



# Multi-species simple exclusion processes

Matthew J. Simpson<sup>\*</sup>, Kerry A. Landman, Barry D. Hughes

*Department of Mathematics and Statistics, University of Melbourne, Victoria 3010, Australia*

## ARTICLE INFO

### Article history:

Received 28 May 2008

Received in revised form 16 October 2008

Available online 31 October 2008

### PACS:

87.10.-e

87.10.Ed

87.10.Hk

### Keywords:

Exclusion process

Multi-species

Multi-scale modelling

## ABSTRACT

A motility mechanism based on a simple exclusion process, where the movement of discrete agents on a lattice is either unbiased (symmetric) or biased (asymmetric) is considered. Estimates of diffusivities from tracking data do not describe the population-level response of the system. This mismatch between the individual-level and population-level behaviour can be resolved by averaging the individual-level mechanism in terms of an expected site occupancy. New insight into simple exclusion processes is obtained by representing the system as a series of interacting subpopulations. This formalism leads to a system of nonlinear advection–diffusion equations which can be interpreted in terms of the agent fluxes. These interactions have consequences for both agent-based modelling and continuum modelling in cell biology, such as tracking subpopulations of cells within a total cell population.

Crown Copyright © 2008 Published by Elsevier B.V. All rights reserved.

## 1. Introduction

Interacting motility mechanisms are widely used to represent various physical and biological processes such as the movement of crowds [1], embryonic morphogenesis [2] and traffic movement [3]. These motility mechanisms involve individual agents making decisions about potential movements that are dictated by local-level cues that could arise from interactions between the individuals [3–8] or external influences [9–11,13,14].

Our interest in modelling interacting motility mechanisms originated from studying the motility of cells in biological systems. Individual-level (microscopic) models are naturally suited to describe cell motility, since experimental time-lapse data give direct observations of complex motility behaviours [15,16], which are readily incorporated into individual-level models. Furthermore, we are interested to model chick-quail graft experiments where a total population of motile agents is composed of interacting subpopulations [17].

A critical aspect of modelling with individual-level motility mechanisms is to consider how the individual-level mechanism gives rise to the population-level (macroscopic) response [8,10,13,18–20]. This kind of multi-scale modelling is advantageous relative to either individual-level or population-level modelling conducted in isolation. In cell biology, for example, the relationship between individual-level signals (chemical, molecular or environmental cues) and the resulting population-level outcome (tissue and organ-level behaviour) is a critical aspect of understanding various biological systems [8,9,21].

In this work we consider a model of motility based on an unbiased (symmetric) or biased (asymmetric) simple exclusion process [5,22–24]. Averaged agent density profiles are related to the solution of continuum equations. Our work describes and explains discrepancies between individual-level and population-level properties of the same mechanism. A population of agents is represented as a system of interacting subpopulations. This leads to nonstandard advection–diffusion equations. Rewriting these continuum models in terms of agent fluxes gives new physical insight into simple exclusion processes.

<sup>\*</sup> Corresponding author. Tel.: +61 3 83446517; fax: +61 3 83444599.

E-mail address: [m.simpson@ms.unimelb.edu.au](mailto:m.simpson@ms.unimelb.edu.au) (M.J. Simpson).

## 2. Discrete motility mechanism

The motility mechanism is implemented on a lattice with dimensions  $L_x \times L_y$ . The lattice discretisation is square with spacing  $\Delta$ . Each site is either empty or occupied by a single agent and during each time step, of length  $\tau$ , all agents are given an opportunity to move with probability  $P$ . A motile agent at  $(x, y)$  steps to  $(x, y \pm \Delta)$  each with probability  $1/4$ , or to  $(x \pm \Delta, y)$  with probability  $(1 \pm \rho)/4$ . Here  $|\rho| \leq 1$  is a parameter controlling the motility bias. Setting  $\rho = 0$  gives an unbiased motility mechanism.

If, for any particular attempted move, the target site is occupied, that move is aborted. This is called a simple exclusion process [5]. This mechanism imposes a strong form of agent–agent interaction and differs radically from a standard noninteracting random walk, where agents do not interact and multiple agents can reside on the same site [18,19,25].

Our aim is to explore how the simple exclusion process can be used to represent a system of interacting subpopulations. To motivate this work we will first consider a system with just one subpopulation of agents.

## 3. Single species motility

Various properties of simple exclusion processes for a single population will motivate our analysis of multi-species systems. We first consider known results for unbiased motility ( $\rho = 0$ ) that relate the discrete mechanism with estimates of a population-level diffusivity [26–29]. Previous simulations, where each lattice site is occupied with probability  $C$ , have used trajectory data to obtain a relationship between the normalized lattice density  $C$  and diffusivity [7], giving

$$D(C) = \frac{P}{4}(1 - C). \quad (1)$$

Solutions of the diffusion equation using this estimate of the diffusivity will be compared with agent density profiles extracted from the discrete simulations. Two methods to average the simulation data will be used. If  $N_n^m(i, j)$  is the occupancy of site  $(i, j)$  after  $n$  steps of the  $m$ th realisation, then for  $M$  realisations starting from the same initial condition, the average occupancy of site  $(i, j)$  is given by

$$\langle N_n(i, j) \rangle = \frac{1}{M} \sum_{m=1}^M N_n^m(i, j). \quad (2)$$

It is also convenient to estimate the occupancy of any site within a particular column by constructing a double average, denoted  $C_n(i)$ , evaluated over the height of each column  $L_y$  and over  $M$  realisations that start from the same initial condition,

$$C_n(i) = \frac{1}{L_y M} \sum_{m=1}^M \sum_{j=1}^{L_y} N_n^m(i, j). \quad (3)$$

Both  $\langle N_n(i, j) \rangle$  and  $C_n(i)$  lie in the closed interval  $[0, 1]$ .

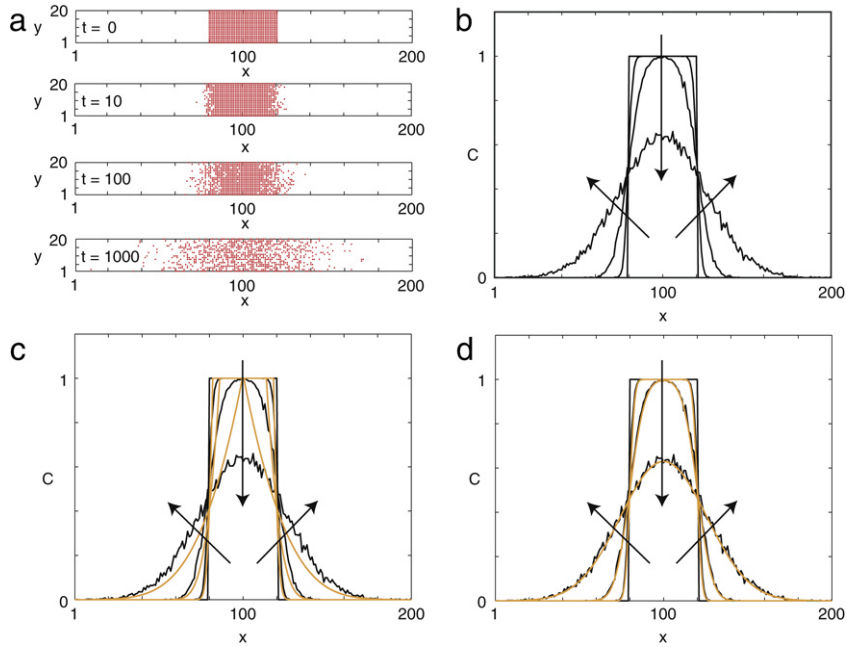
Simulations of unbiased motility are used to generate density profiles. Agents are initially placed with every site between  $80 \leq x \leq 120$  occupied, and then the system is allowed to evolve with time. Reflecting boundary conditions are imposed at  $x = 1$  and  $x = L_x$ , and periodic boundary conditions are imposed at  $y = 1$  and  $y = L_y$ . Eq. (3) is used to calculate the column density at various times. Snapshots of the agent distributions during a single realisation and the corresponding column density profiles  $C_n(i)$  are given in Fig. 1a–b respectively.

The discrete density profiles are compared with solutions of the diffusion equation

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial x} \left( D(C) \frac{\partial C}{\partial x} \right). \quad (4)$$

Following Eq. (1), we assume  $D(C) = P(1 - C)/4$  and solve Eq. (4) numerically with boundary and initial conditions that match the discrete simulations (Fig. 1b). A finite difference method with constant grid spacing  $\delta x$  and implicit Euler stepping with constant time steps  $\delta t$  is used to solve Eq. (4). Picard linearisation, with tolerance  $\epsilon$ , is used to solve the nonlinear equations. All one-dimensional results presented in this work correspond to  $\delta x = 0.1$ ,  $\delta t = 0.05$  and  $\epsilon = 1 \times 10^{-4}$ . Comparing discrete and continuum profiles (Fig. 1c) shows a poor match. Instead, solutions of a linear diffusion model with  $D(C) = P/4$  (Fig. 1d) provide a perfect match. This result, although known [26], is counter-intuitive, since individual-level data gave a direct estimate of  $D(C)$ , given by Eq. (1), yet this estimate does not predict the population-level response. These differences can be reconciled by averaging the discrete mechanism to give a continuum description.

In any single realisation of the discrete algorithm, it is incorrect to treat the occupancy of adjacent sites as independent. However, when averaging over a large collection of identically prepared systems, effective independence of the occupancy status of adjacent sites is a natural approximation. By invoking this independence assumption, a conservation of occupancy



**Fig. 1.** Unbiased agents spread from  $80 \leq x \leq 120$  on a lattice with  $L_x = 200$  and  $L_y = 20$ . (a) Simulation with  $P = 1$ . (b) Averaged column density profiles. (c) Density profiles (black) do not compare well with the solution of Eq. (4) with  $D(C) = (1 - C)/4$  (orange). (d) Density profiles (black) compare very well with the solution of Eq. (4) with  $D(C) = 1/4$  (orange). Arrows indicate the direction of increasing time, and profiles are given at  $t = 0, 10, 100$  and  $1000$ . All discrete results correspond to  $\Delta = \tau = 1$  and are averaged using Eq. (3) with  $M = 40$ . (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

statement for site  $(i, j)$  is constructed by considering transitions into and out of that site between step  $n$  and  $n + 1$ :

$$\begin{aligned}
 & \langle N_{n+1}(i, j) \rangle - \langle N_n(i, j) \rangle \\
 &= -\frac{P}{4} \langle N_n(i, j) \rangle [1 - \langle N_n(i, j + 1) \rangle] - \frac{P}{4} \langle N_n(i, j) \rangle [1 - \langle N_n(i, j - 1) \rangle] \\
 & \quad - \frac{P}{4} (1 + \rho) \langle N_n(i, j) \rangle [1 - \langle N_n(i + 1, j) \rangle] - \frac{P}{4} (1 - \rho) \langle N_n(i, j) \rangle [1 - \langle N_n(i - 1, j) \rangle] \\
 & \quad + \frac{P}{4} \langle N_n(i, j + 1) \rangle [1 - \langle N_n(i, j) \rangle] + \frac{P}{4} \langle N_n(i, j - 1) \rangle [1 - \langle N_n(i, j) \rangle] \\
 & \quad + \frac{P}{4} (1 + \rho) \langle N_n(i - 1, j) \rangle [1 - \langle N_n(i, j) \rangle] + \frac{P}{4} (1 - \rho) \langle N_n(i + 1, j) \rangle [1 - \langle N_n(i, j) \rangle].
 \end{aligned} \tag{5}$$

The positive terms on the right of Eq. (5) represent transitions into site  $(i, j)$  while the negative terms represent transitions out of site  $(i, j)$ . For example, the first term on the right hand side of Eq. (5) represents transitions out of site  $(i, j)$  into the site  $(i, j + 1)$ . This term consists of three factors:  $P/4$  is the probability that agents are motile in the upward vertical direction within the time period  $\tau$ ;  $\langle N_n(i, j) \rangle$  is the probability that site  $(i, j)$  is occupied;  $[1 - \langle N_n(i, j + 1) \rangle]$  is the probability that the site  $(i, j + 1)$  is unoccupied. The other terms on the right hand side of Eq. (5) have the same form and can be interpreted in a similar way.

Rearranging Eq. (5), dividing through by  $\tau$ , and letting  $\Delta$  and  $\tau$  tend to zero jointly while holding  $\Delta^2/\tau$  constant [25], the discrete differences can be written as derivatives of a continuous variable  $N(x, y, t)$  that satisfies

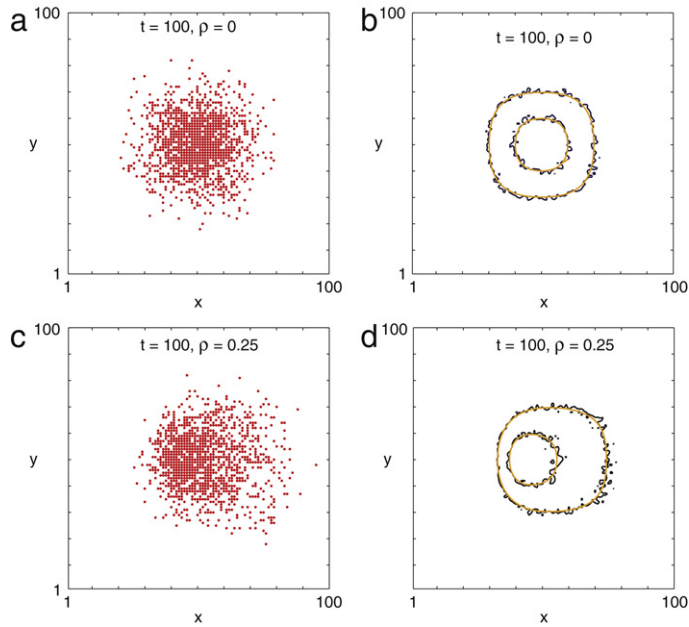
$$\frac{\partial N}{\partial t} = D \nabla^2 N - v \frac{\partial}{\partial x} [N(1 - N)], \tag{6}$$

where

$$D = \lim_{\Delta, \tau \rightarrow 0} \frac{P \Delta^2}{4\tau}, \quad v = \lim_{\Delta, \tau \rightarrow 0} \frac{\rho P \Delta}{2\tau}. \tag{7}$$

For biased motility,  $\rho = \mathcal{O}(\Delta)$  to obtain a well-defined continuum limit [25].

We note that the term associated with unbiased motility in the continuum model (Eq. (6)) is a linear diffusion term. This explains why our previous column-averaged simulation data matched a linear rather than a nonlinear diffusion equation (Fig. 1c–d). This term is linear since all nonlinear interactions associated with unbiased transitions in the discrete



**Fig. 2.** Agents spread from the square region  $35 \leq x, y \leq 65$  on a lattice with  $L_x = L_y = 100$ . Reflecting boundary conditions are imposed on all boundaries. (a) The distribution of unbiased agents at  $t = 100$ . (b) The 25% and 75% contours from the unbiased discrete model (black) together with the corresponding contours from the solution of Eq. (6) (orange). (c) The distribution of biased agents with  $\rho = 0.25$  at  $t = 100$ . (d) The 25% and 75% contours from the biased discrete model (black) together with the corresponding contours from the solution of Eq. (6) (orange). All discrete data are generated with  $\Delta = \tau = 1$  and averaged using Eq. (2) with  $M = 100$ . (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

conservation equation (Eq. (5)) cancel. The term associated with biased motility in the continuum model (Eq. (6)) is a nonlinear convection term that is associated with models of traffic flow [30] and the Burgers equation [23,24].

Discrete data is compared with solutions of Eq. (6) for a lattice with  $L_x = L_y = 100$  with all sites initially unoccupied, except for the square region  $35 \leq x, y \leq 65$  in which all sites are initially occupied (Fig. 2). Biased and unbiased simulations are performed and the distribution of agent occupancy is evaluated with Eq. (2). Contoured solutions of Eq. (6) are superimposed on contours of the discrete data, showing an excellent match (Fig. 2). All two-dimensional continuum results are obtained numerically with  $\delta x = \delta y = 0.5$ ,  $\delta t = 0.1$  and  $\epsilon = 1 \times 10^{-4}$ .

It is also convenient to average the discrete mechanism along each column in the lattice. If all terms in Eq. (5) are replaced with their corresponding column averages and we let  $\Delta \rightarrow 0$  and  $\tau \rightarrow 0$  while holding  $\Delta^2/\tau$  constant, the discrete differences can be written as derivatives of a continuous variable  $C(x, t)$  that satisfies

$$\frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial x^2} - v \frac{\partial}{\partial x} [C(1 - C)], \quad (8)$$

where  $D$  and  $v$  are previously defined by Eq. (7). If the initial distribution of agents in all columns on the lattice is uniform, then it is appropriate to compare column-averaged simulation data with a one-dimensional continuum model, as shown in Fig. 1. For more general initial conditions, it is appropriate to compare simulation results with a two-dimensional continuum model, as illustrated in Fig. 2.

The discrete density profiles for systems corresponding to Eq. (8) do not depend on the exact details of the initial distribution of agents, provided that the average occupancy within each column is identical for all Monte Carlo realisations. This is because the average occupancy of any site within a particular column is independent of the details of the locations of agents within that column.

Our approach to connect the discrete motility mechanism with a continuum model is related to other approaches in the mathematical biology literature. The novelty of our approach is that we start with a biologically realistic discrete exclusion process and then construct a conservation statement. This conservation statement can be interpreted in terms of transition probabilities and can be written as a continuum model in an appropriate limit. This approach will be extended to a system of interacting subpopulations in Section 4. Other researchers have taken alternative approaches, for example Othmer and Stevens [31] and Painter and coworkers [32,33] use a spatially discrete, continuous in time, nearest neighbour master equation. These previous approaches do not start from an exclusion process but instead from a probability distribution function. Transition probabilities that depend on volume filling lead to comparable continuum equations to Eq. (8) [32]. Other researchers have taken the opposite approach by assuming that a particular continuum description of the motility is valid, then a discretised version of the continuum model is interpreted as transition probabilities to define the discrete

motility mechanism [11,12]. We believe that our approach, starting with a biologically motivated exclusion process which effectively imposes transition probabilities and leading to continuum models, is well suited to interpreting time lapse data where individual-level “rules” or mechanisms can be qualitatively inferred through visual inspection of biological phenomena.

To gain further insight into the simple exclusion process, we now focus on a more general problem where the total population is composed of a mixture of two or more interacting subpopulations.

#### 4. Multi-species motility

We now consider an exclusion process with  $s$  subpopulations. Each subpopulation has an average site occupancy  $\langle N_n^{(r)}(i, j) \rangle$  for all  $r = 1, 2, 3, \dots, s$ . A discrete conservation of occupancy statement for each subpopulation is constructed. In the appropriate limit, these discrete conservation equations lead to  $s$  two-dimensional nonlinear advection–diffusion equations which account for the exclusion process. Here we limit our discussion to cases where the initial distribution of all subpopulations is uniform in all columns across the lattice. The resulting continuum equations can be written as a multi-species system with  $s$  subpopulations of density  $C^{(r)}$  for all  $r = 1, 2, 3, \dots, s$ . Subpopulation  $r$  moves with probability  $P^{(r)}$  per time step, and has a drift parameter  $\rho^{(r)}$ .

In conservation form, the governing continuum model is given by

$$\frac{\partial C^{(r)}}{\partial t} = -\frac{\partial J^{(r)}}{\partial x}, \quad r = 1, 2, 3, \dots, s, \quad (9)$$

where  $J^{(r)}$  is the flux of subpopulation  $r$ , that can be written as

$$J^{(r)} = -D^{(r)} \left( 1 - \sum_{q=1}^s C^{(q)} \right) \frac{\partial C^{(r)}}{\partial x} - D^{(r)} C^{(r)} \frac{\partial}{\partial x} \left( \sum_{q=1}^s C^{(q)} \right) + v^{(r)} C^{(r)} \left( 1 - \sum_{q=1}^s C^{(q)} \right), \quad (10)$$

where

$$D^{(r)} = \lim_{\Delta, \tau \rightarrow 0} \frac{P^{(r)} \Delta^2}{4\tau}, \quad v^{(r)} = \lim_{\Delta, \tau \rightarrow 0} \frac{\rho^{(r)} P^{(r)} \Delta}{2\tau}. \quad (11)$$

Again, we require  $\rho^{(r)} = \mathcal{O}(\Delta)$  for all  $r = 1, 2, 3, \dots, s$ . Each term in Eq. (10) represents a particular aspect of the net motility. The first term represents the diffusive flux of subpopulation  $r$ , and the factor  $(1 - \sum_{q=1}^s C^{(q)})$  represents a decrease in diffusivity due to exclusion by the total population. The second term is a first order advection term, representing the flux of subpopulation  $r$  down the gradient of the total agent density. The third term corresponds to the advective flux of subpopulation  $r$ , and the factor  $(1 - \sum_{q=1}^s C^{(q)})$  represents a decrease in the advective velocity due to exclusion by the total population. We now discuss some interesting points that arise from this formulation.

Suppose that all subpopulations are identical, that is, all subpopulations have the same diffusion coefficient  $D^{(r)}$  for all  $r = 1, 2, 3, \dots, s$ , and the same advection coefficient  $v^{(r)}$  for all  $r = 1, 2, 3, \dots, s$ . Under these conditions, each of the conservation equations can be summed to produce a single equation for the total population  $C = \sum_{q=1}^s C^{(q)}$ . Each of the nonlinear terms in the diffusive flux cancel, giving a linear diffusive flux for the total population  $C$ , while the advective term remains nonlinear. In summary, the equation for  $C$  reduces to Eq. (8). This cancellation of nonlinear diffusive terms is a generalisation of the individual tracking data that gave rise to a nonlinear diffusivity (Eq. (1)), yet these nonlinearities cancelled when we considered the evolution of the total population density in Fig. 1.

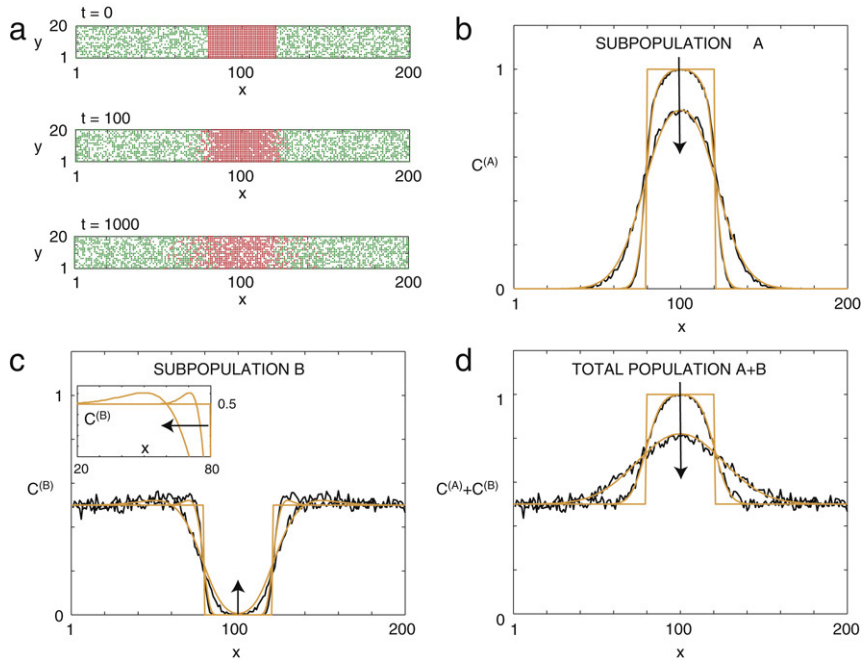
We now consider an alternative way of writing Eq. (10). With some cancellation it is possible to rewrite the flux as

$$J^{(r)} = -D^{(r)} \left( 1 - \sum_{\substack{q=1 \\ q \neq r}}^s C^{(q)} \right) \frac{\partial C^{(r)}}{\partial x} - D^{(r)} C^{(r)} \frac{\partial}{\partial x} \left( \sum_{\substack{q=1 \\ q \neq r}}^s C^{(q)} \right) + v^{(r)} C^{(r)} \left( 1 - \sum_{q=1}^s C^{(q)} \right). \quad (12)$$

Written in this way, the unbiased component of the flux of subpopulation  $r$  is not hindered by the presence of subpopulation  $r$ . The physical interpretation of this alternative expression for the flux is less satisfying than the physical interpretation of Eq. (10) since, in that equation, each subpopulation excludes all other subpopulations, regardless of whether it is from the same subpopulation or a different subpopulation.

We now compare discrete and continuum results for two specific examples of a two-species system which, for convenience, we call subpopulations  $A$  and  $B$ . First consider a system with identical subpopulations so that  $P^{(A)} = P^{(B)} = P$  and  $\rho^{(A)} = \rho^{(B)} = \rho$ . Eqs. (9) and (10) give a system of conservation equations for both subpopulations in terms of  $C^{(A)}(x, t)$  and  $C^{(B)}(x, t)$ . Summing these equations and setting  $C = C^{(A)} + C^{(B)}$  produces the same equation for the single population given by Eq. (8).

Simulations are performed for identical unbiased subpopulations with  $\rho^{(A)} = \rho^{(B)} = 0$  and  $P^{(A)} = P^{(B)} = 1$ . Each lattice site between  $80 \leq x \leq 120$  is initially fully occupied by subpopulation  $A$  (red agents) while the remaining lattice sites are



**Fig. 3.** Two unbiased identical subpopulations with  $\rho^{(A)} = \rho^{(B)} = 0$  and  $P^{(A)} = P^{(B)} = 1$  on a lattice with  $L_x = 200$  and  $L_y = 20$ . (a) Subpopulation A (red) and subpopulation B (green) mix over time. Density profiles (black) and continuum results (orange) are shown for (b) subpopulation A, (c) subpopulation B and (d) the total population. All results are given at  $t = 0, 100$  and  $1000$  with arrows indicating the direction of increasing time. All discrete data are generated with  $\Delta = \tau = 1$  and averaged using Eq. (3) with  $M = 40$ . (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

initially occupied by subpopulation B (green agents) at a density of 0.5. The averaged discrete results compare well with the corresponding continuum results (Fig. 3).

The solution profiles in Fig. 3 reveal some unusual and unexpected features. Density profiles of subpopulation A (Fig. 3b) are qualitatively similar to the single species unbiased problem (Fig. 1b). In contrast, the density profiles for subpopulation B (Fig. 3c) are strikingly different from those associated with a linear diffusion mechanism, since the  $C^{(B)}(x, 100)$  and  $C^{(B)}(x, 1000)$  profiles are nonmonotone, with maximum values greater than the initial condition (Fig. 3c). Clearly these solutions do not obey a maximum principle [34], which is intriguing given that the total population density  $C^{(A)} + C^{(B)}$  (Fig. 3d) is governed by linear diffusion and does obey a maximum principle. These simulations reveal differences that would not be obvious without considering both the discrete and continuous description of the simple exclusion process in a multi-species framework.

Next consider a biased system with distinct subpopulations,  $P^{(A)} = 0.5$  and  $P^{(B)} = 1$ , and  $\rho^{(A)} = \rho^{(B)} = 0.75$ . Each site between  $61 \leq x \leq 100$  is initially completely occupied by subpopulation A (red agents) while sites between  $20 \leq x \leq 60$  are initially occupied by subpopulation B (green agents) at a density of 0.5. Simulation data and associated continuous solutions are given in Fig. 4.

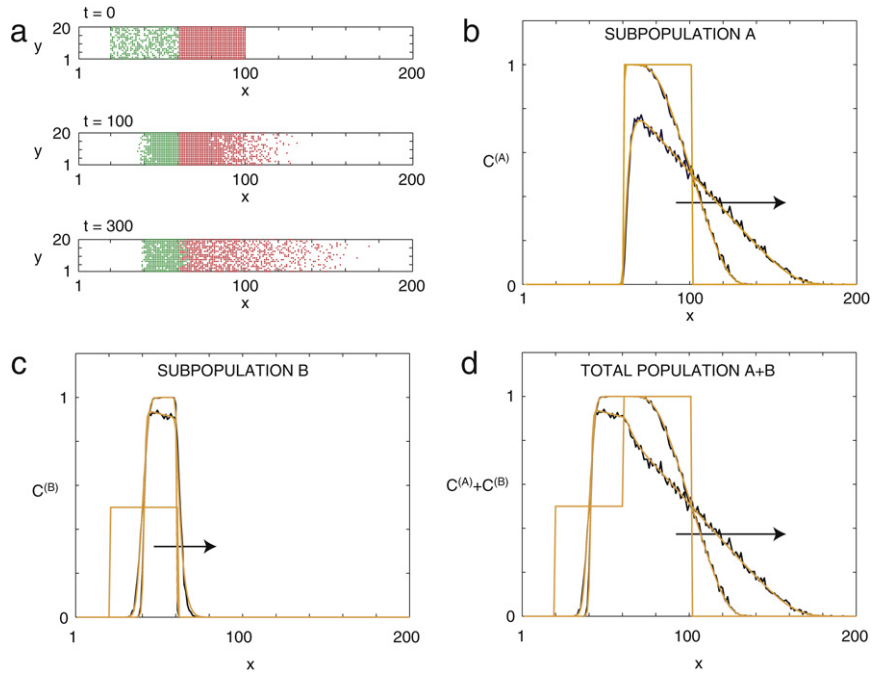
During the early part of the simulation, the agents in subpopulation A near  $x = 100$  move in the positive  $x$  direction (Fig. 4b), while the agents in subpopulation B move to the right within the region  $20 \leq x \leq 60$ . During the early part of the simulation, the movement of subpopulation B is restricted at  $x = 60$  due to exclusion by subpopulation A (Fig. 4a and c). At later times, after the density of subpopulation A at  $x = 60$  reduces sufficiently, subpopulation B is able to move further in the positive  $x$  direction past  $x = 60$ . This transient crowding and restriction of subpopulation B behind subpopulation A corresponds to a situation where a standard maximum principle does not apply to the equation governing subpopulation B.

## 5. Discussion and conclusions

This work investigates several aspects of motility based on a simple exclusion process. Diffusivity estimates from tracking data do not match the global behaviour of the system. It is important to emphasise this mismatch, since the generation and analysis of tracking data in cell biology is common [12]. Our work shows that transport coefficients extracted from this data may not describe the population-level behaviour. This mismatch is reconciled by averaging the discrete mechanism to give the appropriate population-level description.

Considering a population of agents as a series of interacting subpopulations is a useful way to probe deeper into the details of the simple exclusion process. In the simplest case, we represent a single population as two identical subpopulations. For unbiased motility, the total population obeys a linear diffusion mechanism, yet the subpopulations obey nonlinear





**Fig. 4.** Data for a biased multi-species simulation with distinct subpopulations on a lattice with  $L_x = 200$  and  $L_y = 20$ . Subpopulation A (red) is distinct from subpopulation B (green) with  $\rho^{(A)} = \rho^{(B)} = 0.75$ ,  $p^{(A)} = 0.5$  and  $p^{(B)} = 1$ . (a) Subpopulations interact over time. Density profiles (black) and continuum results (orange) are shown for (b) subpopulation A, (c) subpopulation B and (d) the total population. All results are given at  $t = 0, 100$  and  $1000$  with arrows indicating the direction of increasing time. All discrete data are generated with  $\Delta = \tau = 1$  and averaged using Eq. (3) with  $M = 40$ . (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

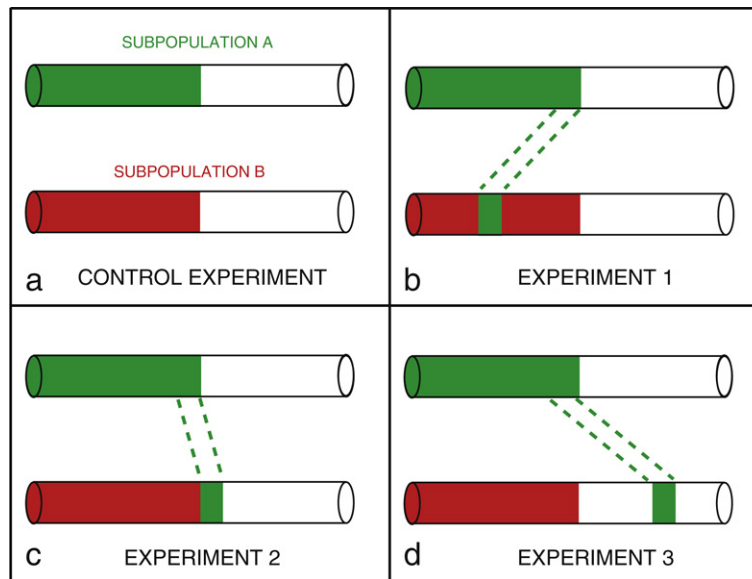
conservation equations. This is interesting, given that the subpopulations are identical, yet the law governing the total population is very different to the law governing each subpopulation. When the equations for the subpopulations are added, the nonlinear terms cancel, and we arrive back at a linear diffusion mechanism for the total population. This property of summing the equations for each subpopulation and arriving at a linear diffusion mechanism for the total population is observed in other systems, such as the work by Sherratt [35], who studied malignant invasion, and also for models in polymer chemistry [36] where certain individual-level interactions cancel out and do not appear at the population-level.

Our work relates to understanding cell motility within a population of cells. A standard problem in experimental cell biology is to track, through time, a subpopulation of cells within a larger population of cells. This provides an understanding of how cells interact with each other. Experimental techniques can be used to label an entire population of cells indistinguishably [16]. If the aim of an experiment is to understand how a subpopulation of these particular cells interact with the total population, then it is impossible to deduce this, if all cells are labelled identically. One way to overcome this difficulty is to perform a chimera experiment, for example, where a donor subpopulation labelled with a particular colour from one type of animal model is grafted into a host population labelled with a different colour from a different animal model. This approach has been used to help understand the movement of neural crest cells along the developing intestine, where donor quail cells were grafted into a chick host [17]. Alternatively, a new labelling technique has been developed where all cells express a photoconvertible fluorescent protein, KikGR, which allows the colour of a subpopulation to be selectively changed through photoconversion [15]. This technique allows a subpopulation of cells to be tracked without any grafting.

Either of these experimental approaches can be used to perform a range of cell labelling assays that are schematically outlined in Fig. 5. In these assays, two subpopulations, that are either identical or distinct, are tracked as they interact. In the control experiment, the behaviour of a single subpopulation is tracked over time. In experiment 1, a subpopulation of agents that is completely contiguous with the other subpopulation (two shared interfaces) is considered, whereas in experiment 2, a subpopulation that is partly contiguous with the other subpopulation (one shared interface) is tracked over time. The final experiment involves two noncontiguous subpopulations (no shared interface). These four cases have been investigated in previous experiments [17]. The multi-species models developed and investigated in this work are relevant to understanding and interpreting this kind of experimental data.

## Acknowledgments

This work is supported by the Australian Research Council (ARC). Mat Simpson is an ARC Postdoctoral Fellow and Kerry Landman is an ARC Professorial Fellow.



**Fig. 5.** Multi-species experiments from cell biology where a total population of cells is represented as a series of interacting subpopulations [17]. Several experiments are possible: (a) The control experiment involves tracking a subpopulation of cells with no interaction between subpopulations. (b) A labelled subpopulation that is completely contiguous with the other subpopulation (two shared interfaces). (c) A labelled subpopulation is partly contiguous with the other subpopulation (one shared interface). (d) A labelled subpopulation is noncontiguous with the total population (no shared interface).

## References

- [1] W.G. Weng, T. Chen, H.Y. Yuan, W.C. Fan, *Phys. Rev. E* 74 (2006) 036102.
- [2] D. Longo, S.M. Peirce, T.C. Skalak, L. Davidson, M. Marsden, B. Dzamba, D.W. DeSimone, *Dev. Biol.* 271 (2004) 210–222.
- [3] A. Schadschneider, *Physica A* 313 (2002) 153–187.
- [4] T. Callaghan, E. Khain, L.M. Sander, R.M. Ziff, *J. Stat. Phys.* 122 (5) (2006) 909–923.
- [5] D. Chowdhury, A. Schadschneider, K. Nishinari, *Phys. Life. Rev.* 2 (2005) 318–352.
- [6] R. Lipowsky, S. Klumpp, *Physica A* 352 (2005) 53–112.
- [7] M.J. Simpson, A. Merrifield, K.A. Landman, B.D. Hughes, *Phys. Rev. E* 76 (2007) 021918.
- [8] S. Turner, *Phys. Rev. E* 71 (2005) 041903.
- [9] A. Alarcón, H.M. Byrne, P.K. Maini, *J. Theor. Biol.* 225 (2003) 257–274.
- [10] M. Alber, N. Chen, T. Glimm, P.M. Lushnikov, *Phys. Rev. E* 73 (2006) 051901.
- [11] A.R.A. Anderson, M.A.J. Chaplain, *Bull. Math. Biol.* 60 (1998) 857–900.
- [12] A.Q. Cai, K.A. Landman, B.D. Hughes, *J. Theor. Biol.* 245 (2007) 576–594.
- [13] R. Erban, H.G. Othmer, *SIAM J. Appl. Math.* 65 (2004) 361–391.
- [14] A.F.M. Marée, P. Hogeweg, *Proc. Natl. Acad. Sci.* 98 (2001) 3879–3883.
- [15] P.M. Kulesa, J.M. Teddy, D.A. Stark, S.E. Smith, R. McLennan, *Dev. Biol.* 316 (2008) 275–287.
- [16] H.M. Young, A.J. Bergner, R.B. Anderson, H. Enomoto, J. Milbrandt, D.F. Newgreen, P.M. Whittington, *Dev. Biol.* 270 (2004) 455–473.
- [17] M.J. Simpson, D.C. Zhang, M. Mariani, K.A. Landman, D.F. Newgreen, *Dev. Biol.* 302 (2007) 553–568.
- [18] H.C. Berg, *Random Walks in Biology*, Expanded Edition, Princeton University Press, Princeton, USA, 1983.
- [19] A. Deutsch, S. Dormann, *Cellular Automaton Modeling of Biological Pattern Formation*, Birkhäuser, Boston, 2005.
- [20] S. Turner, J.A. Sherratt, K.J. Painter, N.J. Savill, *Phys. Rev. E* 69 (2004) 021910.
- [21] B.C. Thorne, A.M. Bailey, D.W. DeSimone, S.M. Peirce, *Birth Defects Res. C* 81 (2007) 344–353.
- [22] M.R. Evans, D.P. Foster, C. Godrèche, D. Mukamel, *Phys. Rev. Lett.* 74 (1995) 208–211.
- [23] G.M. Schütz, E. Domany, *J. Stat. Phys.* 72 (1993) 277–296.
- [24] G.M. Schütz, *J. Stat. Phys.* 88 (1997) 427–445.
- [25] B.D. Hughes, *Random Walks and Random Environments*, vol. 1, Oxford University Press, Oxford, UK, 1995.
- [26] F. Spitzer, *Adv. Math.* 5 (1970) 256–290.
- [27] T.M. Liggett, *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*, Springer-Verlag, Berlin, 1999.
- [28] A. Arratia, *Ann. Probab.* 11 (1983) 362–373.
- [29] E. Saada, *Ann. Probab.* 15 (1987) 375–381.
- [30] G.B. Whitham, *Linear and Nonlinear Waves*, Wiley, New York, 1974.
- [31] H.G. Othmer, A. Stevens, *SIAM J. Appl. Math.* 57 (4) (1997) 1044–1081.
- [32] K.J. Painter, T. Hillen, *Canad. Appl. Math. Q.* 10 (2002) 501–543.
- [33] K. Painter, J.A. Sherratt, *J. Theor. Biol.* 225 (2003) 325–337.
- [34] M.H. Protter, H.F. Weinberger, *Maximum Principles in Differential Equations*, Prentice Hall, Englewood Cliffs, New Jersey, 1967.
- [35] J.A. Sherratt, *Proc. R. Soc. Lond. A* 456 (2000) 2365–2386.
- [36] C.P. Hiemenz, *Polymer Chemistry: The Basic Concepts*, Marcel Dekker, New York, 1984.