Macroscopic dynamics of biological cells interacting via chemotaxis and direct contact

Pavel M. Lushnikov,[1,*] Nan Chen,[2] and Mark Alber[2]
[1]*Department of Mathematics and Statistics, University of New Mexico, Albuquerque, New Mexico 87131, USA*
[2]*Department of Mathematics, University of Notre Dame, Notre Dame, Indiana 46656, USA*
(Received 18 August 2008; revised manuscript received 12 October 2008; published 3 December 2008)

A connection is established between discrete stochastic model describing microscopic motion of fluctuating cells, and macroscopic equations describing dynamics of cellular density. Cells move towards chemical gradient (process called chemotaxis) with their shapes randomly fluctuating. Nonlinear diffusion equation is derived from microscopic dynamics in dimensions one and two using excluded volume approach. Nonlinear diffusion coefficient depends on cellular volume fraction and it is demonstrated to prevent collapse of cellular density. A very good agreement is shown between Monte Carlo simulations of the microscopic cellular Potts model and numerical solutions of the macroscopic equations for relatively large cellular volume fractions. Combination of microscopic and macroscopic models were used to simulate growth of structures similar to early vascular networks.

DOI: 10.1103/PhysRevE.78.061904 PACS number(s): 87.18.Ed, 05.40.−a, 05.65.+b, 87.18.Hf

I. INTRODUCTION

So far most models used in biology have been developed at specific scales. Establishing a connection between discrete stochastic microscopic description and continuous deterministic macroscopic description of the same biological phenomenon would allow one to switch when needed from one scale to another, considering events at individual (microscopic) cell level such as cell-cell interaction or cell division to events involving thousands of cells such as organ formation and development. Due to the fast calculation speed possible with the continuous model, one can quickly test wide parameter ranges and determine satiability conditions and then use this information for running Monte Carlo simulations of the stochastic discrete dynamical systems. Also, continuous models provide very good approximation for systems containing a biologically realistic (i.e., large) number of cells, for which numerical simulations of stochastic trajectories can be prohibitive.

Most continuous biological models have been postulated either by requiring certain biologically relevant features from the solutions or making it easier to analyze behavior of solutions using certain mathematical techniques. In particular, a system of nonlinear partial differential equations (PDE) model with chemotactic term was used in [1,2] to simulate the de novo blood vessel formation from the mesoderm. The rational for the model was provided by the experimental observations [3] demonstrating that chemotaxis played an important role in guiding cells during early vascular network formation. Discrete models have been also applied to simulating vasculogenesis and angeogenesis [4–6].

In this paper we derive continuous macroscopic limits of the one-dimensional (1D) and two-dimensional (2D) microscopic cell-based models with extended cell representations, in the form of nonlinear diffusion equations coupled with chemotaxis equation. We demonstrate that combination of the discrete model and derived continuous model can be used for simulating biological phenomena in which a nonconfluent population of cells interact directly and via diffusible factors, forming an open network structure in a way similar to formation of networks during vasculogenesis [1,2] and pattern formation in limb cell cultures [7].

Continuous limits of microscopic models of biological systems based on pointwise cell representation were extensively studied over the last 30 years. The classical Keller-Segel PDE model has been derived in [8] from a discrete model with pointwise cells undergoing random walk in chemotactic field and then studied in [9–12]. Cells in this model secrete a diffusing chemical at a constant rate and detect local concentration c of this chemical due to a process called chemotaxis. The chemical is called an attractant or repellent depending on whether the cell moves towards chemical gradient or in the opposite direction. Aggregation occurs if attraction exceeds diffusion of cells. For pointwise cells aggregation results in infinite cellular density corresponding to the solution of the macroscopic Keller-Segel equation becoming infinite in finite time (also called blow up in finite time or collapse of the solution) [13,14].

There have been few attempts to derive macroscopic limits of microscopic models which treat cells as extended objects consisting of several points. In [15] the diffusion coefficient for a collection of noninteracting randomly moving cells was derived from a one-dimensional cellular Potts model (CPM). A microscopic limit of subcellular elements model [16] was derived in the form of an advection-diffusion PDE for cellular density. In our previous papers [17–19] we studied the continuous limit of the CPM describing individual cell motion in a medium and in the presence of an external field with contact cell-cell interactions in mean-field approximation. However, mean-field approximation does not allow one to consider high density of cells when cellular volume fraction (fraction of volume occupied by cells) φ is of the order of 1.

In this paper we go beyond mean-field approximation. Namely, we take into account the finite size of cells in the CPM, use exclusion volume principle (meaning that two cells cannot occupy the same volume) and derive the follow-

*Author to whom correspondence should be addressed; plushnik@math.unm.edu

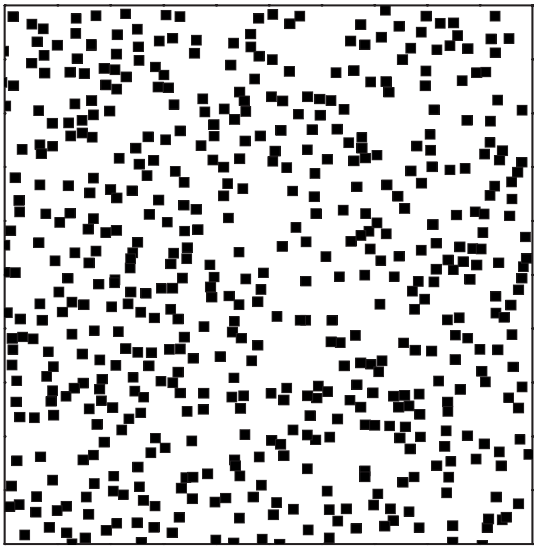

FIG. 1. Representation of cells in the 2D CPM in the form of fluctuating rectangles. 539 cells (15% volume fraction) with $L_x^{(0)}=L_y^{(0)}=1.666\,667$ are shown.

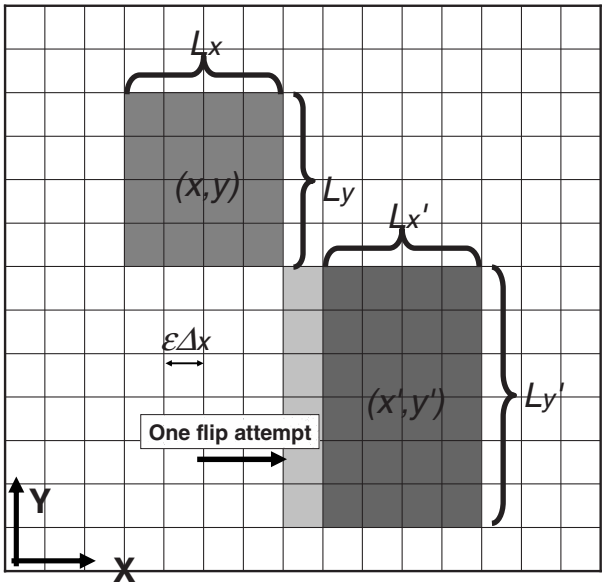

FIG. 2. 2D CPM cell representation. Grey and white colors are used to indicate cell body and surrounding extracellular matrix, respectively. The cell can grow or shrink in x and y directions by adding or removing one row (or column) of pixels.

ing macroscopic nonlinear diffusion equation for evolution of cellular density $p(\mathbf{r},t)$ in one dimension (1D),

$$\partial_t p = D_2 \nabla_{\mathbf{r}} \cdot \left(\frac{1+\varphi^2}{(1-\varphi)^2} \nabla_{\mathbf{r}} p \right) - \chi_0 \nabla_{\mathbf{r}} \cdot [p \nabla_{\mathbf{r}} c(\mathbf{r},t)], \quad (1)$$

and two dimensions (2D),

$$\partial_t p = D_2 \nabla_{\mathbf{r}} \cdot \left(\frac{1+\varphi}{1-\varphi+\varphi \ln(\varphi)} \nabla_{\mathbf{r}} p \right) - \chi_0 \nabla_{\mathbf{r}} \cdot [p \nabla_{\mathbf{r}} c(\mathbf{r},t)], \quad (2)$$

which do not have blow up in finite time. These nonlinear diffusion equations are coupled with the equation for evolution of chemical field $c(\mathbf{r},t)$,

$$\partial_t c(\mathbf{r},t) = D_c \nabla_{\mathbf{r}}^2 c + ap - \gamma c. \quad (3)$$

Here φ is a volume fraction (fraction of volume occupied by cells). In the 1D case cells have a form of fluctuating rods and $\varphi=L_x^{(0)}p(\mathbf{r},t)$, where $L_x^{(0)}$ is an average length of cells. In the 2D case we assume that cells are fluctuating rectangles and $\varphi=L_x^{(0)}L_y^{(0)}p(\mathbf{r},t)$, where $(L_x^{(0)},L_y^{(0)})$ are average length and width of a cell. Here $p(\mathbf{r},t)$ is the density of cells normalized by the total number of cells N: $\int p(\mathbf{r},t)d\mathbf{r}=N$ and $\mathbf{r}$ is a vector of spatial coordinates in 1D or 2D. D_2 is the diffusion coefficient for a motion of an isolated cell, χ_0 defines strength of chemotactic interactions, D_c is the diffusion coefficient of a chemical, a is the production rate of a chemical, and γ is the decay rate of a chemical. A typical microscopic picture of distribution of individual cells is shown in Fig. 1 for the 2D case. Solutions of Eqs. (1)–(3) describe coarse-grained macroscopic cellular dynamics. Below we show very good agreement even for relatively large densities $\varphi \simeq 0.3$, between solutions of Eqs. (1)–(3) and ensemble average of stochastic trajectories of the microscopic CPM.

In Sec. II we introduce general microscopic equations describing motion of cells based on random fluctuations of their shapes and their interactions. We assume that each cell has a rectangular shape and consider stochastic differential equations for motion of cells as well as the Smoluchowski equation (see, e.g., [20]) for multicellular probability density function. In Sec. III we consider a particular case of microscopic cellular dynamics represented by the CPM without excluded volume interactions. Coefficients of the Smoluchowski equation are derived from the CPM and stochastic dynamics of shapes and finding positions of cells is reduced to solving the closed equations describing positions of cells. In Sec. IV we consider cell motion and cell-cell interactions with collisions resolved through the jump processes resulting in Eqs. (1)–(3). In Sec. V numerics for the continuous macroscopic equation is compared with Monte Carlo simulations of the microscopic CPM. In Sec. VI we summarize the main results and discuss future directions.

II. MICROSCOPIC MOTION OF CELLS

Motion of many eukaryotic cells and bacteria is accompanied by the random fluctuations of their shapes [21–23] resulting in a random diffusion of center of mass of an isolated cell. Coefficient of such diffusion can be measured experimentally (see, e.g., [24]). Fluctuations of cellular membrane in the presence of a chemical field are more likely in the direction of chemical gradient (for chemoattractant) or in the opposite direction (for chemorepellent). Cells can also interact through direct contact which includes cell-cell adhesion and can be modeled using excluded volume principle. Cellular environment is highly viscous and inertia of a cell can be ignored.

In this paper we assume that each cell has a fluctuating rectangular shape and allow random fluctuation of the dimensions of each rectangular cell (see Figs. 1 and 2). Posi-

tions and shapes of cells are completely characterized by a finite set of dynamical variables in the configuration space: $X=(\mathbf{R}_1,\mathbf{L}_1,\mathbf{R}_2,\mathbf{L}_2,\ldots,\mathbf{R}_N,\mathbf{L}_N)$, where N is the total number of cells, $\mathbf{R}_j$ is a position of center of mass of the jth cell, $\mathbf{L}_j$ is the size of the jth cell in D spatial dimensions. We consider $D=1$ and $D=2$ in which case cells are moving over substrate but results can be extended to the $D=3$ case. Microscopic description is provided by the multicellular probability density function (PDF) $P(X,t)$ defined as ensemble average $\langle\cdots\rangle$ over stochastic trajectories $[\mathbf{R}'_j(t),\mathbf{L}'_j(t)]$, $j=1,\ldots,N$; in the configuration space X: $P(X,t)=\langle\Pi_{j=1}^{N}\delta[\mathbf{R}_j-\mathbf{R}'_j(t)]\delta[\mathbf{L}_j-\mathbf{L}'_j(t)]\rangle$ determined by a solution of the coupled stochastic equation

$$\frac{dX}{dt}=\mathbf{A}(X,c,t)+\hat{B}(X,c,t)\xi(t), \tag{4}$$

and chemotaxis equation for the chemical field $c(\mathbf{r},t)$. Here $\mathbf{r}$ is the spatial coordinate, $\mathbf{A}(X,t)$ is $2DN$-component vector, $\hat{B}(X,t)$ is $2DN\times 2DN$ matrix, and $\xi(t)$ is the $2DN$-component stationary Gaussian stochastic process with zero correlation time and zero mean

$$\langle\xi_j\rangle=0,\quad \langle\xi_i(t)\xi_j(t')\rangle=\delta_{i,j}\delta(t-t'),\quad i,j=1,\ldots,2DN, \tag{5}$$

where $\delta_{i,j}$ is a Kronecker's symbol.

Application of the Stratonovich stochastic calculus to Eq. (4) results in a "multicellular" Fokker-Planck equation [20] in configuration space X,

$$\partial_t P(X,t)=-\nabla\cdot[\mathbf{v}P]+\nabla\cdot[\hat{D}\cdot\nabla P],$$

$$\mathbf{v}=\mathbf{A}-\tfrac{1}{2}\hat{B}\cdot\nabla\cdot\hat{B}^T,$$

$$\hat{D}=\tfrac{1}{2}\hat{B}\cdot\hat{B}^T, \tag{6}$$

where $\hat{D}(X,t)$ is a $2DN\times 2DN$ diffusion matrix, $\nabla=(\partial_{\mathbf{R}_1},\partial_{\mathbf{L}_1},\ldots,\partial_{\mathbf{R}_N},\partial_{\mathbf{L}_N})$ is the gradient operator in $2DN$ dimensions, and $\hat{B}^T$ represents transposed matrix $\hat{B}$, and in components second equations reads as $v_j=A_j-\frac{1}{2}\Sigma_{k,n}B_{jn}\partial_k B_{kn}$.

Dynamics of the chemical field $c(\mathbf{r},t)$ is described by a diffusion equation

$$\partial_t c(\mathbf{r},t)=\nabla_{\mathbf{r}}D_c\cdot\nabla_{\mathbf{r}}c+\int A(X,\mathbf{r},t)P(X,t)dX-\gamma c, \tag{7}$$

where $\nabla_{\mathbf{r}}$ is the gradient operator for the spatial coordinate and diffusion coefficient for the chemical field D_c in the general case can depend on $\mathbf{r}$ and t. Term $\int A(X,\mathbf{r},t)P(X,t)dX$ describes production of a chemical by cells at the rate of $A(X,\mathbf{r},t)$ and γ is a decay rate of the chemical. Cells produce chemical into intercellular space through their membranes. Therefore, in the three-dimensional (3D) case $A(X,\mathbf{r},t)$ is nonzero only at the cellular membrane. However, in 1D and 2D cases cells are moving on a substrate, while a chemical diffuses over the entire three-dimensional space. Therefore, in 1D and 2D cases $A(X,\mathbf{r},t)$ is nonzero inside cells.

We assume that $\mathbf{v}$ has a form of a potential $\mathbf{v}=-\beta\hat{D}\cdot\nabla\Phi(X,t)$, where $\beta=1/T$ is the inverse effective temperature T of the cellular shape fluctuations. Multicellular Fokker-Planck equation (6) is then reduced to the multicellular Smoluchowski equation,

$$\partial_t P(X,t)=\nabla\cdot\hat{D}\cdot[\nabla P+\beta(\nabla\Phi)P]. \tag{8}$$

(General case of $\mathbf{v}$, in the form of a function can be studied using a similar approach.) If we neglect fluctuations of the cellular size $\mathbf{L}_j$ and chemotaxis, $c=0$, then Eq. (8) is similar to the Smoluchowski equation for the Brownian dynamics of colloidal particles (see, e.g., [25] for a review) and Eq. (4) has a form of the Langevin equation for interacting Brownian particles with term $\hat{B}\xi$ representing thermal forces from solution in colloids. However, in the general case considered here, both chemotaxis and fluctuations of cellular shape are taken into account. Mechanisms of random fluctuation of cellular shape and cellular motion are still not completely clear and subject of active research [21–23].

We impose the excluded volume constraint by choosing $\Phi(X,c)=\infty$ if any two cells overlap. We assume in what follows that all direct interactions between cells are of this type. We also allow indirect interactions between cells mediated by chemotaxis for which we choose

$$\Phi(X,c)=-\sum_{j=1}^{N}\chi(\mathbf{L}_j)c(\mathbf{r},t)\Big|_{\mathbf{r}=X_j}+\Phi(X,c)|_{c=0}\quad\text{if cells do not overlap},$$

$$\Phi(X,c)=\infty\quad\text{if any pair of cells overlaps},$$

where $\chi(\mathbf{L}_j)$ represents strength of chemotactic interaction as a function of cellular sizes $\mathbf{L}_j$ and $\Phi(X,c)|_{c=0}$ represents chemotaxis-independent terms of the potential. We assume for simplicity that chemotactic interaction depends only on gradient of $c(\mathbf{r},t)$ at the center of mass of each cell. $\Phi(X,c)|_{c=0}$ is also responsible for preserving cellular shape close to some equilibrium shape. Without this term, shape (size) of each cell would experience unbounded random fluctuation which is nonbiological. We consider the specific form of $\Phi(X,c)$ in Sec. III C.

Our main goal is to derive a macroscopic equation describing dynamics of (total) cell probability density function

$$p(\mathbf{r},t)=\sum_{j=1}^{N}p_j(\mathbf{r},t), \tag{9}$$

coupled with $c(\mathbf{r},t)$, from microscopic equations (7) and (8). Here $p_j(\mathbf{r},t)=\int P(X,t)\Pi_{l=1,l\neq j}^{N}d\mathbf{R}_l\Pi_{m=1}^{N}d\mathbf{L}_m|_{\mathbf{R}_j=\mathbf{r}}$ is a single-cell probability density function of the position of center of mass. After approximating $A(X,\mathbf{r},t)=a\Sigma_{j=1}^{N}\delta(\mathbf{r}-\mathbf{R}_j)$, $a=$const and assuming that $D_c=$const, Eq. (7) is reduced with the help of Eq. (9) to Eq. (3). This approximation is justified because typical diffusion of a chemical is much faster than cell diffusion $D_2/D_c\ll 1$. For example, $D_2/D_c\sim 1/40-1/400$ for the cellular slime mold *Dictyostelium* [26], $D_2/D_c\sim 1/30$ for microglia cells and neutrophils [27,28].

III. MICROSCOPIC CELLULAR DYNAMICS IN CELLULAR POTTS MODEL

Stochastic discrete models are used in a variety of problems dealing with biological complexity. One motivation for this approach is the enormous range of length scales of typical biological phenomena. Treating cells as simplified interacting agents, one can simulate the interactions of tens of thousands to millions of cells and still have within reach the smaller-scale structures of tissues and organs that would be ignored in continuum (e.g., partial differential equation) approaches. At the same time, discrete stochastic models including the cellular Potts model (CPM) can be made sophisticated enough to reproduce almost all commonly observed types of cell behavior [29–34]. Reference [35] reviews many cell-based models.

The cell-based stochastic discrete CPM, which is an extension of the Potts model from statistical physics, has become a common technique for simulating complex biological problems including embryonic vertebrate limb development [29,36], tumor growth [37], and vasculogenesis [4]. The CPM can be made sophisticated enough to reproduce almost all commonly observed types of cell behavior. It consists of a list of biological cells with each cell represented by several pixels, a list of generalized cells, a set of chemical diffusants, a description of their biological and physical behaviors, and interactions embodied in the effective energy E, with additional terms to describe absorption and secretion of diffusants and extracellular materials. Distribution of multidimensional indices associated with lattice cites determines the current cell system configuration. The effective energy of the system, E, mixes true energies, like cell-cell adhesion, and terms that mimic energies, e.g., volume constraint and the response of a cell to a gradient of an external field (including chemotactic field) and cell's area constraint.

A. Cellular Potts model for cells of rectangular shape

For simplicity, we use in this paper CPM with rectangular cellular shapes. We also assume that all cells are of the same type. The results can be extended to the general case of the CPM with arbitrary cellular shapes. Also, the approach is not limited to using CPM. For example, one could use microscopic off-lattice models [16,38], where each cell is represented by a collection of subcellular elements with postulated interactions between them.

Notice that reduced representation of cells as fluctuating rectangles corresponds to intermediate level of description where fluctuations of cellular shapes are replaced by fluctuations of cellular sizes. Stochastic equation (4) or Smoluchowski equation (8) coupled with (3) can be used for modeling cell aggregation.

In the CPM, change of a cell shape evolves according to the classical Metropolis algorithm based on Boltzmann statistics and the following effective energy:

$$E = E_{\mathrm{ECM}} + E_{\mathrm{Adhesion}} + E_{\mathrm{Perimeter}} + E_{\mathrm{Field}}. \tag{10}$$

If an attempt to change index of a pixel in a cell leads to energy change ΔE, it is accepted with the probability

$$\Phi(\Delta E) = \begin{cases} 1, & \Delta E \leqslant 0, \\ e^{-\beta \Delta E}, & \Delta E > 0, \end{cases} \tag{11}$$

where $1/\beta = T$ represents an effective boundary fluctuation amplitude of model cells in units of energy. Since the cells' environment is highly viscous, cells move to minimize their total energy consistent with imposed constraints and boundary conditions. If a change of a randomly chosen pixels' index causes cell-cell overlap it is abandoned. Otherwise, the acceptance probability is calculated using the corresponding energy change. The accepted pixel change attempt results in changing location of the center of mass and dimensions of the cell.

We consider the 2D case with rectangular shape of each cell with sizes $\mathbf{L}_j = (L_{x,j}, L_{y,j})$ and position of center of cellular mass at $\mathbf{R}_j = (x_j, y_j)$. Cell motion and changing shape are implemented by adding or removing a row or column of pixels (see Fig. 2). We assume that cells can come into direct contact and that they interact over long distances through chemotaxis. Term E_{ECM} in the Hamiltonian (10) phenomenologically describes net adhesion or repulsion between the cell surface and surrounding extracellular matrix: $E_{\mathrm{ECM}} = \sum_{j=1}^{N} 2J_{\mathrm{ECM}}(L_{x,j} + L_{y,j})$, where J_{ECM} is the binding energy per unit length of an interface. Term $E_{\mathrm{Adhesion}} = J_a L_{\mathrm{contact}}$ in the Hamiltonian (10) corresponds to the cell-cell adhesion, where J_a is the binding energy per unit length of an interface and L_{contact} is the total contact area between cells. Term $E_{\mathrm{Perimeter}}$ defines an energy penalty function for dimensions of a cell deviating from the target values $L_{T_{x(y)}}$: $E_{\mathrm{Perimeter}} = \sum_{j=1}^{N} \lambda_x (L_{x,j} - L_{T_x})^2 + \lambda_y (L_{y,j} - L_{T_y})^2$, where λ_x and λ_y are Lagrange multipliers. Cells can move up or down gradients of both diffusible chemical signals (chemotaxis) and insoluble ECM molecules (haptotaxis) described by $E_{\mathrm{Field}} = \sum_{j=1}^{N} \mu c(\mathbf{R}_j, t) L_{x,j} L_{y,j}$, where μ is an effective chemical potential.

In this paper we neglect cell-cell adhesion $J_a = 0$ and the Hamiltonian (10) is reduced in 2D to the following expression:

$$E(X,t) = \left. \sum_{j=1}^{N} E(\mathbf{r}, \mathbf{L}, t) \right|_{\mathbf{r}=\mathbf{R}_j, \mathbf{L}=\mathbf{L}_j},$$

$$E(\mathbf{r}, \mathbf{L}, t) = 2J_{cm}(L_x + L_y) + \lambda_x (L_x - L_{T_x})^2 + \lambda_y (L_y - L_{T_y})^2 + \mu c(\mathbf{r}) L_x L_y. \tag{12}$$

B. Master equation for discrete cellular dynamics

We now represent CPM dynamics by using $P(\mathbf{r}, \mathbf{L}, t)$, probability density for any cell with its center of mass at $\mathbf{r}$ to have dimensions $\mathbf{L}$ at time t. Which means that we consider one-cell PDF $P(\mathbf{r}, \mathbf{L}, t)$ rather than N-cell PDF $P(X, t)$. Let $\epsilon\Delta r \times \epsilon\Delta r$ be the size of a lattice site with $\epsilon \ll 1$ and let vectors $\mathbf{e}_{1,2}$ indicate changes in x and y dimensions, $\mathbf{e}_1 = \Delta r(1,0), \mathbf{e}_2 = \Delta r(0,1)$. We normalize the total probability to the number of cells, $\int P(\mathbf{r}, \mathbf{L}, t) d\mathbf{r} d\mathbf{L} = N$. Excluded volume constraint implies that position $\mathbf{r}'$ and size $\mathbf{L}'$ of any neighboring cell should satisfy $2|x - x'| \geqslant L_x + L_x'$, $2|y - y'| \geqslant L_y + L_y'$.

Discrete stochastic cellular dynamics under these conditions is described by the following master equation [19]:

$$P(\mathbf{r},\mathbf{L},t+\Delta t)=\sum_{j=1}^{2}\left\{\left[\frac{1}{2}-\Omega_{j,l}\left(\mathbf{r}-\frac{\epsilon}{2}\mathbf{e}_j,\mathbf{L}+\epsilon\mathbf{e}_j;\mathbf{r},\mathbf{L},t\right)-\Omega_{j,r}\left(\mathbf{r}+\frac{\epsilon}{2}\mathbf{e}_j,\mathbf{L}+\epsilon\mathbf{e}_j;\mathbf{r},\mathbf{L},t\right)-\Omega_{j,l}\left(\mathbf{r}+\frac{\epsilon}{2}\mathbf{e}_j,\mathbf{L}-\epsilon\mathbf{e}_j;\mathbf{r},\mathbf{L},t\right)\right.\right.$$
$$\left.-\Omega_{j,r}\left(\mathbf{r}-\frac{\epsilon}{2}\mathbf{e}_j,\mathbf{L}-\epsilon\mathbf{e}_j;\mathbf{r},\mathbf{L},t\right)\right]P(\mathbf{r},\mathbf{L},t)+\Omega_{j,l}\left(\mathbf{r},\mathbf{L};\mathbf{r}+\frac{\epsilon}{2}\mathbf{e}_j,\mathbf{L}-\epsilon\mathbf{e}_j,t\right)P\left(\mathbf{r}+\frac{\epsilon}{2}\mathbf{e}_j,\mathbf{L}-\epsilon\mathbf{e}_j,t\right)$$
$$+\Omega_{j,r}\left(\mathbf{r},\mathbf{L};\mathbf{r}-\frac{\epsilon}{2}\mathbf{e}_j,\mathbf{L}-\epsilon\mathbf{e}_j,t\right)P\left(\mathbf{r}-\frac{\epsilon}{2}\mathbf{e}_j,\mathbf{L}-\epsilon\mathbf{e}_j,t\right)+\Omega_{j,l}\left(\mathbf{r},\mathbf{L};\mathbf{r}-\frac{\epsilon}{2}\mathbf{e}_j,\mathbf{L}+\epsilon\mathbf{e}_j,t\right)P\left(\mathbf{r}-\frac{\epsilon}{2}\mathbf{e}_j,\mathbf{L}+\epsilon\mathbf{e}_j,t\right)$$
$$\left.+\Omega_{j,r}\left(\mathbf{r},\mathbf{L};\mathbf{r}+\frac{\epsilon}{2}\mathbf{e}_j,\mathbf{L}+\epsilon\mathbf{e}_j,t\right)P\left(\mathbf{r}+\frac{\epsilon}{2}\mathbf{e}_j,\mathbf{L}+\epsilon\mathbf{e}_j,t\right)\right\}. \tag{13}$$

We incorporate dynamics into MC algorithm by defining the time step Δt. Individual biological cells experience diffusion through random fluctuations of their shapes. Diffusive coefficient can be measured experimentally (see, e.g., [24]). We choose Δt to match experimental diffusion coefficient. Equation (13) would determine a version of kinetic-dynamic MC algorithm (see, e.g., [39]) if Δt were to be allowed to fluctuate. For simplicity we assume that $\Delta t=$const. Also $\Omega_{j,l}(\mathbf{r},\mathbf{L};\mathbf{r}',\mathbf{L}',t)$ and $\Omega_{j,r}(\mathbf{r},\mathbf{L};\mathbf{r}',\mathbf{L}',t)$ denote probabilities of transitions from a cell of length L' and center of mass at $\mathbf{r}'$ to a cell of dimensions L and center of mass at $\mathbf{r}$. Subscripts l and r correspond to transitions by addition and/or removal of a row and/or colomn of pixels from the rear and/or lower and front and/or upper ends of a cell, respectively.

We define $\Omega_{j,l(r)}(\mathbf{r},\mathbf{L};\mathbf{r}',\mathbf{L}')\equiv T_{l(r)}(\mathbf{r},\mathbf{L};\mathbf{r}',\mathbf{L}')[1+\varphi_{j,l(r)}(\mathbf{r},\mathbf{L},t)]$ where $T_l(r)(\mathbf{r},\mathbf{L};\mathbf{r}',\mathbf{L}',t)$ denote probabilities of transitions from a cell of length $\mathbf{L}'$ and center of mass at $\mathbf{r}'$ to a cell of dimensions $\mathbf{L}$ and center of mass at $\mathbf{r}$ without taking into account excluded volume principle and cell-cell adhesion. According to the CPM we have that $T_l(\mathbf{r},\mathbf{L};\mathbf{r}',\mathbf{L}')=T_r(\mathbf{r},\mathbf{L};\mathbf{r}',\mathbf{L}')=\frac{1}{8}\Phi[E(\mathbf{r},\mathbf{L})-E(\mathbf{r}',\mathbf{L}')]$ where the factor of $1/8$ is due to the fact that there are potentially eight possibilities for increasing or decreasing of L_x and L_y (it means that we can add or remove pixels from any four sides of a rectangular cell). The second term $\varphi_{j,l(r)}(\mathbf{r},\mathbf{L},t)$ takes into account contact interactions between cells. It includes contributions from three possible types of stochastic jump processes due to contact interactions between cells: (a) a cell adheres to another one, (b) two adhered cells dissociate from each other due to membrane fluctuations, (c) membranes of two adhered cells are prevented from moving inside each other (due to excluded volume constraint) resulting in a negative sign of a contribution to a jump probability. If neither of these three processes happens at a given time step then $\varphi_{j,l(r)}(\mathbf{r},\mathbf{L},t)=0$.

C. Macroscopic limit of master equation and mean-field approximation

Equation (13) is not closed because one must know multicellular probability density to determine $\varphi_{j,l(r)}(\mathbf{r},\mathbf{L},t)$. One could use BBGKY-type hierarchy [40] similar to the one used in kinetic theory of gases, which expresses iteratively n-cell PDF through $n+1$-cell PDF with truncation at some order. This is however extremely difficult and ineffective for large n. Instead, we develop in Sec. IV a nonperturbative approach to derivation of Eqs. (1) and (2).

In previous work [17–19] we studied a macroscopic limit $\epsilon\ll 1$ of the master equation (13) by both neglecting contact interactions between cells [17,18] and including contact interactions in mean-field approximation [19], which represents simplest closure for BBGKY-type hierarchy. If contact interactions are neglected, then $\varphi_{j,l(r)}(\mathbf{r},\mathbf{L},t)=0$ and (13) yields in macroscopic limit $\epsilon\ll 1$ [17,18],

$$\partial_t P(\mathbf{r},\mathbf{L},t)=D_2(\nabla_{\mathbf{r}}^2+4\nabla_{\mathbf{L}}^2)P+8D_2\beta\lambda_x\partial_{L_x}(\tilde{L}_xP)$$
$$+8D_2\beta\lambda_y\partial_{L_y}(\tilde{L}_yP)+D_2\beta L_xL_y\mu\partial_{\mathbf{r}}(P\nabla_{\mathbf{r}}c),$$

$$\tilde{L}_x=\frac{1}{\lambda_x}\left(J_{cm}+\lambda_x(L_x-L_{T_x})+\frac{1}{2}L_y\mu c(\mathbf{r})\right),$$

$$\tilde{L}_y=\frac{1}{\lambda_y}\left(J_{cm}+\lambda_y(L_y-L_{T_y})+\frac{1}{2}L_x\mu c(\mathbf{r})\right),$$

$$D_2=\frac{(\epsilon\Delta r)^2}{16\Delta t},\quad \nabla_{\mathbf{r}}^2=\partial_x^2+\partial_y^2,\quad \nabla_{\mathbf{L}}^2=\partial_{L_x}^2+\partial_{L_y}^2. \tag{14}$$

Similar equation in the 1D case (with only coordinate x and length L_x present) was obtained in Ref. [17].

Mean-field approximation assumes the following decoupling of multicellular PDF:

$$P(X,t)=\left.N^{-N}\prod_{j=1}^{N}P(\mathbf{r},\mathbf{L}_j,t)\right|_{\mathbf{r}=\mathbf{R}_j}, \tag{15}$$

where factor N^{-N} is due to an assumed normalization $\int P(\mathbf{r},\mathbf{L},t)d\mathbf{r}d\mathbf{L}=N$. Mean-field approximation is exact if contact interactions are neglected. This holds since we assume above that chemotaxis depends only on average density (9). Equation (15) results in decoupling of Eq. (8) into independent equations for $P(\mathbf{R}_j,\mathbf{L}_j,t)$ for all j. This allows direct comparison of Eq. (14) with Eq. (8) and yields that

diffusion matrix has a diagonal form with main diagonal $D_2(1,1,4,4,1,1,4,4,\ldots)$ in 2D. In 1D and 3D cases numbers 1 and 4 repeat themselves with period 1 and 3, respectively.

Further comparison of (8), (9), and (14) leads to expression $\chi(\mathbf{L}_j)=L_{x,j}L_{y,j}\mu$ and

$$\Phi(X,t)=E(X,t), \tag{16}$$

where $E(X,t)$ is given by (12).

Taking into account contact cell-cell interaction (excluded volume) yields that Eq. (15) is not exact any more. Also, potential $\Phi(X,t)$ is given by (16) only if cells do not overlap. Otherwise $\Phi(X,t)=\infty$ according to (9).

D. Boltzmann-like distribution and macroscopic equation for cellular density

From Eqs. (12), (11), and (14) it follows that typical fluctuations of cell sizes $\tilde{L}_{x(y)}$ are determined by $\beta\lambda_{x(y)}\tilde{L}^2_{x(y)}\sim 1$. Suppose x_0 and y_0 are typical scales of $P(\mathbf{r},\mathbf{L},t)$ with respect to x and y. We assume that $\beta x_0^2\lambda_x \gg 1$ and $\beta y_0^2\lambda_y \gg 1$, meaning that $x_0 \gg \tilde{L}_x, y_0 \gg \tilde{L}_y$. We also assume that chemical field $c(\mathbf{r},t)$ is a slowly varying function of $\mathbf{r}$ on the scale of typical cell length meaning that $x_c/L_x \gg 1, y_c/L_y \gg 1$, where x_c and y_c are typical scales for variation of $c(\mathbf{r},t)$ in x and y. We also make an additional biologically relevant assumption that $4\lambda_x\lambda_y \gg \mu^2 c(\mathbf{r},t)^2$ meaning that change of typical cell size due to chemotaxis $\delta L^{(\mathrm{chemo})}_{x(y)}$ is small $|\delta L^{(\mathrm{chemo})}_{x(y)}| \ll L^{(\min)}_{x(y)}$.

If all these conditions are satisfied then we found by using both solutions of Eq. (14) and MC simulations of CPM with general initial conditions, that PDF $P(\mathbf{r},\mathbf{L},t)$ quickly converges in t at each spatial point $\mathbf{r}$ to the following Boltzmann-type form:

$$P(\mathbf{r},\mathbf{L},t)=P_{\mathrm{Boltz}}(\mathbf{r},\mathbf{L})p(\mathbf{r},t), \tag{17}$$

where

$$P_{\mathrm{Boltz}}(\mathbf{r},\mathbf{L})=\frac{1}{Z(\mathbf{r})}\exp(-\beta\Delta E_{\mathrm{length}}) \tag{18}$$

is a Boltzmann-type distribution depending on $\mathbf{r}$ and t only through $c(\mathbf{r},t)$, and

$$\Delta E_{\mathrm{length}}=E(\mathbf{r},\mathbf{L})-E_{\min}=\lambda_x L_x'^2+\lambda_y L_y'^2+L_x'L_y'\mu c(\mathbf{r}),$$

$$\mathbf{L}'=\mathbf{L}-\mathbf{L}^{(\min)}. \tag{19}$$

Here $E_{\min}=E(\mathbf{r},\mathbf{L}^{(\min)})$ is the minimal value of (12) achieved at $\mathbf{L}=\mathbf{L}^{(\min)}$,

$$E_{\min}=E(\mathbf{r},\mathbf{L}^{(\min)}),$$

$$L_x^{(\min)}=\frac{-4\lambda_y(J_{cm}-\lambda_x L_{T_x})+2(J_{cm}-\lambda_y L_{T_y})\mu c(\mathbf{r})}{4\lambda_x\lambda_y-\mu^2 c(\mathbf{r})^2},$$

$$L_y^{(\min)}=\frac{-4\lambda_x(J_{cm}-\lambda_y L_{T_y})+2(J_{cm}-\lambda_x L_{T_x})\mu c(\mathbf{r})}{4\lambda_x\lambda_y-\mu^2 c(\mathbf{r})^2}, \tag{20}$$

and $Z(\mathbf{r},t)=(2\epsilon\Delta r)^2\Sigma_{\mathbf{L}}\exp(-\beta\Delta E_{\mathrm{length}})\simeq\frac{2\pi}{\beta\sqrt{4\lambda_x\lambda_y-\mu^2c(\mathbf{r},t)^2}}$ is an asymptotic formula for a partition function as $\epsilon\to 0$. (See [17] for details about convergence rate for the case without contact interactions.)

Also, under these conditions we can use Eq. (17) to integrate Eq. (14) over $\mathbf{L}$ which results in the following evolution equation for the the cellular probability density $p(\mathbf{r},t)=\int[P(\mathbf{r},\mathbf{L},t)]d\mathbf{L}$ [17,18]:

$$\partial_t p=D_2\nabla^2_{\mathbf{r}}p-\chi_0\nabla_{\mathbf{r}}\cdot[p\nabla_{\mathbf{r}}c(\mathbf{r},t)], \tag{21}$$

where $\chi_0=-D_2\mu\beta L_x^{(0)}L_y^{(0)}$ and

$$L_x^{(0)}=L_{T_x}-J_{cm}/\lambda_x,$$

$$L_y^{(0)}=L_{T_y}-J_{cm}/\lambda_y \tag{22}$$

correspond to $L^{(\min)}_{x(y)}$ from (20) provided we neglect chemotaxis.

Equation (21) together with (3) form a closed set which coincides with the classical Keller-Segel system [8]. It has a finite time singularity (collapse) and was extensively used for modeling aggregation of bacterial colonies [13,14]. We show below that near singularity contact interactions between cells could prevent collapse.

Equations similar to (21) and (3) for the case of contact interactions (excluded volume constraint) has been obtained in the mean-field approximation [19]. These equations significantly slow down collapse in comparison with (23) and (3). They still have collapsing solutions if initial density is not small. (These equations are applicable only for small densities.)

IV. BEYOND MEAN-FIELD APPROXIMATION AND REGULARIZATION OF COLLAPSE

The main purpose of this paper is to derive macroscopic equations which do not have collapse (blow up of solutions in finite time) for arbitrary initial densities and are in good agreement with microscopic stochastic simulations for large cellular densities. This requires one to go beyond mean-field approximation.

We conclude from the preceding section that random changes of cellular lengths result in random walks of centers of mass of cells during the time between cell "collisions." Significant simplification in comparison with Eq. (6) is that we have now explicit dependence on spatial coordinate $\mathbf{r}$ but not on $\mathbf{L}$. Below we use notation $(L_x^{(0)},L_y^{(0)})$ for the average size of a cell neglecting change of that size due to chemotaxis. In CPM $(L_x^{(0)},L_y^{(0)})$ are given by (22). If we neglect chemotaxis (i.e., set $c=0$) then during time between collision cell probability density $p(\mathbf{r},t)$ is described by linear diffusion equation which follows from Eq. (21):

$$\partial_t p=D_2\nabla^2_{\mathbf{r}}p. \tag{23}$$

In a similar way, two-cell probability density $p(\mathbf{r}_1,\mathbf{r}_2,t)$ is described by linear diffusion in two independent variables $\mathbf{r}_1,\mathbf{r}_2$,

$$\partial_t p=D_2(\nabla^2_{\mathbf{r}_1}+\nabla^2_{\mathbf{r}_2})p. \tag{24}$$

$p(\mathbf{r}_1,\mathbf{r}_2,t)$ represents probability density of cells 1 and 2 having centers of mass at $\mathbf{r}_1$ and $\mathbf{r}_2$ at time t, respectively. After making change of variables

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2, \quad \mathbf{R} = (\mathbf{r}_1 + \mathbf{r}_2)/2, \tag{25}$$

where variable $\mathbf{r}$ describes relative motion of cells and variable $\mathbf{R}$ describes motion of "center of mass" of two cells, Eq. (24) has the following form:

$$\partial_t p = 2D_2 \nabla_{\mathbf{r}}^2 p + \frac{D_2}{2} \nabla_{\mathbf{R}}^2 p. \tag{26}$$

Each collision involving cell 1 or 2 modifies both $p(\mathbf{r},t)$ and $p(\mathbf{r}_1,\mathbf{r}_2,t)$. In other words, it effects random walk of each colliding cell.

We describe first the effect of collisions due to excluded volume constraint between cells in the 1D case. Consider a pair of neighboring cells which in 1D always remain neighbors. We assume that at the time of collision two colliding cells have the same size. Generally this is not true because sizes of cells continuously fluctuate with length-scale $\delta L_x \sim 1/(\beta^{1/2}\lambda_x^{1/2})$. However, we assume as before that these fluctuations are small $|\delta L_x| \ll L_x^{(0)}$ which justifies our approximation. Collision is defined as two cells being in direct contact at a given moment of time and one of them trying to penetrate into another. Collison is prevented by excluded volume constraint. In the continuous limit $\epsilon \to 0$ each collision takes infinitesimally small time. After collision, cells move away from each other so they are not in direct contact anymore. Instead of explicitly describing each of these collisions we use an assumption that two cells are identical and view each collision as exchange of positions of two cells [41,42]. From this point of view cells do not collide at all but simply pass through each other. They both experience random walk as point objects (cells) without collisions according to Eqs. (25) and (26) in the domain free from other cells (free domain). (The volume of such free domain per cell, which has dimension of length in 1D, is on average $[1 - L_x^{(0)} p(x,t)]/p(x,t)$.)

This means that we are considering collective diffusion of cellular density instead of trajectories of individual cells (see, e.g., Ref. [25] for a review). Another type of diffusion is self-diffusion which describes mean-square displacement of an individual cell as a function of time [25]. Self-diffusion might be important for describing propagation of one cell type through space occupied by cells of another type. In this paper we consider only collective diffusion.

There is an important qualitative difference between dependence of collective diffusion and self-diffusion on volume fraction. It is known from the study of diffusion of colloids in the hard-sphere model that diffusion coefficient of self-diffusion decreases with increase of the volume fraction due to the cage effect [25]. Namely, hard spheres prevent a given hard sphere from moving at distances larger than a typical distance between nearest hard spheres. Effect of volume fraction increase on collective diffusion is the opposite. Frequent collisions of hard spheres in case of large volume fractions result in hard spheres moving on average in the direction of negative density gradient. In that case relative displacement of neighboring hard spheres is small but change of the density of hard spheres in the direction of the density gradient might be large. Collective diffusion also means that one uses a long wavelength limit of the density fluctuations [43]. Decrease of self-diffusion and increase of collective diffusion with increase of the volume fraction has been shown in the limit of diluted gas of diffusing hard spheres in Refs. [44–46]. It was also observed in experiments described in, e.g., Refs. [47,48] and many subsequent studies (see, e.g., [25] for a review). Notice that increase of collective diffusion of hard spheres in colloidal suspension is significantly reduced by hydrodynamic interactions [25]. In this paper we consider biological cells which move at random with their shapes also randomly fluctuating which is different from the mechanism of diffusion in colloidal suspensions. However, we expect that previous results for diffusing hard spheres will qualitatively agree with results for the collective and self-diffusion of cells.

Note also that the macroscopic model derived in this paper has qualitative analogies with a hydrodynamic limit of interacting particles and cellular automata (see, e.g., [49,50]).

Because we study collective diffusion and adopt a point of view that each collision is an exchange of positions of two cells, we conclude that while trajectories of cells in a free domain are continuous, positions of cells in physical space change instantaneously at each collision by L_x for the cell colliding from the left and by $-L_x$ for the cell colliding from the right. The effective rate of cell diffusion is enhanced as free space becomes smaller with growth of cellular volume fraction. Let us assume that at the initial time $t=t_0$ centers of mass of two cells are separated by an average distance $1/p(x,t)$ and that these two cells collide for the first time (meaning that the preceding collision of any of these two cells involved collision with another cell). This yields that their centers of mass are separated by distance $L_x^{(0)}$ at $t=t_0$. If a moving reference frame is set at one of the cells then the other will experience random walk with doubled diffusion coefficient $2D_2$ [as seen from Eq. (26)], where D_2 is a diffusion coefficient of each cell in the stationary reference frame. Relative motion of two cells in a moving reference frame corresponds to random walk of a pointwise cell with diffusion coefficient $2D_2$.

Consider continuous random walk in free space. The number of returns to the initial position $x=L_x^{(0)}$ during any finite time interval after t_0 is infinite meaning that the number of collisions between a given pair of cells in physical space is infinite. However, successive collisions effectively cancel each other since they change positions of cells by $\pm L_x^{(0)}$. What really matters is whether the total number of collisions is even or odd. For an even number of collisions the total contribution of collisions is zero. While for an odd number of collisions the contribution to the flux of probability for the left cell (at time t_0) $\propto L_x^{(0)}$ and for the right cell $\propto -L_x^{(0)}$. Because the total number of collisions is infinite, the probabilities of having even or odd number of collisions are equal to $1/2$. While the average distance between center of mass of cells is $1/p(x,t)$, the average distance between boundaries of two neighboring cells in the physical domain is

$$\Delta x = 1/p(x,t) - L_x^{(0)}. \tag{27}$$

When separation between surfaces of two cells after collision reaches Δx from (27) we determine that "extended collision" between the pair of cells is over. Namely, two cells are not

closer to each other than to other neighboring cells anymore. Therefore, probability of them colliding with each other is not higher any more than probability of them colliding with other cells. This extended collision includes infinitely many "elementary" collisions but its final contribution depends only on whether the total number of such collisions is even or odd.

To find the average time of extended collision we solve the diffusion equation in the moving frame

$$\partial_t p_m = 2D_2 \nabla_x^2 p_m \tag{28}$$

with reflecting boundary condition $\partial_x p_m(L_x^{(0)},t)=0$ at $x=L_x^{(0)}$ and absorbing boundary condition $p_m(L_x^{(0)}+\Delta x,t)=0$ at $x=L_x^{(0)}+\Delta x$. Reflecting boundary condition means that the cell does not cross point $x=L_x^{(0)}$. Instead of crossing $x=L_x^{(0)}$ cells exchange positions at each collision. Absorbing boundary condition corresponds to the "escape point" of a cell from the extended collision. Initial condition is $p_m(x,t_0)=\delta(x-L_x^{(0)})$ which is defined by initial zero distance between surfaces of two cells. Solution of Eq. (28) with these initial and boundary conditions yields the mean first-passage time (escape time) T_{esc}, which is equal to the extended collision time in our case. As shown in the Appendix, the extended collision time in 1D is given by

$$T_{\mathrm{esc},1} = \frac{(\Delta x)^2}{4D_2}. \tag{29}$$

Consider mean-square displacements $(\Delta x)_1^2$ and $(\Delta x)_2^2$ of cells 1 and 2, respectively, in the stationary frame during extended collision. Using (25) we obtain that $(\Delta x)_1^2+(\Delta x)_2^2=\frac{(\Delta x)^2}{2}+2(\Delta X)^2$, where $X=\mathbf{R}$ in 1D. Because of the symmetry between cells 1 and 2 the following relation holds $(\Delta x)_1^2=(\Delta x)_2^2=\frac{(\Delta x)^2}{4}+(\Delta X)^2$. Center of mass of two cells experiences random walk with diffusion coefficient $D_2/2$ [see Eq. (26)] and $(\Delta X)^2=T_{\mathrm{esc},1}D_2=\frac{(\Delta x)^2}{4}$, $(\Delta x)_1^2=(\Delta x)_2^2=\frac{(\Delta x)^2}{2}$.

Now recall that with probability 1/2 cells exchange positions during extended collision meaning that in that expression for $(\Delta x)_{1,2}^2$ with probability 1/2 term $(\Delta x)^2$ should be replaced by $(2L_x^{(0)}+\Delta x)^2$ resulting in the following total mean-square displacement for cell 1 or 2:

$$(\Delta x)_{\mathrm{total}}^2 = \frac{1}{2}\left(\frac{(\Delta x)^2}{2}+\frac{(2L_x^{(0)}+\Delta x)^2}{2}\right). \tag{30}$$

From Eqs. (27), (29), and (30) we obtain effective (nonlinear) diffusion coefficient for the cell probability density

$$D_{1,\mathrm{effective}} = \frac{(\Delta x)_{\mathrm{total}}^2}{2T_{\mathrm{esc},1}} = D_2\left(\frac{1+(L_x^{(0)})^2p^2}{(1-L_x^{(0)}p)^2}\right). \tag{31}$$

In simulations described below we typically use a not very large number of cells N. In case of finite N each given cell can collide with $N-1$ other cells resulting in extra factor $(N-1)/N$ and allows one to rewrite the effective diffusion coefficient (31) as follows:

$$D_{1,\mathrm{effective}} = D_2\left(\frac{1+(1-N^{-1})(L_x^{(0)})^2p^2}{(1-(1-N^{-1})L_x^{(0)}p)^2}\right). \tag{32}$$

The effective diffusion (32) results in 1D nonlinear diffusion equation

$$\partial_t p = D_2\nabla_{\mathbf{r}}\cdot\left(\frac{1+(1-N^{-1})(L_x^{(0)})^2p^2}{(1-(1-N^{-1})L_x^{(0)}p)^2}\nabla_{\mathbf{r}}p\right) - \chi_0\nabla_{\mathbf{r}}\cdot[p\nabla_{\mathbf{r}}c(\mathbf{r},t)]. \tag{33}$$

Here we added chemotaxis term from Eq. (21) which is well justified provided $x_c \gg \Delta x$ (see also the preceding section). For example, typical spatial scale of developmental gradients is $x_c \sim 0.1-0.5$ mm [51]. Which can be large in comparison with a typical size of eukaryotic cell 2–100 μm.

Now consider the 2D case. First we look at cells with a disk-shaped form instead of cells of rectangular shapes (i.e., cellular shapes fluctuating around a disk) with average diameter $L_x^{(0)}$. Assume (as in 1D case) that t_0 is the time of the first collision between two given cells. Previous collisions of each of these cells involved cells other then these two. We also average over angles relative to the center of one of these two cells which implies that cells collide at time $t=t_0$ with equal probability at every angle. In that approximation there is rotational symmetry in the moving frame and all variables depend only on the radial variable $r\equiv|\mathbf{r}|$ but do not depend on angular variables. Change of variables in Fokker-Planck equation $\partial_t p=2D_2\nabla_{\mathbf{r}}^2 p$ results in

$$\partial_t \tilde{p} = -2D_2\nabla_r\frac{\tilde{p}}{r}+2D_2\nabla_r^2\tilde{p}, \tag{34}$$

where $\tilde{p}\equiv rp$. Equation (34) is equivalent to the 1D Fokker-Planck equation with potential $U(r)=-2D_2\ln(r)$. As shown in the Appendix, the extended collision time in the 2D case is given by

$$T_{\mathrm{esc},2} = \frac{\Delta x(2L_x^{(0)}+\Delta x)}{8D_2} - \frac{[L_x^{(0)}]^2}{4D_2}\ln\left(1+\frac{\Delta x}{L_x^{(0)}}\right). \tag{35}$$

We assume again that during extended collision time cells on average span entire space, i.e.,

$$\frac{\pi}{4}(L_x^{(0)}+\Delta x)^2 = p^{-1}. \tag{36}$$

Using Eqs. (30), (35), and (36) we obtain the effective diffusion coefficient in 2D (here and below the average disk diameter is $L\equiv L_x^{(0)}$),

$$D_{2,\mathrm{effective}} = \frac{(\Delta x)_{total}^2}{4T_{\mathrm{esc},2}} = D_2\left(\frac{1+\pi L^2p}{1-\pi L^2p+\pi L^2p\ln(\pi L^2p)}\right), \tag{37}$$

which, after modification to include effect of a finite N similar to the one used in 1D case, results in

$$D_{2,\text{effective}} = D_2\left(\frac{1+(1-N^{-1})\pi L^2 p}{1-(1-N^{-1})\pi L^2 p+(1-N^{-1})\pi L^2 p\ln(\pi L^2 p)}\right), \tag{38}$$

Eq. (38) yields the following nonlinear diffusion equation for cells fluctuating around disk-shaped form

$$\partial_t p = D_2\nabla_{\mathbf{r}}\cdot\left(\frac{1+(1-N^{-1})\pi L^2 p}{1-(1-N^{-1})\pi L^2 p+(1-N^{-1})\pi L^2 p\ln(\pi L^2 p)}\nabla_{\mathbf{r}} p\right)-\chi_0\nabla_{\mathbf{r}}\cdot[p\nabla_{\mathbf{r}}c(\mathbf{r},t)]. \tag{39}$$

Notice that both 1D effective diffusion coefficient (32) and 2D effective diffusion coefficient (38) depend only on the volume fraction φ ($\varphi=L_x^{(0)}p$ in 1D and $\varphi=\pi L^2 p$ in 2D). Based on that we propose that the effective diffusion coefficient for 2D rectangles also depends only on the volume fraction $\varphi=L_x^{(0)}L_y^{(0)}p$, which results in the nonlinear diffusion equation for cells of rectangular shape

$$\partial_t p = D_2\nabla_{\mathbf{r}}\cdot\left(\frac{1+(1-N^{-1})L_x^{(0)}L_y^{(0)}p}{1-(1-N^{-1})L_x^{(0)}L_y^{(0)}p+(1-N^{-1})L_x^{(0)}L_y^{(0)}p\ln(L_x^{(0)}L_y^{(0)}p)}\nabla_{\mathbf{r}} p\right)-\chi_0\nabla_{\mathbf{r}}\cdot[p\nabla_{\mathbf{r}}c(\mathbf{r},t)]. \tag{40}$$

Numerical simulations in the next section confirm that Eq. (40) agrees very well with the MC simulations of microscopic dynamics.

In the macroscopic limit, the $N\gg 1$ factor $(1-N^{-1})$ is replaced by 1 in Eqs. (33), (39), and (40), which yields Eqs. (1) and (2).

If the volume fraction $\varphi\to 1$ then nonlinear diffusion in Eqs. (1) and (2) diverges. Thus, we do not expect any blow up of the solutions of Eqs. (1)–(3). This is in contrast to blow up in Eqs. (21) and (3) in 2D. Global existence of Eqs. (1) and (2) together with Eq. (3) can be studied in a way similar to [52].

V. NUMERICAL RESULTS AND APPLICATION TO VASCULOGENESIS

A. One-dimensional cell motion

Figure 3 demonstrates simulation of 1D motion of cells represented in the form of fluctuating rods. Numerical solution of the continuous model was obtained using a pseudospectral scheme. For both the CPM simulation and numerics of the continuous model we used periodic boundary conditions, simulation time was $t_{\text{end}}=200$ and values of other parameters were chosen as follows: $\Delta r=1$, $L_{T_x}=L_{T_x}=3$, $\lambda_x=\lambda_y=1.5$, $\epsilon=0.01J_{cm}=2$, $\beta=15$, $N=8$, $\mu=0$. Simulation was performed on the spatial domain $0\leqslant x<100$ and the initial cell density distribution was $p(\mathbf{r},0)=k_0 e^{-[(x-50)/10]^4}$ with k_0 determined by normalization $N=8$. Figure 3 shows simulations of the CPM (curve *a*), numerical solutions of the macroscopic model without excluded volume interactions (21) (curve *b*), macroscopic Eq. (41) (curve *c*), and macroscopic Eq. (33) (curve *d*). Here,

$$\partial_t p = D_2\nabla_{\mathbf{r}}\cdot\left(\frac{1}{(1-(1-N^{-1})L_x^{(0)}p)^2}\nabla_{\mathbf{r}} p\right)-\chi_0\nabla_{\mathbf{r}}\cdot[p\nabla_{\mathbf{r}}c(\mathbf{r},t)], \tag{41}$$

was derived from the equation of state for the 1D hard rod fluid [53] which allows one to determine collective diffusion coefficient from static structure factor and compressibility (see, e.g., [25,43]).

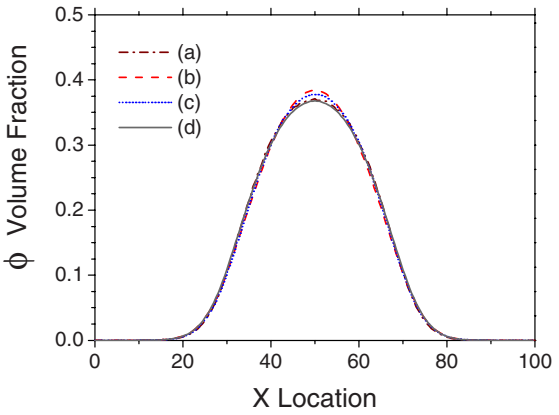

FIG. 3. (Color online) Volume fraction $\varphi=L_x^{(0)}p(x,t_{\text{end}})$ for the 1D cellular motion as a function of x. Curve *a*, Monte Carlo simulations of the CPM. Curve *b*, solution of Eq. (21). Curve *c*, solution of Eq. (41). Curve *d*, solution of the Eq. (33).

Figure 3 demonstrates that solution of Eq. (33) is in much better agreement with CPM than both Eqs. (21) and (41). While the difference between the CPM simulation and solutions of Eqs. (21) and (41) is small but clearly exceeds the error in MC simulations. The difference between MC simulation and solution of Eq. (33) is within an accuracy of the MC simulations.

The difference between CPM and Eq. (41) is due to the fact that the equation of state for the 1D hard rod fluid was calculated in Ref. [53] from grand canonical partition function [54] while diffusion of cells is a nonequilibrium phenomenon resulting in the corrections to the equilibrium partition function. Numerous attempts have been made to describe dynamics of interacting Brownian particles (see [55] and references therein). Note also that the difference between CPM and Eq. (21) results is not so dramatic in 1D as in 2D because in 1D Keller-Segel model does not support collapse [14].

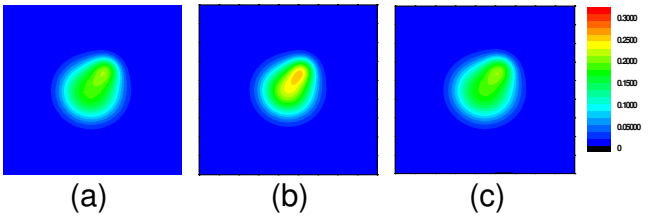

FIG. 4. (Color online) $p(x,y,t)$ in 2D as a function of (x,y) for (a) Monte Carlo simulation of CPM, (b) Eq. (21), and (c) Eq. (40).

B. Two-dimensional cell motion with chemotaxis

Figures 4 and 5 demonstrate a very good agreement between typical CPM simulation and numerical solution of the continuous model Eq. (40). Both simulations were performed on a square domain $0 \leqslant x,y \leqslant 100$ over simulation time $t_{\text{end}}=400$. Simulation parameters' values are as follows: $\Delta r=1$, $L_{T_x}=L_{T_y}=4.4$, $\lambda_x=\lambda_y=1.5$, $J_{cm}=2$, $\beta=15$, $\mu=0.1$, $\epsilon=0.01$, and $N=15$. Chemical field concentration is chosen in the form of $c(x,y)=0.2(1-e^{-[(x-65)^2+(y-60)^2]/144})$ and does not depend on time. Initial cell density is chosen in the form of $p_0(x,y)=k_0 e^{-\{[(x-50)^2+(y-50)^2]/100\}^5}$, where k_0 is a constant that normalizes the integral of the cell density to $N=15$. Numerical solution of the continuous model has been obtained using pseudo spectral scheme with 200×200 Fourier modes. A large number of CPM simulations (600 000) have been run on a parallel computer cluster to guarantee a representative statistical ensemble.

Numerics for Eq. (21) significantly differs from CPM simulations indicating that excluded volume interactions are important in the chosen range of values of parameters.

C. Application to vasculogenesis

To test the model we studied the effect of the chemical production rate on the network formation. Our previous results [19] were limited to relatively small chemical production rate $a \leqslant 0.2$ in Eq. (3) because otherwise chemotaxis resulted in cellular density which was too high for applying mean-field approximation. Here we use macroscopic equations (40) and (3) and compare numerical results with the

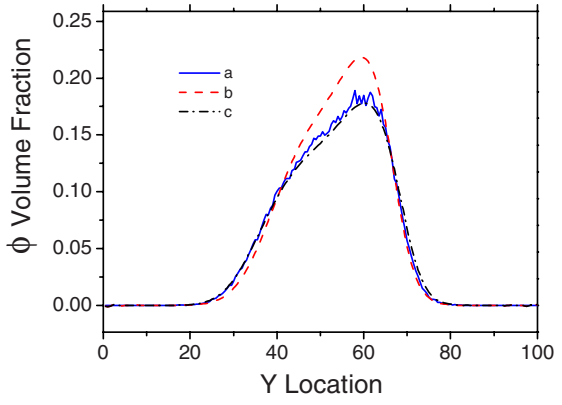

FIG. 5. (Color online) Cross sections of the volume fraction $\varphi = L_x^{(0)} L_y^{(0)} p(x_0,y,t)$ in 2D for (a) CPM simulation, (b) Eq. (21), and (c) Eq. (40). All cross sections are at the same position, $x_0=57.75$.

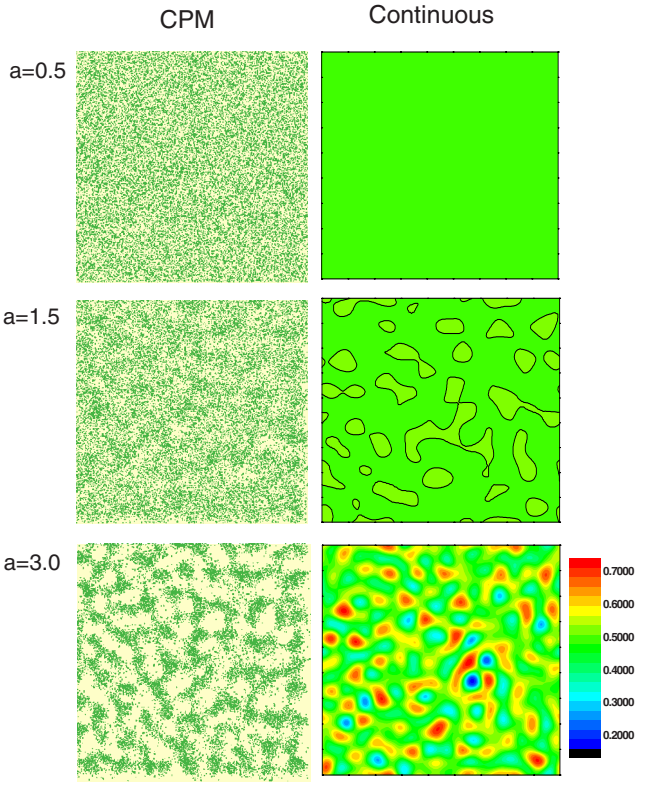

FIG. 6. (Color online) Simulation of early vascular network formation with different chemical production rates α. $\Delta r=1$, $L_{T_x}=L_{T_y}=0.6$, $\lambda_x=\lambda_y=1.5$, $J_{cm}=0.002$, $\beta=15$, $\mu=-0.1$, $D_c=0.5$, $\gamma=0.014$, $\Delta t_c=\epsilon^2\Delta t=0.01$, $\epsilon=0.1$, $t_{\text{end}}=60$. In the Monte Carlo CPM simulation $N=15\,000$ cells were randomly distributed in a domain $0 \leqslant x,\ y \leqslant 100$ with initial chemical field at zero. In the continuous model, a uniform initial cell density distribution with 5% random fluctuation was used. Scale bar is in units of the volume fraction φ.

CPM simulations. Figure 6 shows a series of simulations with different chemical production rates $a=0.5$, $a=1.5$, and $a=3.0$. Simulations start with initially dilute populations of cells moving on a substrate in a chemotactic field, subject to an excluded volume constraint. Both CPM and continuous model simulation results indicate that stripe patterns are obtained for high chemical production rates. Higher chemical production rate, by strengthening the chemotaxis and cell aggregation process, eventually leads to higher pattern density with smaller average distance between two neighboring stripes. Structures of the resulting networks obtained using discrete and continuous models are very similar to each other as well as to the one obtained experimentally for a population of endothelial cells cultured on a Matrigel film [2].

VI. SUMMARY AND DISCUSSION

We have derived macroscopic continuous Eqs. (1) and (2) coupled with Eq. (3) for describing evolution of cellular density in the chemical field, from microscopic cellular dynamics. Microscopic cellular model includes many individual cells moving on a substrate by means of random fluctuations of their shapes, chemotactic and contact cell-cell interactions. Contrary to classical Keller-Segel model, solutions of

the obtained Eqs. (1)–(3) do not collapse in finite time and can be used even when relative volume occupied by cells φ is quite large. This makes them much more biologically relevant then earlier introduced systems. We compared numerics for macroscopic equations with Monte Carlo simulations of microscopic cellular dynamics for the CPM and demonstrated a very good agreement for $\varphi \simeq 0.3$. For larger density we expect transition to a glass state [25]. It was demonstrated that combination of the CPM and derived continuous model can be applied to studying network formation in early vasculogenesis. We are currently working on an important problem in vasculogenesis of simulating self-diffusion [25] of one type of cells through dense population of other types of cells.

This work was partially supported by NSF Grants No. DMS 0719895 and No. IBN-0344647.

APPENDIX: CALCULATION OF MEAN ESCAPE TIME T_{esc}

For convenience of the reader we provide here a derivation of extended collision time T_{esc} which is similar to the calculation of mean escape time in Ref. [20]. To find T_{esc} we use a backward Fokker-Planck equation

$$\partial_{t_1} p(x_2,t_2|x_1,t_1) = U'(x_1)\partial_{x_1} p(x_2,t_2|x_1,t_1) - D_0 \nabla^2_{x_1} p(x_2,t_2|x_1,t_1) \tag{A1}$$

(see, e.g., [20]), where $p(x_2,t_2|x_1,t_1)$ is the conditional probability density for transition from (x_1,t_1) to (x_2,t_2). D_0 is the diffusion coefficient and $U(x)$ is an arbitrary potential. In the 1D case $U(x)=0$ and $D_0=2D_2$.

The probability that a cell remains inside an interval $(0,\Delta x)$ at time t, provided it was located at x at $t=t_0$, is given by $G(x,t) \equiv \int_0^{\Delta x} p(x',t|x,t_0)dx'$. Using the stationarity of the random walk $p(x',t|x,t_0)=p(x',0|x,t_0-t)$ we obtain from (A1) that

$$\partial_t G = -U'(x)\partial_x G + D_0 \nabla^2_x G. \tag{A2}$$

Mean escape time $T_{esc}(x)$ for a cell located at x at $t=t_0$ is

$$T_{esc}(x) = -\int_{t_0}^{\infty} (t-t_0)\partial_t G(x,t)dt = \int_{t_0}^{\infty} G(x,t)dt. \tag{A3}$$

Integrating Eq. (A2) over t from t_0 to ∞ and using initial normalization $G(x,t_0)=1$ results in

$$-U'(x)\partial_x T_{esc}(x) + D_0 \nabla^2_x T_{esc}(x) = -1. \tag{A4}$$

Reflecting and absorbing boundary conditions for p result in similar boundary conditions for $T_{esc}(x)$,

$$\partial_x T_{esc}(x)|_{x=L_x^{(0)}} = 0, \quad T_{esc}(x)|_{x=L_x^{(0)}+\Delta x} = 0, \tag{A5}$$

which allows us to solve boundary value problem (A4) and (A5) explicitly,

$$T_{esc}(x) = D_0^{-1} \int_x^{L_x^{(0)}+\Delta x} \exp[D_0^{-1}U(x')]dx' \times \int_{L_x^{(0)}}^{x'} \exp[-D_0^{-1}U(x'')]dx''. \tag{A6}$$

Initial condition implies that $T_{esc,1}=T_{esc}(L_x^{(0)})$ and yields extended collision time in 1D given by Eq. (29).

Radially symmetric Fokker-Planck equation (34) is equivalent to the 1D Fokker-Planck equation with potential $U(r)=-2D_2 \ln(r)$. Backward Fokker-Planck equation (A1) with $x=r$, and the same $U(r)$ yields from Eq. (A6), an extended collision time (mean escape time) given by Eq. (35).

[1] A. Gamba, D. Ambrosi, A. Coniglio, A. de Candia, S. DiTalia, E. Giraudo, G. Serini, L. Preziosi, and F. Bussolino, Phys. Rev. Lett. **90**, 118101 (2003).
[2] G. Serini, D. Ambrosi, E. Giraudo, A. Gamba, L. Preziosi, and F. Bussolino, EMBO J. **22**, 1771 (2003).
[3] P. Carmeliet, Nat. Med. **6**, 389 (2000).
[4] R. M. H. Merks, S. Brodsky, M. Goligorksy, S. Newman, and J. Glazier, Dev. Biol. **289**, 44 (2006).
[5] P. A. Rupp, A. Czirok, and C. D. Little, Development **131**, 2887 (2004).
[6] A. Szabo, E. D. Perryn, and A. Czirok, Phys. Rev. Lett. **98**, 038102 (2007).
[7] S. Christley, M. Alber, and S. Newman, PLOS Comput. Biol. **3**, e76 (2007).
[8] E. F. Keller and L. A. Segel, J. Theor. Biol. **30**, 225 (1971).
[9] W. Alt, J. Math. Biol. **9**, 147 (1980).
[10] H. G. Othmer and A. Stevens, SIAM J. Appl. Math. **57**, 1044 (1997).
[11] A. Stevens, SIAM J. Appl. Math. **61**, 172 (2000).
[12] T. J. Newman and R. Grima, Phys. Rev. E **70**, 051916 (2004).
[13] M. A. Herrero and J. J. L. Velázquez, Math. Ann. **306**, 583 (1996).
[14] M. P. Brenner, P. Constantin, L. P. Kadanoff, A. Shenkel, and S. C. Venkataramani, Nonlinearity **12**, 1071 (1999).
[15] S. Turner, J. A. Sherratt, K. J. Painter, and N. J. Savill, Phys. Rev. E **69**, 021910 (2004).
[16] T. J. Newman, Math. Biosci. Eng. **2**, 611 (2005).
[17] M. Alber, N. Chen, T. Glimm, and P. M. Lushnikov, Phys. Rev. E **73**, 051901 (2006).
[18] M. Alber, N. Chen, T. Glimm, and P. M. Lushnikov, in *Single-Cell-Based Models in Biology and Medicine, Series: Mathematics and Biosciences in Interaction*, edited by A. R. A. Anderson, M. A. J. Chaplain, and K. A. Rejniak (Birkhauser Verlag Basel, Switzerland, 2007).
[19] M. Alber, N. Chen, P. M. Lushnikov, and S. A. Newman, Phys. Rev. Lett. **99**, 168102 (2007).
[20] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences* (Springer-Verlag, New York, 2004).
[21] H. G. Döbereiner, B. J. Dubin-Thaler, J. M. Hofman, H. S. Xenias, T. N. Sims, G. Giannone, M. L. Dustin, C. H. Wiggins, and M. P. Sheetz, Phys. Rev. Lett. **97**, 038102 (2006).

[22] J. Coelho Neto and O. N. Mesquita, Philos. Trans. R. Soc. London, Ser. A **366**, 319 (2008).
[23] U. Agero, C. H. Monken, C. Ropert, R. T. Gazzinelli, and O. N. Mesquita, Phys. Rev. E **67**, 051904 (2003).
[24] J. P. Rieu, J.-P. Rieu, A. Upadhyaya, J. A. Glazier, N. B. Ouchi, and Y. Sawada, Biophys. J. **79**, 1903 (2000).
[25] R. B. Jones and P. N. Pusey, Annu. Rev. Phys. Chem. **42**, 137 (1991).
[26] T. Hofer, J. A. Sherratt, and P. K. Maini, Physica D **85**, 425 (1995).
[27] M. Luca, A. Chavez-Ross, L. Edelstein-Keshet, and A. Mogilner, Bull. Math. Biol. **65**, 693 (2003).
[28] R. Grima, Phys. Rev. Lett. **95**, 128103 (2005).
[29] T. M. Cickovski, C. Huang, R. Chaturvedi, T. Glimm, H. Hentschel, M. S. Alber, J. A. Glazier, S. A. Newman, and J. A. Izaguirre, IEEE/ACM Trans. Comput. Biol. Bioinf. **2**, 273 (2005).
[30] T. M. Cickovski, K. Aras, M. Swat, R. Merks, T. Glimm, H. Hentschel, M. S. Alber, J. A. Glazier, S. A. Newman, and J. A. Izaguirre, Comput. Sci. Eng. **9**, 50 (2007).
[31] R. Chaturvedi, C. Huang, B. Kazmierczak, T. Schneider, J. A. Izaguirre, T. Glimm, H. G. Hentschel, J. A. Glazier, S. A. Newman, and M. S. Alber, J. R. Soc., Interface **2**, 237 (2005).
[32] M. A. Knewitz and J. C. M. Mombach, Comput. Biol. Med. **36**, 59 (2006).
[33] O. Sozinova, Y, Jiang, D, Kaiser, and M, S. Alber, Proc. Natl. Acad. Sci. U.S.A. **103**, 17255 (2006).
[34] Z. Xu, N. Chen, M. M. Kamocka, E. D. Rosen, and M. S. Alber, J. R. Soc., Interface **5**, 705 (2008).
[35] *Single Cell-Based Models in Biology and Medicine*, Mathematics and Biosciences in Interaction, edited by A. R. A. Anderson, M. A. J. Chaplain, and K. A. Rejmau (Birkhauser, Switzerland, 2007).
[36] S. Newman, S. Christley, T. Glimm, H. Hentschel, B. Kazmierczak, Y. Zhang, J. Zhu, and M. S. Alber, Curr. Top Dev. Biol. **81**, 311 (2008).
[37] Y. Jiang, J. Pjesivac-Grbovic, C. Cantrell, and J. P. Freyer, Biophys. J. **89**, 3884 (2005).
[38] Y. Wu, Y. Jiang, D. Kaiser, and M. Alber, PLOS Comput. Biol. **3**, e253 (2007).
[39] A. B. Bortz, M. H. Kalos, and J. L. Lebowitz, J. Comput. Phys. **17**, 10 (1975).
[40] D. A. McQuarrie, *Statistical Mechanics* (University Science Books, New York, 2000).
[41] H. Rost, Lect. Notes Math. **1059**, 127 (1984).
[42] M. Bodnar and J. J. L. Velázquez., Math. Methods Appl. Sci. **28**, 1757 (2005).
[43] R. Gomer, Rep. Prog. Phys. **53**, 917 (1990).
[44] B. J. Ackerson and L. Fleishman, J. Chem. Phys. **76**, 2675 (1982).
[45] S. Hanna, W. Hess, and R. Klein, Physica A **111**, 181 (1982).
[46] R. B. Jones and G. S. Burfield, Physica A **111**, 562 (1982); **111**, 577 (1982).
[47] M. M. Kops-Werkhoven and H. M. Fijnaut, J. Chem. Phys. **77**, 2242 (1982).
[48] P. N. Pusey and W. van Megen, J. Phys. (Paris) **44**, 285 (1983).
[49] J. L. Lebowitz, E. Presutti, and H. Spohn, J. Stat. Phys. **51**, 841 (1988).
[50] H. Spohn, *Large Scale Dynamics of Interacting Particles* (Springer-Verlag, New York, 1991).
[51] A. D. Lander, Q. Nie, and F. Y. M. Wan, Dev. Cell **2**, 785 (2002).
[52] Z. A. Wang and T. Hillen, Chaos **17**, 037108 (2007).
[53] J. K. Percus, J. Stat. Phys. **15**, 505 (1976).
[54] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, 3rd ed. (Butterworth-Heinemann, London, 1984).
[55] B. Kim and K. Kawasaki, J. Stat. Mech.: Theory Exp. (2008) P02004.