

Supplementary Material

Oleksii M Matsiaka · Catherine J
Penington · Ruth E Baker · Matthew J
Simpson

Received: date / Accepted: date

Contents

1	Comparing results from the two-dimensional and one-dimensional stochastic models	2
2	Derivation of the $p_1(x, t)$ and $p_2(x, y, t)$ governing equations for the one-dimensional model	5
3	Discretisation scheme for the one-dimensional MFA model	10
4	Discretisation scheme for the one-dimensional KSA model	11

1 Comparing results from the two-dimensional and one-dimensional stochastic models

In this section we present results from solving the Langevin equations (Equation (2), Main paper) for the scratch assay geometry in two dimensions, and compare them with the corresponding solutions to the Langevin equations in one dimension. Employing the vector notation introduced in the Main paper we can write the Langevin equations for the two-dimensional model as follows,

$$\frac{d\mathbf{u}^{(i)}}{dt} = \sum_{j \neq i} \mathbf{F}_{ij} + \boldsymbol{\xi}_i, \quad i = 1, \dots, N, \quad (1)$$

where $\mathbf{u}^{(i)}$ is the position of the i th agent, \mathbf{F}_{ij} is the interaction force between agent i and agent j , $\boldsymbol{\xi}_i$ is the stochastic force acting on the i th agent, and N is the number of agents in the simulation.

The initial positions of agents in both the two-dimensional and one-dimensional simulations are chosen from the following distribution

$$\alpha(x) = \begin{cases} 12.5 \times 10^{-3}, & 0 \mu\text{m} \leq x < 600 \mu\text{m}, \\ 0, & 600 \mu\text{m} \leq x \leq 1400 \mu\text{m}, \\ 12.5 \times 10^{-3}, & 1400 \mu\text{m} < x \leq 2000 \mu\text{m}. \end{cases} \quad (2)$$

In the case of the two-dimensional simulation, the vertical positions of agents are placed at random to form a homogeneous density distribution. It ensures that there is no density gradient in the vertical direction.

We fix all model parameters to be the same in both the two-dimensional and one-dimensional simulations to avoid any other possible source of variability. The size of agents is fixed to a typical size of skin cell size of $25 \mu\text{m}$ (Simpson et al. 2013). We choose the number of agents in each one-dimensional simulation to be 15. The number of agents in each two-dimensional simulation

is 534. The number of agents in both simulations is chosen so that the initial nondimensional density outside of the scratched region is $p_1/C = 0.625$, where C is the carrying capacity density of agents with diameter $25 \mu\text{m}$. The size of the domain in the two-dimensional simulations is $2000 \mu\text{m} \times 1400 \mu\text{m}$, which is typical for scratch assay experiments (Figure 1, Main paper). Periodic boundary conditions are imposed in both the two-dimensional and one-dimensional simulations.

Results are summarised in Figure 1. To compare the one-dimensional and two-dimensional simulations we average results from the two-dimensional model in the vertical direction to obtain a one-dimensional agent density distribution. We show that averaged two-dimensional results demonstrate very similar population level behaviour compared to simple one-dimensional model provided all model parameters and initial densities are fixed.

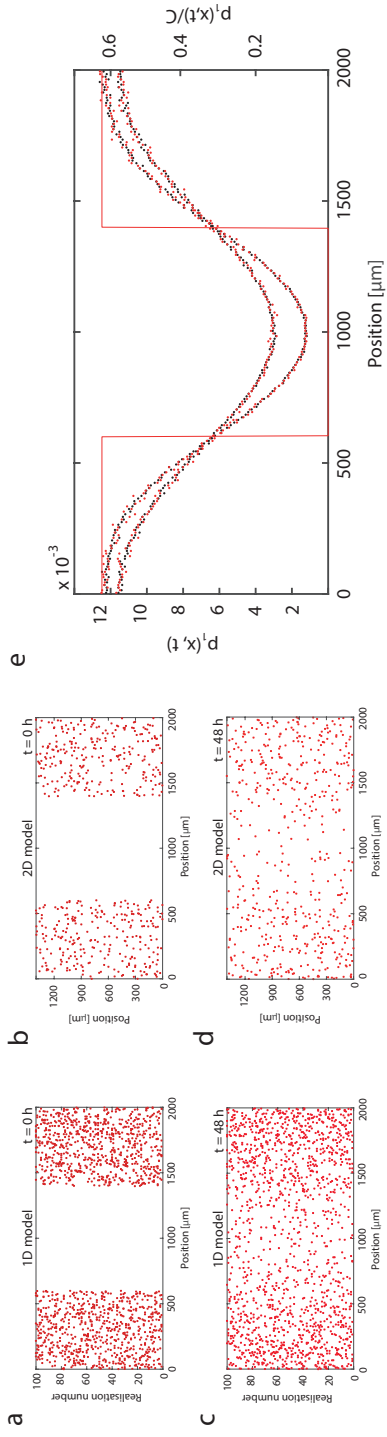


Fig. 1. Comparison of the averaged agent density data from the two-dimensional stochastic simulations and its one-dimensional counterpart. Snapshots in (a) and (c) show 100 identically prepared realisations of the stochastic model, given by Equation (2) (Main paper), at $t = 0$ and 48 h. Snapshots in (b) and (d) represent one realisation of the stochastic model, given by Equation (1), at $t = 0$ and 48 h. Results in (e) show the averaged agent density profiles using an ensemble of 10^5 identically prepared simulations of the one-dimensional model (black dots) with binsize $10 \mu\text{m}$. These results are compared with the averaged ensemble of 10^3 identically prepared realisations of the two-dimensional model, Equation (1), at $t = 24$ and 48 h. Equations (2) (Main paper) and Equation (1) are solved with $\Delta t = 2 \times 10^{-2}$ h. The red solid line shows initial conditions given by Equation (2).

2 Derivation of the $p_1(x, t)$ and $p_2(x, y, t)$ governing equations for the one-dimensional model

In this section we derive the equations for the evolution of the density $p_1(x, t)$ and the pair-correlation density $p_2(x, y, t)$ presented in Section 2.2 (Main paper). These equations represent the first two levels of the full hierarchy of equations that incorporates all spatial moments (Middleton et al. 2014; Mat-siaka et al. 2017). To begin, we introduce an agent density function, $\rho(x, t)$. The evolution of the agent density is given by continuity equation (van Kam-pen, 1975):

$$\frac{\partial \rho(x, t)}{\partial t} = - \sum_{i=1}^N \frac{\partial J_i}{\partial x^{(i)}}, \quad (3)$$

where J_i is the component of the total flux of agents associated with the agent i , and N is the total number of agents. If $J_i = \rho dx^{(i)}/dt$, we have

$$\frac{\partial \rho(x, t)}{\partial t} = - \sum_{i=1}^N \frac{\partial}{\partial x^{(i)}} \left(\sum_{j \neq i} F_{ij} + \xi_i \right) \rho(x, t), \quad (4)$$

where $\sum_{j \neq i} F_{ij} + \xi_i$ is the right hand side of Equation (2) (Main paper) and ξ_i is treated as a fixed parameter (van Kampen, 1975).

Suppose we treat ξ_i as a random variable. We can obtain a solution to Equation (4) with the initial conditions

$$\rho(x, 0) = \prod_{i=1}^N \delta(x - x^{(i)}(0)) = \delta^{(N)}(x - x^{(i)}(0)), \quad (5)$$

where δ is the Dirac delta function, and $x^{(i)}(0)$ is the initial position of the i th agent. For any initial density distribution, $\rho(x, 0)$, we assume that we can obtain the average over many different realisations of the stochastic force ξ_i , $\langle \rho(x, t) \rangle_{\xi}$.

Let $P_1^i(x, t)$ be the probability density function (PDF) for an individual agent. Then the one-agent PDF is given by (van Kampen, 1976)

$$P_1^i(x, t) = \langle \langle \rho(x, t) \rangle_{\xi} \rangle_{\text{IC}}, \quad (6)$$

where the average is taken over different realisations of the initial distribution and the random variable ξ_i . Since the averaged local density can be expressed as

$$\langle \langle \rho(x, t) \rangle_{\xi} \rangle_{\text{IC}} = \langle \langle \delta^{(N)}(x - x^{(i)}(t)) \rangle_{\xi} \rangle_{\text{IC}}, \quad (7)$$

we can define the one-agent PDF in the following way,

$$P_1^i(x, t) = \langle \langle \delta^{(N)}(x - x^{(i)}(t)) \rangle_{\xi} \rangle_{\text{IC}}, \quad (8)$$

where $x^{(i)}(t)$ is the position of the i th agent at time t , as given by Equation (2) (Main paper).

Similarly, we define the two-agent PDF as

$$P_2^{ij}(x, y, t) = \langle \langle \delta^{(N)}(x - x^{(i)}(t)) \delta^{(N)}(y - y^{(j)}(t)) \rangle_{\xi} \rangle_{\text{IC}}. \quad (9)$$

The evolution of $P_1^i(x, t)$ and $P_2^{ij}(x, y, t)$ are given by (García-Ojalvo and Sancho, 1999)

$$\frac{\partial P_1^i(x, t)}{\partial t} = D \Delta P_1^i(x, t) - \frac{\partial}{\partial x} (f_i P_1^i(x, t)), \quad (10)$$

$$\frac{\partial P_2^{ij}(x, y, t)}{\partial t} = D \Delta P_2^{ij}(x, y, t) - \frac{\partial}{\partial x} (f_i P_2^{ij}(x, y, t)) - \frac{\partial}{\partial y} (f_j P_2^{ij}(x, y, t)), \quad (11)$$

where the total force f_i acting on agent i can be expressed as

$$f_i = \sum_{j \neq i} F_{ij}. \quad (12)$$

Combining Equations (8), (10) and (12), and taking the convolution of the interaction force and a δ function centred at $y^{(j)}$, we obtain

$$\begin{aligned} \frac{\partial P_1^i(x, t)}{\partial t} &= D\Delta P_1^i(x, t) \\ &\quad - \nabla \left\langle \sum_{j \in \mathcal{L}, j \neq i} \int_{\Omega} F(x^{(i)} - y, t) \delta(x - x^{(i)}(t)) \delta(y - y^{(j)}(t)) \, dy \right\rangle, \end{aligned} \quad (13)$$

where Ω denotes the domain and \mathcal{L} is the set of agents. The second term on the right hand side of Equation (13) is an advection term. Combining Equations (9) and (13), and interchanging summation and integration, we obtain

$$\frac{\partial P_1^i(x, t)}{\partial t} = D\Delta P_1^i(x, t) - \nabla \int_{\Omega} F(x - y, t) \sum_{j \in \mathcal{L}, j \neq i} P_2^{ij}(x, y, t) \, dy, \quad (14)$$

where, from this point forward, we drop the subscript i on $x^{(i)}$.

To make the transition from individual level behaviour in a discrete simulation to population level dynamics, we define the following quantities,

$$p_1(x, t) = \frac{1}{N} \sum_{i \in \mathcal{L}} P_1^i(x, t), \quad (15)$$

$$p_2(x, y, t) = \frac{1}{N(N-1)} \sum_{i \in \mathcal{L}} \sum_{j \in \mathcal{L}, j \neq i} P_2^{ij}(x, y, t), \quad (16)$$

where $p_1(x, t)$ is the normalised one-agent density distribution, and $p_2(x, y, t)$ is the density-density correlation function that captures correlations in agent positions.

To proceed, we sum over the index i in Equation (14) and apply the definitions given in Equations (15)-(16) to obtain

$$\frac{\partial p_1(x, t)}{\partial t} = D\Delta p_1(x, t) - (N-1) \nabla \left(\int_{\Omega} F(x - y, t) p_2(x, y, t) \, dy \right). \quad (17)$$

To derive an evolution equation for $p_2(x, y, t)$ we begin with the two-agent Fokker-Planck equation,

$$\frac{\partial P_2^{ij}(x, y, t)}{\partial t} = D\Delta P_2^{ij}(x, y, t) - \frac{\partial}{\partial x} \left(f_i P_2^{ij}(x, y, t) \right) - \frac{\partial}{\partial y} \left(f_j P_2^{ij}(x, y, t) \right), \quad (18)$$

where indices i and j denote arbitrary agents in population.

Adopting the interaction force law, Equation (5) (Main paper), using the definition of the two-agent PDF, as given by Equation (9), and evaluating the required convolutions, allows us to rewrite Equation (18) as

$$\begin{aligned} \frac{\partial P_2^{ij}(x, y, t)}{\partial t} &= D\Delta P_2^{ij}(x, y, t) \\ &- \frac{\partial}{\partial x} \left\langle F(x - y, t) \delta(x - x^{(i)}(t)) \delta(y - y^{(j)}(t)) \right\rangle \\ &- \frac{\partial}{\partial y} \left\langle F(y - x, t) \delta(x - x^{(i)}(t)) \delta(y - y^{(j)}(t)) \right\rangle \\ &- \frac{\partial}{\partial x} \left\langle \sum_{g \in \mathcal{L}, g \neq i, j} \int_{\Omega} F(x - z, t) \delta(x - x^{(i)}(t)) \delta(y - y^{(j)}(t)) \delta(z - z^{(g)}(t)) dz \right\rangle \\ &- \frac{\partial}{\partial y} \left\langle \sum_{g \in \mathcal{L}, g \neq i, j} \int_{\Omega} F(y - z, t) \delta(x - x^{(i)}(t)) \delta(y - y^{(j)}(t)) \delta(z - z^{(g)}(t)) dz \right\rangle, \end{aligned} \quad (19)$$

where the second and third terms on the right hand side of Equation (19) represent interactions between agents i and j , the fourth and fifth terms on the right hand side of Equation (19) represent interactions between agents i and j and other agents within the population.

The three-agent normalised density functions can be defined as

$$p_3(x, y, z, t) = \frac{1}{N(N-1)(N-2)} \sum_i \sum_{j \neq i} \sum_{g \neq i, j} P_3^{ijg}(x, y, z, t). \quad (20)$$

We therefore require a definition for the three-agent PDF, $P_3^{ijg}(x, y, z, t)$. Similar to Equation (9),

$$P_3^{ijg}(x, y, z, t) = \langle \delta(x - x^{(i)}(t)) \delta(y - y^{(j)}(t)) \delta(z - z^{(g)}(t)) \rangle. \quad (21)$$

To proceed we divide Equation (19) by $N(N - 1)$, and combine Equations (19)-(21), summing over the indices i and j , to obtain an expression for the evolution of $p_2(x, y, t)$

$$\begin{aligned} \frac{\partial p_2(x, y, t)}{\partial t} = & D\Delta p_2(x, y, t) \\ & - \frac{\partial}{\partial x} \left(F(x - y) p_2(x, y, t) \right) \\ & - \frac{\partial}{\partial y} \left(F(y - x) p_2(x, y, t) \right) \\ & - (N - 2) \frac{\partial}{\partial x} \int_{\Omega} F(x - z, t) p_3(x, y, z, t) dz \\ & - (N - 2) \frac{\partial}{\partial y} \int_{\Omega} F(y - z, t) p_3(x, y, z, t) dz. \end{aligned} \quad (22)$$

This procedure can be repeated to yield the hierarchy of $N - 1$ coupled integro partial differential equations and one Fokker-Planck equation. At any arbitrary level $d \in [1, N - 1]$ of this hierarchy, the d -density, p_d , depends on the higher order density, p_{d+1} . This means that full hierarchy of equations is intractable for analysis. Consequently, we invoke two approximations to simplify the hierarchy of density equations: (i) the standard mean-field approximation (MFA); and (ii) the Kirkwood superposition approximation (KSA).

The MFA approximates the pair correlation function $p_2(x, y, t)$ in terms of $p_1(x, t)$ and $p_1(y, t)$:

$$p_2(x, y, t) = p_1(x, t)p_1(y, t). \quad (23)$$

This expression implies that the probability of finding an agent at x is statistically independent of the probability of finding an agent at y . The KSA

approximates the three-agent normalised density function as the product of two-agent density functions, and can be written as (Singer 2004; Middleton et al. 2014)

$$p_3(x, y, z, t) = \frac{p_2(x, y, t) p_2(x, z, t) p_2(y, z, t)}{p_1(x, t) p_1(y, t) p_1(z, t)}. \quad (24)$$

3 Discretisation scheme for the one-dimensional MFA model

In this section we present the discretisation scheme used to solve Equation (10) (Main paper) governing the evolution of the agent density $p_1(x, t)$ in one dimension. The MFA-based continuum equation is given by

$$\frac{\partial p_1(x, t)}{\partial t} = D \Delta p_1(x, t) - (N - 1) \nabla(p_1(x, t) V(x, t)), \quad (25)$$

where

$$V(x, t) = \int_{\Omega} F(x - y, t) p_1(y, t) dy, \quad (26)$$

and Ω is the integration domain.

We introduce the following quantities

$$\beta(x, y, t) = F(x - y, t) p_1(y, t) = f_0 Z(r, t) \operatorname{sgn}(x - y) p_1(y, t), \quad (27)$$

$$\begin{aligned} I_l &= p_1(x_l, t) \int_{\Omega} \beta(x_l, y, t) dy \\ &= p_1(x_l, t) \frac{h}{2} \sum_s \left[\beta(x_l, y_{s+1}, t) + \beta(x_l, y_s, t) \right] + O(h^2), \end{aligned} \quad (28)$$

where the trapezoidal rule with a stepsize h is used for numerical integration, and the indices l and s denote equally-spaced spatial mesh nodes. Here, the stepsize is chosen to be $h = 4 \mu\text{m}$, while the binsize used in discrete results presented in the Main paper is $10 \mu\text{m}$.

Using the definitions presented in Equations (27)-(28), we now apply the method of lines to Equation (25) and obtain the system of ordinary differential

equations

$$\frac{dp_i^1}{dt} = \frac{D}{h^2} [p_{i+1} - 2p_i + p_{i-1}] - (N-1) \frac{1}{2h} [I_{i+1} - I_{i-1}], \quad (29)$$

where the index i denotes a spatial mesh node. This system of ordinary differential equations is solved using the first order explicit Euler method with a constant time step Δt . This expression is valid for an arbitrary interior node and, since we apply periodic boundary conditions, it can be easily adapted on the boundaries of the domain.

4 Discretisation scheme for the one-dimensional KSA model

We now write down the discretisation scheme used to solve Equation (14) (Main paper) governing the evolution of $p_2(x, y, t)$ in the KSA-based framework. Note that we only solve the equation for $p_2(x, y, t)$ and obtain $p_1(x, t)$ by numerical integration, using

$$p_1(x, t) = \int_{\Omega} p_2(x, y, t) dy. \quad (30)$$

The governing equation that we consider is as follows,

$$\begin{aligned} \frac{\partial p_2(x, y, t)}{\partial t} &= D \Delta p_2(x, y, t) \\ &- f_0 \left(\frac{\partial}{\partial x} - \frac{\partial}{\partial y} \right) \left(Z(|x-y|, t) \operatorname{sgn}(x-y) p_2(x, y, t) \right) \\ &- f_0 (N-2) \frac{\partial}{\partial x} \int_{\Omega} Z(|x-z|, t) \operatorname{sgn}(x-z) \frac{p_2(x, y, t) p_2(x, z, t) p_2(y, z, t)}{p_1(x, t) p_1(y, t) p_1(z, t)} dz \\ &- f_0 (N-2) \frac{\partial}{\partial y} \int_{\Omega} Z(|y-z|, t) \operatorname{sgn}(y-z) \frac{p_2(x, y, t) p_2(x, z, t) p_2(y, z, t)}{p_1(x, t) p_1(y, t) p_1(z, t)} dz. \end{aligned} \quad (31)$$

In order to present the numerical method as briefly as possible, we define the following quantities

$$\gamma(x, y, t) = f_0 Z(|x - y|, t) \operatorname{sgn}(x - y) p_2(x, y, t), \quad (32)$$

$$\phi(x, y, z, t) = f_0 Z(|x - z|, t) \operatorname{sgn}(x - z) \frac{p_2(x, z, t) p_2(y, z, t)}{p_1(z, t)}, \quad (33)$$

$$\psi(x, y, z, t) = f_0 Z(|y - z|, t) \operatorname{sgn}(y - z) \frac{p_2(x, z, t) p_2(y, z, t)}{p_1(z, t)}. \quad (34)$$

Upon substituting Equations (32)-(34) into Equation (31), the evolution equation for $p_2(x, y, t)$ becomes

$$\begin{aligned} \frac{\partial p_2(x, y, t)}{\partial t} &= D \Delta p_2(x, y, t) - \frac{\partial}{\partial x} \gamma(x, y, t) + \frac{\partial}{\partial y} \gamma(x, y, t) \\ &\quad - (N - 2) \frac{\partial}{\partial x} \left[\frac{p_2(x, y, t)}{p_1(x, t) p_1(y, t)} \int_{\Omega} \phi(x, y, z, t) \, dz \right] \\ &\quad - (N - 2) \frac{\partial}{\partial y} \left[\frac{p_2(x, y, t)}{p_1(x, t) p_1(y, t)} \int_{\Omega} \psi(x, y, z, t) \, dz \right]. \end{aligned} \quad (35)$$

We now introduce the discretised quantities

$$\begin{aligned} Q_{l,k} &= \frac{p_2(x_l, y_k, t)}{p_1(x_l, t) p_1(y_k, t)} \int_{\Omega} \phi(x_l, y_k, z, t) \, dz \\ &= \frac{p_2(x_l, y_k, t)}{p_1(x_l, t) p_1(y_k, t)} \frac{h}{2} \sum_s \left[\phi(x_l, y_k, z_{s+1}, t) + \phi(x_l, y_k, z_s, t) \right] + O(h^2), \end{aligned} \quad (36)$$

$$\begin{aligned} Y_{l,k} &= \frac{p_2(x_l, y_k, t)}{p_1(x_l, t) p_1(y_k, t)} \int_{\Omega} \psi(x_l, y_k, z, t) \, dz \\ &= \frac{p_2(x_l, y_k, t)}{p_1(x_l, t) p_1(y_k, t)} \frac{h}{2} \sum_s \left[\psi(x_l, y_k, z_{s+1}, t) + \psi(x_l, y_k, z_s, t) \right] + O(h^2), \end{aligned} \quad (37)$$

where the trapezoidal rule with stepsize h on an equally spaced mesh is used to approximate the integrals. We now apply the method of lines to Equation

(35) and obtain the following system of equations

$$\begin{aligned} \frac{dp_{ij}^2}{dt} = & \frac{D}{h^2} \left[p_{i+1,j} - 2p_{ij} + p_{i-1,j} + p_{i,j+1} - 2p_{ij} + p_{i,j-1} \right] \\ & - \frac{1}{2h} \left[\gamma_{i+1,j} - \gamma_{i-1,j} \right] + \frac{1}{2h} \left[\gamma_{i,j+1} - \gamma_{i,j-1} \right] \\ & - (N-2) \frac{1}{2h} \left[Q_{i+1,j} - Q_{i-1,j} \right] - (N-2) \frac{1}{2h} \left[Y_{i,j+1} - Y_{i,j-1} \right], \quad (38) \end{aligned}$$

where indices i, j denote spatial mesh nodes, and $\gamma_{l,m} = \gamma(x_l, y_m, t)$. This expression is valid for an arbitrary interior node and, since we apply periodic boundary conditions, it can be easily adapted on the boundaries of the domain. This systems is then solved using the first order explicit Euler method with a constant time step of duration Δt .

References

1. García-Ojalvo J, Sancho JM (1999). Noise in spatially extended systems. Springer New York. New York. USA.
2. Van Kampen NG (1976). Stochastic differential equations. Physics Reports. 24c (3): 171-228.
3. Matsiaka OM, Penington CJ, Baker RE, Simpson MJ (2017). Continuum approximations for lattice-free multi-species models of collective cell migration. Journal of Theoretical Biology. 422: 1-11.
4. Middleton AM, Fleck C, Grima R (2014). A continuum approximation to an off-lattice individual-cell based model of cell migration and adhesion. Journal of Theoretical Biology. 359: 220-232.
5. Simpson MJ, Treloar KK, Binder BJ, Haridas P, Manton KJ, Leavesley DI, McElwain DLS, Baker RE (2013). Quantifying the roles of cell motility and cell proliferation in a circular barrier assay. Journal of the Royal Society Interface. 10: 20130007.
6. Singer A (2004). Maximum entropy formulation of the Kirkwood superposition approximation. The Journal of Chemical Physics. 121: 3657-3666.