic and macroscopic levels of descrip-

\*Corresponding author. Electronic address: malber@nd.edu

tion of cellular dynamics has been described in [24] where the diffusion coefficient for a collection of noninteracting randomly moving cells has been derived from a one-dimensional CPM. Recently a microscopic limit of subcellular elements model [21] was derived in the form of continuous advection-diffusion equation for cellular density. In the present paper, we establish a connection between a one-dimensional CPM of a cell moving in a media and reacting to a chemical field and a Fokker-Planck equation for the cell probability density function. This equation is then reduced to the classical macroscopic Keller-Segel equation. In particular, we derive all coefficients of the Keller-Segel model from parameters of the CPM. We also compare Monte Carlo [9] simulations for the CPM with numerics for the Keller-Segel model to support our theoretical results.

A unified multiscale approach, described in this paper and based on combining microscopic and macroscopic models, can be applied to studying such biological phenomena as streaming in *Dictyostelium discoideum*. In starved populations of *Dictyostelium* amoebae, cells produce and detect a communication chemical (cAMP). The movement of *Dictyostelium* cells changes from a random walk to a directed walk up the cAMP gradient resulting in formation of streams of cells towards the aggregation center [see Fig. 1(a)] and subsequent formation of multicellular fruiting body. Figure 1(b) shows cells' movement from left to right in response to waves of cAMP traveling through the aggregation stream from right to left. The cAMP gradient of the up-down direction is very small and could be ignored. Figure 1(c) schematically demonstrates the main features of the cell movement.

Unlike differential adhesion [12,13], chemotactic cell motion is highly organized over a length scale significantly larger than the size of a single cell. (For details about modeling *Dictyostelium discoideum* fruiting body formation see, e.g., [14,15,25,26].)

The paper is organized as follows. In Sec. II we describe a one-dimensional (1D) CPM with chemotaxis. In Sec. III, we derive from the Monte Carlo dynamics of the CPM the discrete master equation for the probability density function  $P(x, L, t)$ —that is, the probability that at time  $t$ , there is a cell whose length is  $L$  and whose center of mass is located at  $x$ . In Sec. IV we use the discrete master equation to derive a partial differential equation for  $P(x, L, t)$  in a continuous limit which assumes that cell changes its position and length at each Monte Carlo step by a small amount. We show that the dependence of  $P(x, L, t)$  on  $L$  is very close to the Boltzmann distribution. This is used in Sec. V for the derivation of a Fokker-Planck equation for the probability density function  $p(x, t)$  of a cell's center of mass being at  $x$  which is the main result of the paper. In Sec. VI it is shown that the addition of the time dependence of chemical field reduces the Fokker-Planck equation to the Keller-Segel equations. Sec. VII deals with numerical verification of the theoretical results of the previous sections and compares the Monte Carlo simulations for our CPM and Keller-Segel models.

## II. CELLULAR POTTS MODEL

The cellular Potts model, an extension of the Potts model from statistical mechanics, is a flexible and powerful way to

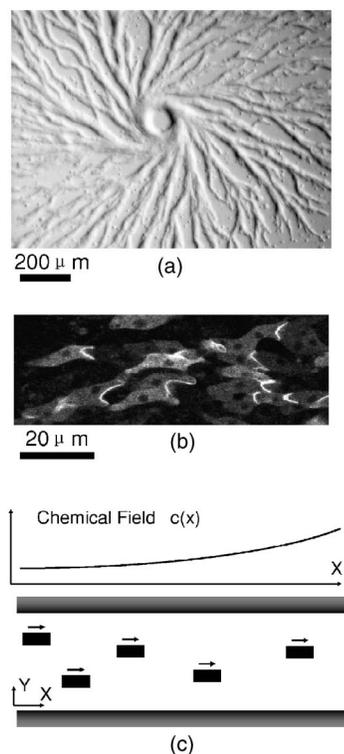


FIG. 1. (a) Streaming of *Dictyostelium discoideum* towards the aggregation center. Cells move chemotactically towards the aggregation center, leading to the formation of cell streams and finally mounds. (Reproduced from [27] with permission.) (b) Example of a quasi-one-dimensional motion of *Dictyostelium discoideum* inside a stream [this picture is on much smaller scale compared with (a)]. Cells are moving parallel to each other in the direction of chemical gradient (from left to right). Chemical gradient also causes polarization of cells so that they become elongated in the direction of a gradient. (Reproduced from [11] with permission.) (c) Schematic picture of cell motion in a gradient of chemical field (e.g., chemoattractant cAMP). The concentration of the chemical field is shown schematically above the main figure.

model cellular patterns. Its core mechanism is the competition between the minimization of various energy terms in some generalized functional of the cellular configuration—e.g., surface minimization, cell-cell contact and chemotactic interactions, and global geometric constraints. It simulates stochastic fluctuations of cell shapes as simple thermal fluctuations.

The CPM is defined on a rectangular lattice  $\mathcal{L}$ , which is of the form  $[0, m_x]$  (for one dimension),  $[0, m_x] \times [0, m_y]$  (for two dimensions) or  $[0, m_x] \times [0, m_y] \times [0, m_z]$  (for three dimensions). (Here  $[0, m] = \{0, 1, \dots, m\}$ .) The elements of  $\mathcal{L}$  are called the *lattice sites* (intervals in 1D, pixels in 2D, voxels in 3D). A lattice site is denoted by an index  $\mathbf{i} \in \mathcal{L}$ .

Each lattice site has an assigned “spin”  $\sigma(\mathbf{i})$  which can have values  $s=0, 1, \dots, Q$ , where  $s=0$  corresponds to absence of any cell at the given site and the value  $1 \leq s \leq Q$  means that the given site is occupied by the  $s$ th cell, where  $Q$  is the total number of cells in the system. Assume that we fix the values of  $\sigma(\mathbf{i})$  at each lattice site; then, we refer to that set of values as a *configuration*. The best way to visualize a

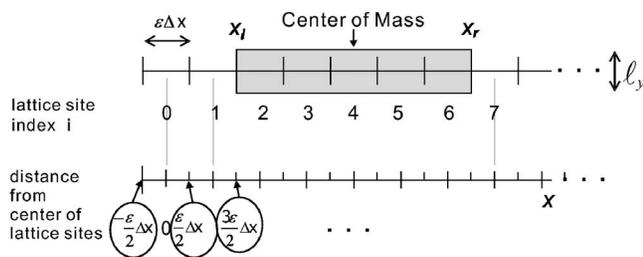


FIG. 2. Example of a cell in one-dimensional CPM. The cell (shaded domain) occupies lattice sites 2,...,6. It has a length of  $5\epsilon\Delta x$ , its center of mass is located at  $x=4\epsilon\Delta x$ , and its end points are  $x_l=1.5\epsilon\Delta x$  and  $x_r=6.5\epsilon\Delta x$ .

configuration is to regard the different spins as different colors. Each lattice site  $\mathbf{i}$  has a color  $\sigma(\mathbf{i})$ . The cells are the collections of lattice sites that have the same spin (color), so that each lattice site can be occupied by a single cell only. White color corresponds to absence of any cell at given site:  $\sigma(\mathbf{i})=0$ . In the model considered here we assume that cells cannot divide so that sites with the same color are always connected.

We assume periodic boundary conditions so that pixels at zero position in  $x$ ,  $y$ , or  $z$  are identical to sites with  $i_x=m_x+1$ ,  $i_y=m_y+1$ , and  $i_z=m_z+1$ , respectively. However, in our numerical simulation the cell never crossed the boundary, so there is no influence of our boundary conditions on result of simulations.

The temporal dynamics of the system is defined by certain probabilistic transition rules between the configurations, giving rise to a Markov chain of configurations—i.e., a sequence of configurations  $\sigma^0, \sigma^1, \sigma^2, \dots$ . To describe the transition rules, we associate to each configuration  $\sigma$  an energy  $E(\sigma)$ , also referred to as the Hamiltonian. The state changes from one configuration to the next are governed by an energy minimization principle with effective temperature  $T$ . This is implemented by means of the Metropolis algorithm for Monte Carlo Boltzmann dynamics [9]. The algorithm works as follows.

Given a configuration  $\sigma^n$ , we randomly select a lattice site  $\mathbf{i} \in \mathcal{L}$  such that not all of its nearest lattice neighbors have the same spin. We then randomly choose a lattice neighbor  $\mathbf{i}'$  of  $\mathbf{i}$  with  $\sigma^n(\mathbf{i}') \neq \sigma^n(\mathbf{i})$ . Let  $\sigma'$  be the configuration we obtain by “flipping” the spin of  $\mathbf{i}$ ; i.e., we have  $\sigma'(\mathbf{j}) = \sigma^n(\mathbf{j})$  for all  $\mathbf{j} \neq \mathbf{i}$  and  $\sigma'(\mathbf{i}) = \sigma^n(\mathbf{i}')$ . The new configuration  $\sigma^{n+1}$  is then either  $\sigma^n$  or the configuration  $\sigma'$ . The probability  $\Phi(\Delta E)$  that  $\sigma'$  is accepted as the next configuration  $\sigma^{n+1}$  depends on the energy difference  $\Delta E = E(\sigma') - E(\sigma^n)$ . The formula is

$$\Phi(\Delta E) = \begin{cases} 1, & \text{if } \Delta E \leq 0, \\ \exp(-\beta\Delta E), & \text{if } \Delta E > 0. \end{cases} \quad (1)$$

Here  $\beta=1/T$  is a positive constant (inverse effective temperature).

In this paper, we consider a quasi-one-dimensional CPM, which means that cells are assumed to move along  $x$  direction only and have fixed thickness  $l_y$  in the  $y$  direction (see Fig. 2). Let  $\epsilon\Delta x$  denote the size of lattice site, where  $0 < \epsilon \ll 1$ ,  $\epsilon$  is the small dimensionless constant, and  $\Delta x$  is a di-

mensional constant of the order of 1. Each lattice site is described by its index  $\mathbf{i}=0,1,\dots$ , so that the center of each lattice site is located at  $x=\mathbf{i}\epsilon\Delta x$  with the lattice site left border at  $x_l=(\mathbf{i}-\frac{1}{2})\epsilon\Delta x$  and the lattice site right border at  $x_r=(\mathbf{i}+\frac{1}{2})\epsilon\Delta x$  (see Fig. 2).

In what follows, we will consider the dynamics of a single cell so that the spin  $\sigma$  can take two values: 0 if cell is absent at a given site and 1 if cell occupies a given site. However, our results remain valid for an ensemble of  $n$  cells which are well separated from each other, so that the probability that two cells would try to occupy the same volume is negligible. This allows us to neglect cell-cell contact interactions. We assume that cells can interact only with the surrounding medium and the chemical field  $c(x)$  (chemotaxis). The chemical field is assumed to depend only on  $x$  but not on  $y$ . Cells can also produce a chemical which then diffuses. In Sec. VI we discuss production of chemicals by cells.

A natural biological realization of this quasi-one-dimensional model is the motion of biological cells in streams [26]. E.g., the amoeba *Dictyostelium discoideum* under starving condition typically forms streams [25]. The biological cells inside each stream are moving towards the aggregation center [see Fig. 1(a)], which results in complicated 2D patterns [26]. If we zoom in to a small scale, we will see that the motion of cells inside each stream is quasi one dimensional with cells moving parallel to each other in the  $x$  direction [Fig. 1(b)]. The chemical gradient of the other direction ( $y$  direction) could be neglected, and during cells movement there are no cell-cell interactions, such as cell collisions or cell signaling. Figure 1(c) schematically shows such a parallel motion of the cells from left to the right under the action of the gradient of a chemical field (chemoattraction).

For a given configuration  $\sigma$  of spins, let  $N=N(\sigma)$  denote the number of lattice sites that the cell occupies. The length of the cell is equal to  $L=N\epsilon\Delta x$ . We denote the position of the center of mass of the cell by  $x$  and denote the position of the left and right ends of the cell by  $x_l$  and  $x_r$ , respectively. Then  $L=x_r-x_l$ . (See Fig. 2.)

We assume that the chemical field  $c(x)$  is a slow function of time so its typical time scale is much bigger than the time step of a Monte Carlo algorithm. Then the Hamiltonian is given by the formula

$$E = J_{cm}(2L + 2\ell_y) + \lambda(L - L_T)^2 + \mu c(x)L. \quad (2)$$

The first term is a surface energy term which has contributions from both the energy of cell-medium adhesiveness and cell membrane elasticity, where  $J_{cm}$  is an interaction energy between the cell and medium per unit length. The second term is a length-constraint term which penalizes deviations of the cell length  $L$  from the target cell length  $L_T$ . Here  $\lambda$  is a positive constant. The choice of  $\lambda$  and  $\beta$  [see Eq. (1)] is determined by the typical scale of fluctuations of the cellular shape. The third term in Eq. (2) is the coupling chemical energy. This term will favor cell motion *down* or *up* the chemical gradient for  $\mu > 0$  and  $\mu < 0$ , respectively. We assume that the concentration  $c(x)$  is a slow function of  $x$  on a scale of the typical cell's length  $L$ :

$$x_c/L \gg 1, \quad (3)$$

where  $x_c$  is a typical scale for variation of  $c(x)$  in  $x$ . This is consistent with the generally accepted view that cells are typically too small to detect chemical gradients without moving. (See, e.g., [28]; however, recent experimental evidence may put this view in question [29].) Note that the chemical energy could also be defined as  $\mu \int_{x_l}^{x_r} c(x) dx$ . But in the limit (3), this is equivalent to the form used in the Hamiltonian (2).

### III. DISCRETE EVOLUTION EQUATION FOR PROBABILITY DENSITY FUNCTION

In this section, we develop an analytical model for the evolution of the stochastic dynamics of a cell in the CPM.

Let  $P(x, L, t)$  be a probability density for the cell with the center of mass at  $x$  of length  $L$  at time  $t$ . Spins  $\sigma(\mathbf{i})$  are defined on the lattice  $\mathcal{L}$  so that the length of the cell  $L$ , which is the difference between positions of right and left ends of a cell— $L = x_r - x_l$ —can take values  $n\varepsilon\Delta x$ ,  $n = 1, 2, \dots$ . The position of the center of mass  $x = (x_r + x_l)/2$  can take values  $n\varepsilon\Delta x/2$ ,  $n = 1, 2, \dots$ . That is, the CPM grid is twice the size of the grid of center of mass. In particular, if  $2\frac{x}{\varepsilon\Delta x}$  is an even number (i.e.,  $x$  coincides with one of the lattice sites) then the ratio  $\frac{L}{\varepsilon\Delta x}$  is also an even number. Alternatively, if  $2\frac{x}{\varepsilon\Delta x}$  is an odd number (i.e.,  $x$  coincides with a boundary between two neighboring lattice sites), then the ratio  $\frac{L}{\varepsilon\Delta x}$  is an odd number.

For convenience, we choose a normalization for  $P(x, L, t)$  such that the probability for a cell to have its center of mass at  $x$  and length  $L$  at time  $t$  is given by  $(\varepsilon\Delta x)^2 P(x, L, t)$ . The factor  $(\varepsilon\Delta x)^2$  results from the product of  $\varepsilon\Delta x/2$  (the spacing between lattice sites) and  $2\varepsilon\Delta x$  (the spacing in  $L$  for a fixed  $x$ ). With this normalization,  $P(x, L, t)$  becomes a true probability density in the continuous limit  $\varepsilon \rightarrow 0$ .

We choose the time interval between two Monte Carlo steps to be  $\varepsilon^2\Delta t$ , where  $\Delta t$  is a fixed constant of dimension of time. This implies diffusive time-space scaling

$$\frac{\varepsilon^2\Delta t}{(\varepsilon\Delta x)^2} = \frac{\Delta t}{(\Delta x)^2},$$

which is independent of the scaling parameter  $\varepsilon$ . We now switch from measuring time in Monte Carlo steps  $n = 0, 1, \dots$  to a continuous time variable  $t = n\varepsilon^2\Delta t$ .

Suppose at time  $t$  the cell is at a state  $(x, L)$  meaning that it has length  $L$  and its center of mass is at  $x$ . The stochastic discrete system at time  $t + \varepsilon^2\Delta t$  can switch to one of the following four possible states: (a)  $(x + \varepsilon\Delta x/2, L + \varepsilon\Delta x)$  by adding the lattice site  $x_r + \varepsilon\Delta x$  to the right end of cell, (b)  $(x + \varepsilon\Delta x/2, L - \varepsilon\Delta x)$  by taking away the site  $x_l$  from the left end of the cell, (c)  $(x - \varepsilon\Delta x/2, L + \varepsilon\Delta x)$  by adding the lattice site  $x_l + \varepsilon\Delta x$  to the left end of cell, and (d)  $(x - \varepsilon\Delta x/2, L - \varepsilon\Delta x)$  by taking away the site  $x_r$  from the right end of the cell.

Therefore, the most general master equation for evolution of the probability density  $P(x, L, t)$  has the form

$$\begin{aligned} P(x, L, t + \varepsilon^2\Delta t) = & \left[ 1 - T_l\left(x - \frac{\varepsilon}{2}\Delta x, L + \varepsilon\Delta x; x, L, t\right) - T_r\left(x + \frac{\varepsilon}{2}\Delta x, L + \varepsilon\Delta x; x, L, t\right) - T_l\left(x + \frac{\varepsilon}{2}\Delta x, L - \varepsilon\Delta x; x, L, t\right) \right. \\ & \left. - T_r\left(x - \frac{\varepsilon}{2}\Delta x, L - \varepsilon\Delta x; x, L, t\right) \right] P(x, L, t) + T_l\left(x, L; x + \frac{\varepsilon}{2}\Delta x, L - \varepsilon\Delta x, t\right) P\left(x + \frac{\varepsilon}{2}\Delta x, L - \varepsilon\Delta x, t\right) \\ & + T_r\left(x, L; x - \frac{\varepsilon}{2}\Delta x, L - \varepsilon\Delta x, t\right) P\left(x - \frac{\varepsilon}{2}\Delta x, L - \varepsilon\Delta x, t\right) + T_l\left(x, L; x - \frac{\varepsilon}{2}\Delta x, L + \varepsilon\Delta x, t\right) P\left(x - \frac{\varepsilon}{2}\Delta x, L + \varepsilon\Delta x, t\right) \\ & + T_r\left(x, L; x + \frac{\varepsilon}{2}\Delta x, L + \varepsilon\Delta x, t\right) P\left(x + \frac{\varepsilon}{2}\Delta x, L + \varepsilon\Delta x, t\right), \end{aligned} \quad (4)$$

where  $T_l(x, L; x', L')$  and  $T_r(x, L; x', L')$  correspond to transitional probabilities for a cell of length  $L'$  and center of mass at  $x'$  to change into a cell of length  $L$  and center of mass at  $x$ . Subscripts  $l$  and  $r$  correspond to a transition due to the addition (removal) of a pixel from the left (right) side of a cell, respectively. These transition probabilities are given by

$$T_l(x, L; x', L') = T_r(x, L; x', L') = \frac{1}{4} \Phi(E(x, L) - E(x', L')), \quad (5)$$

where  $E(x, L)$  is the Hamiltonian (2) and  $\Phi(\Delta E)$  is given by Eq. (1). The factor of 1/4 in Eq. (5) accounts for transitions

to four possible states (a)–(d). For computational purposes it is convenient to rewrite Eq. (1) in an equivalent form

$$\Phi(\Delta E) = 1 - \{1 - \exp[-\beta\Delta E]\} \Theta(\Delta E). \quad (6)$$

Here  $\Theta(x)$  is a Heaviside step function:  $\Theta(x) = 1$  for  $x > 0$  and  $\Theta(x) = 0$  for  $x < 0$ .

### IV. CONTINUOUS EVOLUTION EQUATION FOR PROBABILITY DENSITY FUNCTION OF THE CPM

Below we assume  $\varepsilon$  to be small,  $\varepsilon \ll 1$ , so that the change of the cell size and position is small at each Monte Carlo step. Now we carry out a Taylor series expansion in  $\varepsilon$  of the

terms in Eq. (4). One has to take special care of  $\Theta(\Delta E)$  terms in the expansion because the Heaviside step function is not analytic. To avoid this difficulty we do not expand the function itself but only its argument instead. There is an important simplification which comes from the fact that  $\Theta(\Delta E) + \Theta(-\Delta E) = 1$  so that in Eq. (4) we obtain that  $T_{l,r}(x, L; x', L', t) + T_{l,r}(x', L', t; x, L, t) = (1/4) \exp[-\beta|E(x, L) - E(x', L')|]$ . This yields a mutual cancellation of nonanalytical terms up to order  $O(\varepsilon^2)$ . Then, equating coefficients in the Taylor expansion in Eq. (4) in order  $O(\varepsilon^2)$  results in the Fokker-Planck equation

$$\partial_t P(x, L, t) = D(\partial_x^2 + 4\partial_L^2)P + 8D\beta\lambda\partial_L(\tilde{L}P) + D\beta L\mu\partial_x[c'(x)P],$$

$$\tilde{L} = \frac{1}{\lambda} \left[ J_{cm} + \lambda(L - L_T) + \frac{1}{2}\mu c(x) \right], \quad D = \frac{(\Delta x)^2}{8\Delta t}. \quad (7)$$

Now, under certain conditions to be described in the end of this section, the terms  $4D\partial_L^2 P + 8D\beta\lambda\partial_L(\tilde{L}P)$  dominate the other terms on the right-hand side of Eq. (7). This means that at the leading order, one can neglect terms with  $x$  derivatives. Under this assumption, the probability density function  $P(x, L, t)$  approaches a Boltzmann distribution for cell length exponentially in time at the rate of  $8D\beta\lambda$ :

$$P(x, L, t) = P_{Boltz}(x, L)p(x, t), \quad (8)$$

where  $p(x, t)$  is a probability density function of finding cell's center of mass at  $x$ .  $P_{Boltz}(x, L)$  is the Boltzmann distribution for the cell length given by

$$P_{Boltz}(x, L) = \frac{1}{Z} \exp(-\beta\Delta E_{length}), \quad (9)$$

$$\Delta E_{length} = E(L) - E_{min} = \lambda\tilde{L}^2, \quad (10)$$

where  $E_{min}$  is a minimum of energy  $E(L)$  as a function of  $L$  for a given  $x$ ,

$$E_{min} = E(L_{min}), \quad L_{min} = L_T - \frac{J_{cm}}{\lambda} - \frac{\mu c(x)}{2\lambda}, \quad (11)$$

and  $Z$  is a partition function

$$Z(x) = 2\varepsilon\Delta x \sum_{L=(1+\alpha)\varepsilon\Delta x, (3+\alpha)\varepsilon\Delta x, (5+\alpha)\varepsilon\Delta x, \dots} \exp(-\beta\Delta E_{length}),$$

$$\alpha = 1 \text{ for } \frac{x}{\varepsilon\Delta x} = n, \quad \alpha = 0 \text{ for } \frac{x}{\varepsilon\Delta x} = n + 1/2, \quad n \in \mathbb{N}.$$

(12)

Here we use the fact that due to discrete nature of our model, the position of the center of mass,  $x$ , could be located at one of the lattice sites  $x = m\varepsilon\Delta x$  ( $m$  being an integer number) if the length of the cell  $L$  is an even number of units  $\varepsilon\Delta x$  or  $x$  could be located at the boundary between two neighboring lattice sites in case of  $L$  being equal to an odd number of units of  $\varepsilon\Delta x$ . The factor  $(\varepsilon\Delta x)^2$  in the definition of the partition function (12) is chosen in such a way as to yield

$\int P(x, L, t)dLdx = 1$  in the continuous limit. We can also normalize  $\int P(x, L, t)dLdx = N$  to the total number of cells in the system  $N$ .

In the continuous limit  $\varepsilon \rightarrow 0$ , the sum in Eq. (12) is transformed into the integral

$$Z \approx \int_{-\infty}^{+\infty} \exp(-\beta\Delta E_{length})dL = \frac{\sqrt{\pi}}{\sqrt{\beta\lambda}}, \quad x \rightarrow 0. \quad (13)$$

Here we have extended the limits of integration from  $(0, +\infty)$  to  $(-\infty, +\infty)$ . Of course, physically, the length of the cell  $L$  is always positive. A typical fluctuation of the cell size  $\delta L = L - L_{min}$  about  $L_{min}$  is determined by the Boltzmann distribution (8) as  $\beta\lambda\delta L^2 \sim 1$ . In what follows we make a biologically motivated assumption about fluctuations of the cell size being much smaller than  $L$ :  $|\delta L| \ll L_{min}$  which results in the condition

$$\beta L_{min}^2 \lambda \gg 1. \quad (14)$$

This justifies the use of the integration limits  $(-\infty, +\infty)$  in Eq. (13) instead of  $(0, +\infty)$  because under this condition  $\exp(-\beta\Delta E_{length})$  peaks around  $L_{min}$  and replacement of integration limits results in an exponentially small correction.

Let us now specify the conditions for the applicability of the Boltzmann distribution approximation (8). For this, consider Eq. (7). We have  $\beta\lambda\delta L^2 \sim 1$ . We now assume in addition the relation

$$\beta x_0^2 \lambda \gg 1, \quad (15)$$

where  $x_0$  is a typical scale of  $P$  with respect to  $x$ . It follows from (15) that  $|\partial_x^2 P| \ll |4\partial_L^2 P|$ , and consequently, we may neglect the first term with the  $x$  derivative,  $\partial_x^2 P$ , on the right-hand side of Eq. (7).

The second condition for the applicability of the Boltzmann distribution approximation (8) is the assumption that the last term with a  $x$  derivative in Eq. (7) is small,  $|\beta L\mu\partial_x[c'(x)P]| \ll |4\partial_L^2 P|$ . This is true if

$$|L_{min}\mu c_0| \left(1 + \frac{x_c}{x_0}\right) \ll \lambda x_0^2, \quad (16)$$

where  $c_0$  is a typical amplitude of  $c(x)$  and  $x_c$  is a typical scale of variation of  $c(x)$  with respect to  $x$ . Last, recall that we derive the continuous equation (7) from the master equation (4) under the condition of the step in  $x$  being small:

$$\varepsilon \ll 1. \quad (17)$$

Notice that diffusion coefficient  $D$  in Eq. (7) does not depend on  $\beta$ . Instead  $\beta$  determines a rate of convergence,  $\tau_r^{-1} = 8D\beta\lambda$ , of  $P(x, L, t)$  to the Boltzmann distribution (8).

We have solved both the master equation (4) and its continuous limit (7) numerically with initial conditions  $P(x, L, 0)$  different from the Boltzmann distribution (4). The simulations described in Sec. VII demonstrate that for each  $x$ , the solution  $P(x, L, t)$  indeed converges in time to the Boltzmann distribution at an exponential rate of  $\sim 8D\beta\lambda$ .

### V. FOKKER-PLANCK EQUATION FOR THE PROBABILITY DENSITY FUNCTION $p(x,t)$

We now turn to calculating the probability density function  $p(x,t)$  of a cell's center of mass being at  $x$ . It is given by the sum over all possible lengths of a cell:

$$p(x,t) = 2\varepsilon\Delta x \sum_{L=(1+\alpha)\varepsilon\Delta x, (3+\alpha)\varepsilon\Delta x, (5+\alpha)\varepsilon\Delta x, \dots} P(x,L,t) \\ \simeq \int_{-\infty}^{+\infty} P(x,L,t)dL, \quad \varepsilon \rightarrow 0,$$

$$\alpha = 1 \text{ for } \frac{x}{\varepsilon\Delta x} = n, \quad \alpha = 0 \text{ for } \frac{x}{\varepsilon\Delta x} = n + 1/2, \quad n \in \mathbb{N}. \quad (18)$$

In the Boltzmann distribution approximation (8), Eq. (18) reduces to trivial condition of the normalization of total probability of the Boltzmann distribution to 1:  $\int_{-\infty}^{+\infty} P_{\text{Boltz}}(x,L,t)dL = 1$ .

To derive closed equation for  $p(x,t)$  we substitute the ansatz (8) into (7) and integrate both the right-hand and left-hand sides of Eq. (7) with respect to  $L$  to obtain

$$\partial_t p = D\partial_x^2 p - \partial_x[\chi(x)p\partial_x c(x)], \\ \chi(x) = \frac{D}{\lambda}\beta\mu \left[ J_{cm} - \lambda L_T + \frac{1}{2}\mu c(x) \right], \quad D = \frac{(\Delta x)^2}{8\Delta t}. \quad (19)$$

This continuous equation is the main result of this paper. The conditions for the applicability of Eq. (19) are given by Eqs. (14)–(17).

### VI. REDUCTION TO THE KELLER-SEGEL MODEL

In this section we add time dependence to the chemical field  $c$  (concentration of chemoattractant or chemorepellant) by including a diffusion equation with the source term  $ap$  which determines the secretion of chemical by a cell:

$$\partial_t c = D_c \partial_x^2 c - \gamma c + ap, \quad (20)$$

where  $D_c$  is a diffusion coefficient of the chemical field,  $\gamma$  is the decay rate of the chemical field, and  $a$  is a production rate of the chemical field.

The system of equations (19) and (20) is applicable under the assumption that the typical time scale  $\tau_c$  of diffusion of  $c(x,t)$ , given by  $\tau_c = \frac{\Delta x_c^2}{D_c}$ , is large in comparison with convergence time  $\tau_r = 1/(8D\beta\lambda)$  of  $P(x,L,t)$  to the Boltzmann distribution (8), where  $x_c$  is a typical spatial width of the distribution of  $c(x,t)$ . Namely, this condition has the form

$$\tau_c / \tau_r = 8D\beta\lambda\tau_c \gg 1. \quad (21)$$

Equations (19) and (20) form a closed set of equations which is equivalent to the classical Keller-Segel model [1] of chemotaxis. If the parameters satisfy condition

$$|J_{cm} - \lambda L_T| \gg \frac{1}{2}|\mu|c(x), \quad (22)$$

then Eq. (19) reduces to the following commonly used form of the Keller-Segel model [2,3]:

$$\partial_t p = D\partial_x^2 p - \chi_0 \partial_x [p\partial_x c], \\ \chi_0 = D\lambda\beta\mu [J_{cm} - \lambda L_T], \quad D = \frac{(\Delta x)^2}{8\Delta t}. \quad (23)$$

The probability density function  $p(x,t)$  corresponds to the microscopic density in the Keller-Segel model. Notice that in both the Keller-Segel model and CPM considered in this paper, there is no direct interaction between cells except through production and reaction to a chemoattractant. In other words, cells are treated in a way similar to a dilute gas with long-range nonlocal interactions due to reaction to a chemical field.

### VII. COMPARISON OF NUMERICAL SIMULATIONS

In this section, we describe numerical tests comparing Monte Carlo simulations of the CPM and simulations of both discrete and continuous models for the probability density functions  $P(x,L,t)$  and  $p(x,t)$ , as given by Eqs. (4), (7), and (19).

#### A. Monte Carlo simulations

The computation of the frequency distribution of the cell center of mass and length for the CPM has been carried out as follows.

(i) We run a large number  $N$  of CPM simulations with one cell with the same initial conditions.

(ii) We fix a time interval  $\delta t = \varepsilon^2 \Delta t$ ; i.e., we fix the time interval between successive Monte Carlo steps. For each simulation we record the locations of the center of mass and lengths of the cell at the times  $t = \delta t, 2\delta t, 3\delta t, \dots$ .

(iii) After the  $N$  runs, the recorded data give a frequency distribution  $M(x,L,t)$  for the location of the center of mass of the cell and length of the cell.

The frequency distribution  $M(x,L,t)$  determines the approximation  $P_{cpm}(x,L,t) = M(x,L,t)/[N(\varepsilon\Delta x)^2]$  of the probability density function  $P(x,L,t)$  for the center of mass of a cell of length  $L$  being at  $x$  at time  $t$ . Therefore, we compare  $P_{cpm}(x,L,t)$  with  $P(x,L,t)$  which is a solution of either the master equation (4) or the Fokker-Planck equation (7). To approximate the probability density function of center of mass  $p(x,t)$  we sum up over all values of  $L$  on the grid in a way used in Eq. (18):

$$P_{cpm}(x,t) = 2\varepsilon\Delta x \sum_{L=(1+\alpha)\varepsilon\Delta x, (3+\alpha)\varepsilon\Delta x, (5+\alpha)\varepsilon\Delta x, \dots} P_{cpm}(x,L,t), \\ \alpha = 1 \text{ for } \frac{x}{\varepsilon\Delta x} = n, \quad \alpha = 0 \text{ for } \frac{x}{\varepsilon\Delta x} = n + 1/2, \quad n \in \mathbb{N}. \quad (24)$$

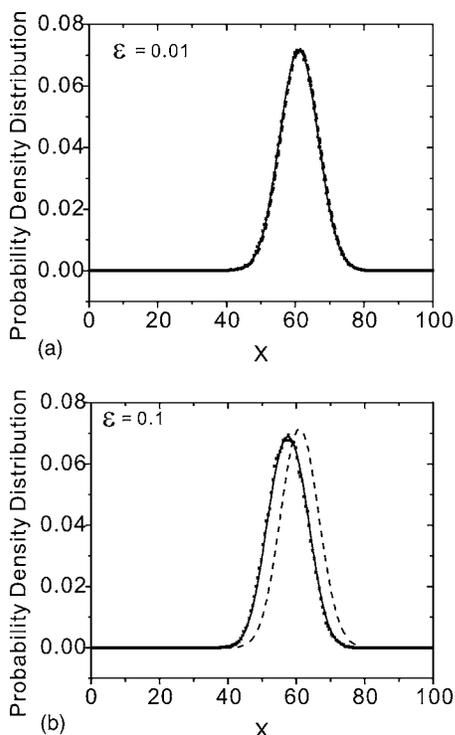


FIG. 3. Probability densities for Monte Carlo simulations  $p_{cpm}(x, t)$  (dotted line),  $p(x, t)$  for the master equation (4) (solid line) and the Fokker-Planck equation (7) (dashed line) versus  $x$  for  $t = t_{end}$ . (a)  $\varepsilon = 0.01$ , (b)  $\varepsilon = 0.1$ . The difference between position of solid curve and a dashed curve is negligibly small in (a). Number of Monte Carlo simulations is  $N = 2 \times 10^5$ . We used  $c(x)$  as given by Eq. (26).

In what follows, we compare  $p_{cpm}(x, t)$  for  $\varepsilon \ll 1$  with  $p(x, t)$ , a solution of the continuous equation (19), corresponding to the following choice of parameters:

$$\lambda = 4, \quad L_T = 5, \quad J_{cm} = 2, \quad \beta = 15, \quad \mu = 0.1, \\ \Delta x = 1, \quad \Delta t = 1. \quad (25)$$

The size of the CPM lattice is chosen to be  $L_{cpm} = 100$ , and the model is typically run from  $t_0 = 0$  to  $t_{end} = 200$ . The number of the CPM lattice sites and the number of Monte Carlo steps are chosen to be  $\frac{L_{cpm}}{\varepsilon \Delta x}$  and  $\frac{t_{end}}{\varepsilon^2 \Delta t}$ , respectively. We use a range of values of  $\varepsilon$  between 0.2 and 0.001.

The initial conditions for each CPM run are chosen as follows. A random pixel in the interval  $[40, 60]$  is selected as a center of mass of a cell, and then the length  $L$  for the cell is chosen with probability  $Z_l^{-1} \exp[-\beta E(L)]$ . Here the normalization constant  $Z_l$  is chosen to have the total probability 1. In most simulations, we use the following static distribution for the chemical field  $c(x)$ :

$$c(x) = \frac{(x - 70)^2}{400}. \quad (26)$$

We use periodic boundary condition to simplify the Monte Carlo and numerical calculations. The chemical field  $c(x)$  which we choose in our paper has the lowest value in the middle of the one-dimensional space domain. The initial cell

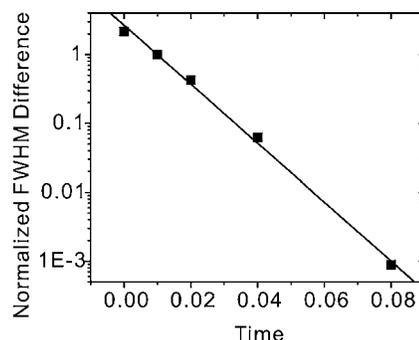


FIG. 4. Exponential convergence of the full width at half maximum (FWHM) of  $P(x, L, t)$  in  $L$  as a function of time for  $x = 50$ . The vertical axis corresponds to the normalized difference  $[W(t) - W_\beta]/W_\beta$ , where  $W(t)$  is the FWHM at time  $t$  and  $W_\beta$  is the FWHM for the Boltzmann distribution (8). Solid squares correspond to the numerical solution of both Eqs. (4) and (7). The solid line is the best linear fit which gives exponential convergence  $e^{-98.55t}$ . The same parameters as in Fig. 3 are used here with  $\varepsilon = 0.01$ .

distribution also is centered in the middle part of the domain. Driven by chemical potential, all cells move toward the middle of the space domain and no cell movement through boundary points is observed in Monte Carlo simulations. The numerical result shows that probability density values at the boundary points are always zero. We checked that a change of periodic boundary conditions into no-flux boundary conditions has no effect on the final result.

### B. Monte Carlo simulations versus numerical solutions of the discrete master equation and the Fokker-Planck equations

We first compare Monte Carlo simulations with the numerics for the master equation (4) and the Fokker-Planck equation (7). Simulations of the Fokker-Planck equation (7) have been performed by using a finite-difference scheme. We verified the result of a finite-difference scheme by varying the grid size and time step. Typically we used 1000 mesh points in  $x$  and time step is 0.0002. We also checked smaller time step (0.00005) to ensure the convergence of our method. Figure 3 shows the probability density functions for all three types of simulations.

The difference between the master equation (4) and the Fokker-Planck equation (7) simulations is negligibly small for  $\varepsilon = 0.01$  [Fig. 3(a)] but can be clearly seen for  $\varepsilon = 0.1$  [Fig. 3(b)]. For the parameters of Fig. 3(b), discreteness causes a shift of results compared with the continuous model. This is a result of the discrete model being compared to the continuous approximation used in the derivation of the Fokker-Planck equation. We suggest a qualitative explanation: this shift for  $\varepsilon = 0.1$  follows from the fact that  $P(x, L, t)$  have only a few points in the discrete model across the width of the Boltzmann distribution as a function of  $L$ . These points generally are not symmetric in respect to the maximum of the Boltzmann distribution. Thus replacement of discrete sum over  $L$  by integral [e.g., as in Eq. (18)] is not a good approximation in that case.

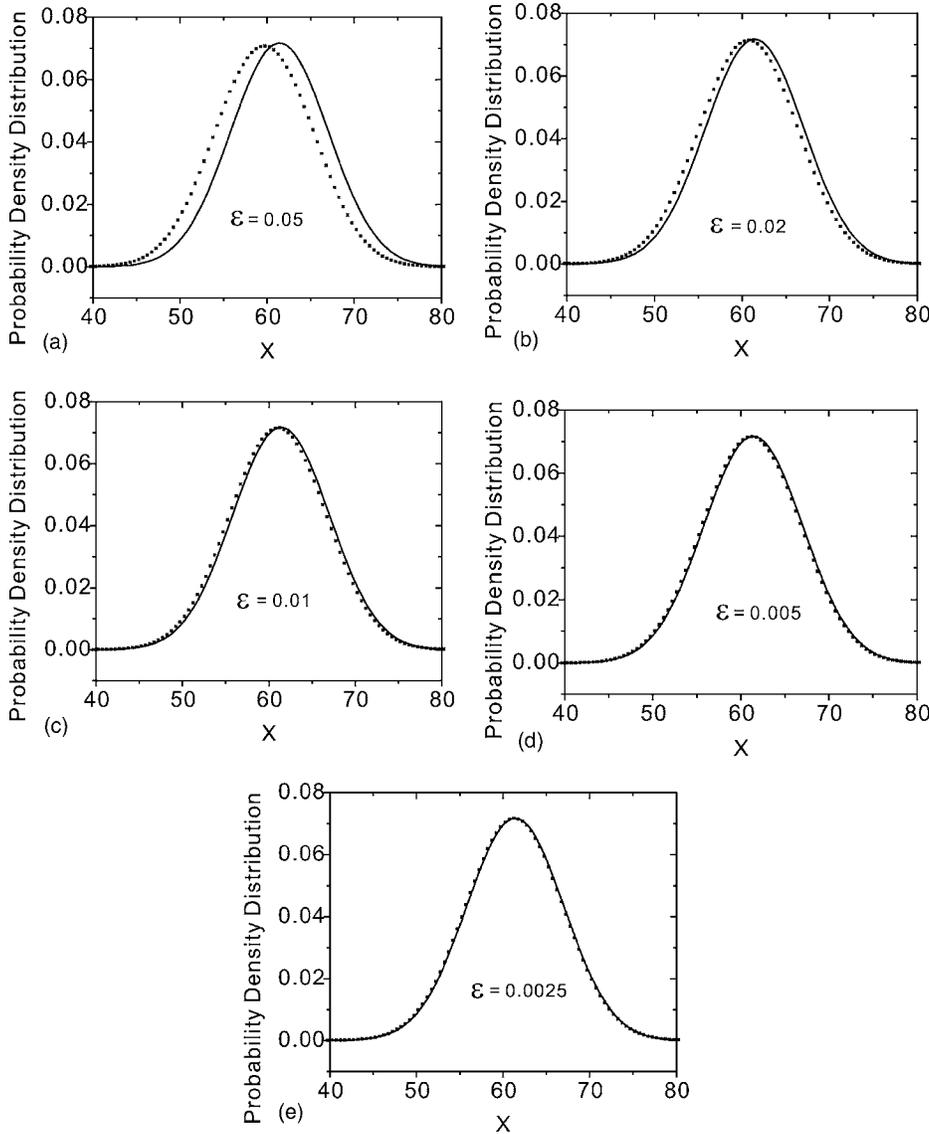


FIG. 5. Plots of  $p_{cpm}$  (dotted line) and  $p_{cont}(x,t)$  (solid line) as functions of  $x$  for a series of decreasing values of  $\varepsilon$  at time  $t=200$ . All other parameters are the same as in Fig. 3.

We conclude that for  $N \rightarrow \infty$ , the Monte Carlo simulations converge to the solution of the master equation (4) for any  $\varepsilon$ . The rate of convergence is about  $N^{-1/2}$ . For small  $\varepsilon \rightarrow 0$ , the solution of the Fokker-Planck equation (7) also converges to the solution of the master equation.

### C. Convergence of the probability density function $P(x,L,t)$ to the Boltzmann distribution

To demonstrate quick convergence of  $P(x,L,t)$  to the Boltzmann distribution (8) (as discussed in Sec. IV) we solve numerically both the master equation (4) and its continuous limit (7) with initial conditions  $P(x,L,0)$  being different from the Boltzmann distribution (4). Namely, we choose initial value  $P(x,L,0)$  to be the Boltzmann distribution with different temperature  $\beta_{ini}=1.5$  so that Fig. 4 shows convergence of initial state with temperature  $1/\beta_{ini}$  to the quasi-equilibrium state with temperature  $1/\beta=1/15$  used in the Monte Carlo algorithm. Linear-log plot in Fig. 4 indicates that convergence is indeed exponential in time with high

convergence rate  $\tau_r^{-1}$  ( $\tau_r^{-1}=98.55$  for parameters of Fig. 4). By a high convergence rate we mean that the typical convergence time  $\tau_r$  is small compared with, e.g., the diffusion time  $x_0^2/D$  in  $x$  [see Eq. (7)]. Because of the  $x$  dependence of the chemical field, the convergence rate  $\tau_r$  is also  $x$  dependent and a closed analytic expression for it is difficult to obtain from Eq. (7) for general  $c(x)$ . However, even a simple estimate  $\tau_r^{-1}=8D\beta\lambda$  of the rate of convergence gives 60 for the parameters of Fig. 4 which is qualitatively close to numerical value 98.55. Here 98.55 is obtained from the linear fit presented in Fig. 4. Also we choose initial value of  $P(x,L,0)$  in  $L$  in a form of step function:  $P(x,L,0)=\text{const}$  for  $0.1 \leq L \leq 5$  and 0 otherwise; then, we found that  $P(x,L,t)$  again converges fast in time (convergence rate is  $\approx 52.38$  in that case) to the Boltzmann distribution (8).

Also, we observe that if we increase temperature  $T$  in Monte Carlo simulations, so that the condition (14) is not true any more, then it results in a significant departure from the Boltzmann distribution (8) which confirms the theoretical results of Sec. IV.

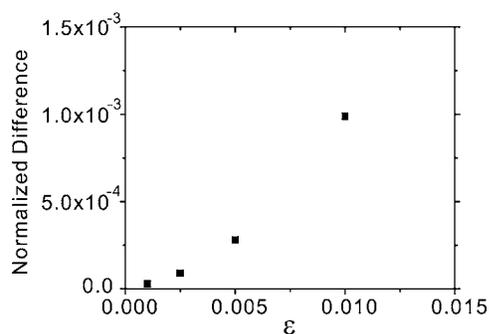


FIG. 6. Normalized difference between solution of the CPM and continuous equation (19) for the same parameters as in Fig. 3 as a function of  $\varepsilon$ . The normalized difference is given by  $1 - \int p_{cpm}(x,t)p_{cont}(x,t)dx / \int p_{cont}(x,t)^2 dx$  for  $t=t_{end}$ .

#### D. $P(x,L,t)$ vs $p(x,t)$ simulations

The ansatz (8) can be used for fast simulations of solutions of the discrete master equation. Summing up over all values of  $L$  in the master equation (4) and taking into account the result of Eq. (8) in a discrete equation for the probability density function  $p(x,t)$  we obtain:

$$\begin{aligned}
 p(x,t + \varepsilon^2 \Delta t) = & \left[ 1 - T\left(x + \frac{\varepsilon}{2} \Delta x; x, t\right) - T\left(x - \frac{\varepsilon}{2} \Delta x; x, t\right) \right] \\
 & \times p(x,t) + T\left(x; x - \frac{\varepsilon}{2} \Delta x, t\right) p\left(x - \frac{\varepsilon}{2} \Delta x, t\right) \\
 & + T\left(x; x + \frac{\varepsilon}{2} \Delta x, t\right) p\left(x + \frac{\varepsilon}{2} \Delta x, t\right), \quad (27)
 \end{aligned}$$

where  $T(x;x',t)$  is a transition probability of a change of position of a center mass from  $x'$  to  $x$  at time  $t$ . Expressions for  $T(x;x',t)$  are described in the Appendix. They are calculated only once at the beginning of a simulation which makes the numerics for the discrete equation (27) very efficient.

We run simulations for the discrete equation (27) and the continuous equation (19) and compared them with the solutions of the discrete (4) and continuous (7) equations, respectively. We find, taking into account Eq. (8), that indeed the differences between these solutions are very small for typical values of the parameters.

We conclude that the Monte Carlo simulations of the CPM are equivalent in the limit of large  $N$  to the simulations of the discrete Eq. (27) for any  $\varepsilon$ .

#### E. Comparison of the continuous model with the CPM

Below we denote as  $p_{cpm}$  both Monte Carlo simulations and numerical solutions of Eq. (27) and as  $p_{cont}(x,t)$  solutions of Eq. (19).

Figure 5 shows a series of simulations of the CPM (dotted line) and numerical solutions of the continuous equation (19) (solid line) for different values of  $\varepsilon$ .

This figure demonstrates that in the limit  $\varepsilon \rightarrow 0$ , the solution of the continuous equation (19) appears to converge to the cell probability density function of the CPM.

Figure 6 shows the normalized difference between solu-

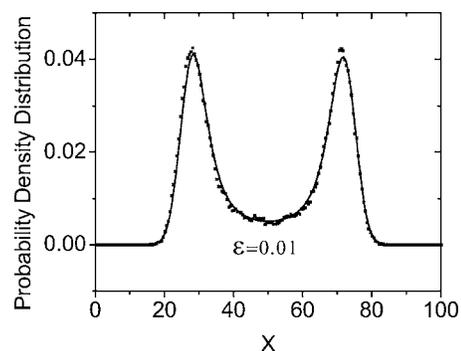


FIG. 7. Typical results of CPM simulations. The same parameters as in Fig. 3 are used except that  $c(x)=\cos(4\pi x/100)$ ,  $\varepsilon=0.01$ . The same notation for solid, dashed, and dotted curves as in Fig. 3 is used here. The difference between position of solid curve and a dashed curve is again negligibly small.

tions of Eq. (19) and the CPM. The normalized difference approaches 0 as  $\varepsilon$  decreases.

We also run a series of tests for different forms of the chemical field  $c(x)$  and demonstrate that solutions of the CPM and continuous equation (19) are close for small values of  $\varepsilon$ . Figure 7 shows a typical result of numerical simulations for a “double-well” chemical concentration  $c(x)=\cos(4\pi x/100)$ .

We conclude that the numerical simulations show excellent agreement between the CPM and the continuous equation (19) provided that the Potts parameters satisfy conditions (14)–(17) and  $\varepsilon \rightarrow 0$ , which correspond to the continuous limit of the CPM.

## VIII. CONCLUSIONS

In this paper we combine microscopic and macroscopic levels of description of one-dimensional cellular dynamics. The microscopic level is represented by a one-dimensional CPM with chemotaxis and without a cell-cell adhesion term. We study a continuous macroscopic limit of our CPM as the size of the Monte Carlo step is made small under the assumption that changes in the cell’s position and length are also small. In this limit, we derive the Fokker-Planck equation (19) for the probability density function  $p(x,t)$  of cells and then further reduce it to the well-known macroscopic continuous Keller-Segel models (20) and (23) for the chemotactic aggregation of cells. All coefficients of the Keller-Segel model are derived from parameters of the CPM.

We use numerical simulations to test hierarchy of models and assumptions which we used to derive the continuous equation (19). In particular, we compare Monte Carlo simulations with simulations of both the discrete master equation (4) and the Fokker-Planck equation (7) for  $P(x,L,t)$ . We find that, as expected from our theoretical analysis, all models agree for small  $\varepsilon$ . Also Monte Carlo simulations agree with the solutions of the discrete master equation (4) for arbitrary  $\varepsilon$ . We verify numerically that the probability density function  $P(x,L,t)$  quickly converges to the Boltzmann distribution (8). And finally, we find that numerical simulations show excellent agreement between Monte Carlo simulations of the

CPM and the continuous macroscopic model (19).

We are currently working on extending our results to a 2D case for modeling chondrogenic patterning in the presence of chemotaxis and fibronactin production [30].

#### ACKNOWLEDGMENTS

This work was partially supported by NSF Grant No. IBN-0083653. Simulations were performed on the Notre Dame Biocomplexity Cluster supported in part by NSF MRI Grant No. DBI-0420980.

#### APPENDIX

The explicit expressions for the transitional probabilities  $T(x; x', t)$  used in Eq. (27) can be obtained by summing over all lengths (or, in other words, over even multiples of  $\varepsilon\Delta x$  [if  $2x/(\varepsilon\Delta x)$  is an even number] and over odd multiples of  $\varepsilon\Delta x$  [if  $2x/(\varepsilon\Delta x)$  is an odd number]). A change in the position of the center of mass from  $x$  to  $x \pm \frac{\varepsilon}{2}\Delta x$  can be made by adding (removing) lattice sites from the left (right) end of a cell which results in

$$\begin{aligned}
 T\left(x; x - \frac{\varepsilon}{2}\Delta x, t\right) &= \frac{1}{4Z\left(x - \frac{\varepsilon}{2}\Delta x\right)} \sum_{L=(1+\alpha)\varepsilon\Delta x, (3+\alpha)\varepsilon\Delta x, (5+\alpha)\varepsilon\Delta x, \dots} \left\{ \exp\left[-\beta\Delta E_{length}\left(x - \frac{\varepsilon}{2}\Delta x, L - \varepsilon\Delta x\right)\right] \Phi\left(E(x, L) - E\left(x - \frac{\varepsilon}{2}\Delta x, L - \varepsilon\Delta x\right)\right) \right. \\
 &\quad \left. + \exp\left[-\beta\Delta E_{length}\left(x - \frac{\varepsilon}{2}\Delta x, L + \varepsilon\Delta x\right)\right] \Phi\left(E(x, L) - E\left(x - \frac{\varepsilon}{2}\Delta x, L + \varepsilon\Delta x\right)\right) \right\}, \\
 T\left(x; x + \frac{\varepsilon}{2}\Delta x, t\right) &= \frac{1}{4Z\left(x + \frac{\varepsilon}{2}\Delta x\right)} \sum_{L=(1+\alpha)\varepsilon\Delta x, (3+\alpha)\varepsilon\Delta x, (5+\alpha)\varepsilon\Delta x, \dots} \left\{ \exp\left[-\beta\Delta E_{length}\left(x + \frac{\varepsilon}{2}\Delta x, L - \varepsilon\Delta x\right)\right] \Phi\left(E(x, L) - E\left(x + \frac{\varepsilon}{2}\Delta x, L - \varepsilon\Delta x\right)\right) \right. \\
 &\quad \left. + \exp\left[-\beta\Delta E_{length}\left(x + \frac{\varepsilon}{2}\Delta x, L + \varepsilon\Delta x\right)\right] \Phi\left(E(x, L) - E\left(x + \frac{\varepsilon}{2}\Delta x, L + \varepsilon\Delta x\right)\right) \right\}, \\
 T\left(x + \frac{\varepsilon}{2}\Delta x; x, t\right) &= \frac{1}{4Z(x)} \sum_{L=(1+\alpha)\varepsilon\Delta x, (3+\alpha)\varepsilon\Delta x, (5+\alpha)\varepsilon\Delta x, \dots} \left\{ \exp\left[-\beta\Delta E_{length}(x, L)\right] \Phi\left(E\left(x + \frac{\varepsilon}{2}\Delta x, L - \varepsilon\Delta x\right) - E(x, L)\right) \right. \\
 &\quad \left. + \exp\left[-\beta\Delta E_{length}(x, L)\right] \Phi\left(E\left(x + \frac{\varepsilon}{2}\Delta x, L + \varepsilon\Delta x\right) - E(x, L)\right) \right\}, \\
 T\left(x - \frac{\varepsilon}{2}\Delta x; x, t\right) &= \frac{1}{4Z(x)} \sum_{L=(1+\alpha)\varepsilon\Delta x, (3+\alpha)\varepsilon\Delta x, (5+\alpha)\varepsilon\Delta x, \dots} \left\{ \exp\left[-\beta\Delta E_{length}(x, L)\right] \Phi\left(E\left(x - \frac{\varepsilon}{2}\Delta x, L - \varepsilon\Delta x\right) - E(x, L)\right) \right. \\
 &\quad \left. + \exp\left[-\beta\Delta E_{length}(x, L)\right] \Phi\left(E\left(x - \frac{\varepsilon}{2}\Delta x, L + \varepsilon\Delta x\right) - E(x, L)\right) \right\}, \\
 \alpha &= 1 \text{ for } \frac{x}{\varepsilon\Delta x} = n, \quad \alpha = 0 \text{ for } \frac{x}{\varepsilon\Delta x} = n + 1/2, \quad n \in \mathbb{N}. \tag{A1}
 \end{aligned}$$

Here the partition function  $Z(x)$  is given by (12).  $Z(x)$  is  $x$  dependent in the discrete case considered in this appendix. This  $x$  dependence is eliminated after going from a discrete summation in Eq. (12) to an integral [as in Eq. (13)]. We evaluate the transitional probabilities  $T(x; x \pm \frac{\varepsilon}{2}\Delta x, t)$  and  $T(x \pm \frac{\varepsilon}{2}\Delta x, t)$  numerically using Eq. (A1) for each value of  $x$  once at the beginning of each simulation and then calculate the discrete evolution of Eq. (27).

Notice that in the limit of small  $\varepsilon \rightarrow 0$ , the continuous equation (19) can be derived directly from Eqs. (12), (27), and (A1). However, this derivation is more tedious compared with the two-step derivation in Secs. IV and V where continuous equation (7) is first derived and then integrated (7) over  $L$  which results in Eq. (19).

- [1] E. F. Keller and L. A. Segel, *J. Theor. Biol.* **26**, 399 (1970).
- [2] M. P. Brenner, L. Levitov, and E. O. Budrene, *Biophys. J.* **74**, 1677 (1998).
- [3] M. P. Brenner, P. Constantin, L. P. Kadanoff, A. Schenkel, and S. C. Venkataramani, *Nonlinearity* **12**, 1071 (1999).
- [4] R. Erban and H. G. Othmer, *SIAM J. Appl. Math.* **65**, 361 (2004).
- [5] E. Ben-Jacob, I. Cohen, and H. Levine, *Adv. Phys.* **49**, 395 (2000).
- [6] D. Kaiser, *Annu. Rev. Microbiol.* **58**, 7598 (2004).
- [7] O. Sozinova, Y. Jiang, D. Kaiser, and M. Alber, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 11308 (2005).
- [8] T. J. Newman and R. Grima, *Phys. Rev. E* **70**, 051916 (2004).
- [9] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [10] M. S. Alber, M. A. Kiskowski, Y. Jiang, and S. A. Newman, in *Dynamics and Bifurcation of Patterns in Dissipative Systems*, edited by G. Dangelmayr and I. Oprea World Scientific Series on Nonlinear Science, Vol. 12 (World Scientific, Singapore, 2004), pp. 274–291.
- [11] C. J. Weijer, *Science* **300**, 96 (2003).
- [12] F. Graner and J. A. Glazier, *Phys. Rev. Lett.* **69**, 2013 (1992).
- [13] J. A. Glazier, and F. Graner, *Phys. Rev. E* **47**, 2128 (1993).
- [14] A. F. M. Marée, A. V. Panfilov, and P. Hogeweg, *J. Theor. Biol.* **199**, 297 (1999).
- [15] A. F. M. Marée, A. V. Panfilov, and P. Hogeweg, *Proc. R. Soc. London, Ser. B* **266**, 1351 (1999).
- [16] R. M. H. Merks, S. A. Newman, and J. A. Glazier, in *Cellular Automata: Proceedings of the 6th International Conference on Cellular Automata for Research and Industry, ACRI 2004, Amsterdam, 2004*, edited by P. M. A. Sloot, B. Chopard, and A. G. Hoekstra, Lecture Notes in Computer Science Vol. 3305 (Springer-Verlag, Heidelberg, 2004), pp. 425–434.
- [17] S. Turner and J. A. Sherratt, *J. Theor. Biol.* **216**, 85 (2002).
- [18] R. Chaturvedi, C. Huang, B. Kazmierczak, T. Schneider, J. A. Izaguirre, T. Glimm, H. G. E. Hentschel, J. A. Glazier, S. A. Newman, and M. Alber, *J. R. Soc., Interface* **2**, 237 (2005).
- [19] T. Cickovski, C. Huang, R. Chaturvedi, T. Glimm, H. G. E. Hentschel, M. Alber, J. A. Glazier, S. A. Newman, and J. A. Izaguirre, *IEEE/ACM Trans. Comput. Biol. Bioinformatics* **3**, 1545 (2005).
- [20] M. S. Alber, M. A. Kiskowski, J. A. Glazier, and Y. Jiang, in *Mathematical Systems Theory in Biology, Communication, and Finance*, edited by J. Rosenthal and D. S. Gilliam IMA Vol. 134 (Springer-Verlag, New York, 2003), pp. 1–39.
- [21] T. J. Newman, *Math. Biosc. Eng.* **2**, 611 (2005).
- [22] W. Alt, *J. Math. Biol.* **9**, 147 (1980).
- [23] A. Stevens, *SIAM J. Appl. Math.* **61**, 183 (2000).
- [24] S. Turner, J. A. Sherratt, K. J. Painter, and N. J. Savill, *Phys. Rev. E* **69**, 021910 (2004).
- [25] D. Dormann, B. Vasiev, and C. J. Weijer, *J. Biol. Phys.* **28**, 765 (2002).
- [26] B. N. Vasiev, P. Hogeweg, and A. V. Panfilov, *Phys. Rev. Lett.* **73**, 3173 (1994).
- [27] <http://www.personal.dundee.ac.uk/~cjweijer/dictyweb/projects.htm>
- [28] J. Adler, *Annu. Rev. Biochem.* **44**, 341 (1975).
- [29] R. Thar and M. Köhl, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 5748 (2003).
- [30] M. A. Kiskowski, M. S. Alber, G. L. Thomas, J. A. Glazier, N. Bronstein, J. Pu, and S. A. Newman, *Dev. Biol.* **271**, 372 (2004).