Incorporating spatial correlations into multispecies mean-field models

Deborah C. Markham,[1,*] Matthew J. Simpson,[2] Philip K. Maini,[1] Eamonn A. Gaffney,[1] and Ruth E. Baker[1]

[1]*Wolfson Centre for Mathematical Biology, Mathematical Institute, University of Oxford, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, United Kingdom*

[2]*School of Mathematical Sciences, Queensland University of Technology, G.P.O. Box 2434, Brisbane, Queensland 4001, Australia*

In biology, we frequently observe different species existing within the same environment. For example, there are many cell types in a tumour, or different animal species may occupy a given habitat. In modeling interactions between such species, we often make use of the mean-field approximation, whereby spatial correlations between the locations of individuals are neglected. Whilst this approximation holds in certain situations, this is not always the case, and care must be taken to ensure the mean-field approximation is only used in appropriate settings. In circumstances where the mean-field approximation is unsuitable, we need to include information on the spatial distributions of individuals, which is not a simple task. In this paper, we provide a method that overcomes many of the failures of the mean-field approximation for an on-lattice volume-excluding birth-death-movement process with multiple species. We explicitly take into account spatial information on the distribution of individuals by including partial differential equation descriptions of lattice site occupancy correlations. We demonstrate how to derive these equations for the multispecies case and show results specific to a two-species problem. We compare averaged discrete results to both the mean-field approximation and our improved method, which incorporates spatial correlations. We note that the mean-field approximation fails dramatically in some cases, predicting very different behavior from that seen upon averaging multiple realizations of the discrete system. In contrast, our improved method provides excellent agreement with the averaged discrete behavior in all cases, thus providing a more reliable modeling framework. Furthermore, our method is tractable as the resulting partial differential equations can be solved efficiently using standard numerical techniques.

PACS number(s): 87.10.Ed, 87.10.Hk, 87.10.Mn

I. INTRODUCTION

Biological systems frequently involve multiple species occupying the same environment. For example, it is well known that cells within a tumor tend not to be phenotypically identical; this heterogeneity has led to questioning the efficacy of many cancer treatments currently in use [1–3]. It has also been observed that the presence of one species can have a detrimental effect on the survival of another, as in the case of red and gray squirrels in Britain [4,5]. In order to understand and accurately predict the behavior of multiple species in an environment, it is essential that we develop robust mathematical frameworks to describe them.

Much work has been devoted to the development of phenomenological descriptions of multispecies systems in biology, leading to many important advances in a number of areas [6,7]. However, these models are frequently based on the mean-field approximation (MFA), which assumes that there are no spatial correlations between the location of individuals; this assumption is not always valid and thus may lead to inaccurate predictions if used inappropriately [8]. This has been demonstrated, for a single species case, where the MFA and a method that incorporated spatial information gave different results when used to calibrate data from experiments with a breast cancer cell line [9]. Additionally, in problems concerning invasive species, ignoring stochastic effects, which result in spatial clusters of individuals, may lead to an overestimation of the invasion speed [10].

Spatial effects can be explicitly included by using agent-based models, whereby each cell type follows a specific set of rules [11–13]. However, this approach is computationally expensive, time-consuming, and generally not amenable to mathematical analysis, especially as the number of species increases. Using pair-wise densities is a simpler option for incorporating spatial information and is frequently used for investigating biological systems [14–26]. In Ref. [14], the authors correct the MFA using pair-wise dynamics on a volume-excluding system with one cell type moving, proliferating, and dying on a regular lattice. This method successfully reproduces the averaged discrete behavior; however, it is not straightforward to implement in two and three dimensions due to the complicated system of ordinary differential equations (ODEs) obtained. A method to circumvent this problem was developed in Ref. [27], where the authors use a partial differential equation (PDE) representation of the spatial correlations between lattice sites. The development of a method such as this for systems of multiple species would be a valuable step forward in our quest for simplified spatial models.

In this paper, we show how to derive PDE descriptions of site occupancy correlation functions for an N-species volume-excluding system on a regular, square lattice. The species do not have a direct effect on one another but are indirectly competing by occupying the same space. We show the failure of the MFA in predicting the averaged individual-level behavior in certain situations. This failure could have severe consequences if the MFA is used to model a situation for which it is not an appropriate choice. In contrast, we demonstrate that our correlations PDE model leads to accurate

---

*markham@maths.ox.ac.uk

predictions of the averaged discrete behavior over a wide range of parameter space.

We begin, in Sec. II, by describing the system we are modeling and the details of the mathematical methods used in deriving the relevant equations. In Sec. III, we discuss some results for the two-species case by comparing our numerical results to those from averaged discrete simulations, for systems both with and without agent death. We discuss scenarios whereby the mean-field model has great shortcomings and show that they are eliminated by incorporation of lattice site occupancy correlations. Finally, in Sec. IV, we conclude and discuss some future directions.

II. THE SYSTEM

We consider a multispecies contact process with diffusion [28,29], whereby we have a volume-excluding system with N different species of agents undergoing movement, proliferation, and death. The Ith species has a movement rate per unit time of P_m^I, a proliferation rate per unit time of P_p^I, and a death rate per unit time of P_d^I. We have a d-dimensional system on a regular, square lattice with a lattice spacing of Δ. The lattice is initially randomly populated, thus lattice site occupancies are initially uncorrelated and the density is approximately spatially uniform. When an agent moves or divides, it chooses the new location at random from its von Neumann neighborhood, and the event is aborted if the chosen site is already occupied.

As with the one species case, we make use of the k-point distribution functions, $\rho^{(k)}$ [14], which can be thought of as the probability of k-tuplets of sites having specific occupancies. When proliferation is present in the system, evolution of $\rho^{(k)}$ depends on the evolution of $\rho^{(k+1)}$, thus the system is described by an impractically large system of ODEs [14]. In order to overcome this, we must choose a point at which to close our system. We close at $\rho^{(3)}$ using the Kirkwood superposition approximation (KSA) [30,31], which expresses $\rho^{(3)}$ in terms of lower-order distribution functions:

$$\rho^{(3)}(\sigma_l, \sigma_m, \sigma_n) = \frac{\rho^{(2)}(\sigma_l, \sigma_m) \rho^{(2)}(\sigma_l, \sigma_n) \rho^{(2)}(\sigma_m, \sigma_n)}{\rho^{(1)}(\sigma_l) \rho^{(1)}(\sigma_m) \rho^{(1)}(\sigma_n)}, \quad (2.1)$$

where σ_l denotes the state of lattice site l. We use correlation functions to determine the level of spatial correlation between lattice site occupancies a given distance apart. We define our correlation functions in keeping with previous work [14,32,33] as

$$F_{\sigma_l, \sigma_m}(|\mathbf{l} - \mathbf{m}|) =: \frac{\rho^{(2)}(\sigma_l, \sigma_m)}{\rho^{(1)}(\sigma_l) \rho^{(1)}(\sigma_m)}, \quad (2.2)$$

where $\mathbf{l}$ and $\mathbf{m}$ are vectors denoting the positions of sites l and m, respectively, on the lattice. The correlation functions depend only on the distance between the lattice sites, as we have assumed an inheritance of the translational invariance and isotropy associated with the initial conditions. This assumption will be made throughout this paper. Note that, using this description, if lattice site occupancies are independent, the correlation functions are equal to unity.

In the one species model, we have an ODE describing the evolution of the agent density, coupled to a system of ODEs for

the correlation functions [14]. This system of ODEs can instead be written as a PDE with a boundary ODE for correlations at a distance of Δ and a far-field boundary condition [27]. With multiple species, we will obviously have a larger system of equations. We now discuss how to determine the number of equations that need solving for a system with N species.

We begin with the equations for evolution of the agent densities. Due to the homogeneous initial conditions, the one-point distribution function for the Ith species is given by the normalized concentration of the Ith species:

$$\rho^{(1)}(I_l) = c_I, \quad (2.3)$$

where c_I is given by

$$c_I(t) = \frac{1}{\eta_{\text{tot}}} \sum_i \mathbb{1}_I\{\sigma_i\}, \quad (2.4)$$

with η_{tot} being the total number of lattice sites. The sites must be occupied by one of the species, or vacant, thus leading to the following conservation statement:

$$\sum_{J=0}^{N} c_J = 1, \quad (2.5)$$

where c_0 is the normalized density of empty sites. Thus, we can reduce the number of unknown densities from $N + 1$ to N using the conservation statement to write c_0 as a function of the other densities:

$$c_0 = 1 - \sum_{J=1}^{N} c_J. \quad (2.6)$$

We use the evolution of the two-point functions to derive the evolution of our correlation functions, according to the definition in Eq. (2.2). The conservation statements for the two-point functions take the form

$$\sum_{J=0}^{N} \rho^{(2)}(I_l, J_m) = \rho^{(1)}(I_l) = c_I. \quad (2.7)$$

In total, we have $N + 1$ second-order conservation statements. We use these to reduce the number of unknown correlation functions, choosing to eliminate all F_{I0} where $I = 0, 1, 2, \dots, N$ using expressions of the following form:

$$F_{I0}(|\mathbf{l} - \mathbf{m}|) = \frac{1 - \sum_{J=1}^{N} c_J F_{IJ}(|\mathbf{l} - \mathbf{m}|)}{1 - \sum_{J=1}^{N} c_J}. \quad (2.8)$$

Therefore, in total, we have N unknown densities and $\sum_{J=1}^{N} J = N(N + 1)/2$ unknown types of correlation functions. Using PDE descriptions of the correlation functions, as in Ref. [27], leads to $N(N + 1)/2$ PDEs, each with a boundary ODE for correlations at a distance of Δ and a far-field boundary condition.

A. Deriving the density and correlation ODEs

First, we use the one-point distribution functions to obtain equations for the evolution of the normalized density of each

species:

$$\frac{d\rho^{(1)}(I_l)}{dt} = \underbrace{P_m^I \sum_n \frac{\alpha_{n,l}}{z} [\rho^{(2)}(0_l, I_n) - \rho^{(2)}(I_l, 0_n)]}_{\text{movement in and out of } l}$$

$$+ \underbrace{P_p^I \sum_n \frac{\alpha_{n,l}}{z} \rho^{(2)}(0_l, I_n)}_{\text{proliferation into } l} - \underbrace{P_d^I \rho^{(1)}(I_l)}_{\text{death at } l}, \quad (2.9)$$

where $z = 2d$ (d being the number of spatial dimensions), and $\alpha_{n,l}$ is unity if n and l are nearest neighbors, otherwise it is zero. This equation simplifies to

$$\frac{dc_I}{dt} = P_p^I c_I \left[1 - \sum_{J=1}^N c_J F_{IJ}(\Delta) \right] - P_d^I c_I. \quad (2.10)$$

If lattice sites are independent (all $F_{IJ}(|\mathbf{l} - \mathbf{m}|) = 1$), these equations reduce to the MFA, a system of logistic equations coupled by the carrying capacity term:

$$\frac{dc_I}{dt} = P_p^I c_I \left[1 - \sum_{J=1}^N c_J \right] - P_d^I c_I. \quad (2.11)$$

We now turn to the evolution of the correlation functions. These will take two different forms depending on whether we are considering the spatial correlations between two agents of the same or different species separated by a radial distance of $|\mathbf{l} - \mathbf{m}|$. This differs from the one-species case where we only have one type of correlation function because we only have one agent type. We first consider the evolution of the auto-correlation functions, which apply to agents of the same species at a distance of $|\mathbf{l} - \mathbf{m}|$, by deriving the two-point distribution functions for the same species case:

$$\frac{d\rho^{(2)}(I_l, I_m)}{dt} = P_m^I \left[\sum_{n \neq l} \frac{\alpha_{n,m}}{z} \rho^{(3)}(I_l, 0_m, I_n) - \sum_{n \neq l} \frac{\alpha_{n,m}}{z} \rho^{(3)}(I_l, I_m, 0_n) + \sum_{n \neq m} \frac{\alpha_{n,l}}{z} \rho^{(3)}(0_l, I_m, I_n) - \sum_{n \neq m} \frac{\alpha_{n,l}}{z} \rho^{(3)}(I_l, I_m, 0_n) \right]$$

$$+ P_p^I \left[\sum_{n \neq l} \frac{\alpha_{n,m}}{z} \rho^{(3)}(I_l, 0_m, I_n) + \sum_{n \neq m} \frac{\alpha_{n,l}}{z} \rho^{(3)}(0_l, I_m, I_n) \right] + P_p^I \frac{\alpha_{l,m}}{z} [\rho^{(2)}(I_l, 0_m) + \rho^{(2)}(0_l, I_m)] - 2 P_d^I \rho^{(2)}(I_l, I_m).$$

$$(2.12)$$

We can simplify to

$$\frac{d\rho^{(2)}(I_l, I_m)}{dt} = 2 P_m^I \sum_{n \neq l} \frac{\alpha_{n,m}}{z} [\rho^{(3)}(I_l, 0_m, I_n) - \rho^{(3)}(I_l, I_m, 0_n)] + 2 P_p^I \sum_{n \neq l} \frac{\alpha_{n,m}}{z} \rho^{(3)}(I_l, 0_m, I_n)$$

$$+ 2 P_p^I \frac{\alpha_{l,m}}{z} \rho^{(2)}(I_l, 0_m) - 2 P_d^I \rho^{(2)}(I_l, I_m). \quad (2.13)$$

Recall, from Eq. (2.2), that we can also write

$$\frac{d\rho^{(2)}(I_l, I_m)}{dt} = c_I^2 \frac{dF_{II}(|\mathbf{l} - \mathbf{m}|)}{dt} + 2 c_I \frac{dc_I}{dt} F_{II}(|\mathbf{l} - \mathbf{m}|). \quad (2.14)$$

Therefore, the evolution of our auto-correlation functions is given by

$$\frac{dF_{II}(|\mathbf{l} - \mathbf{m}|)}{dt} = \frac{2 P_m^I}{c_I^2} \sum_{n \neq l} \frac{\alpha_{n,m}}{z} [\rho^{(3)}(I_l, 0_m, I_n) - \rho^{(3)}(I_l, I_m, 0_n)] + \frac{2 P_p^I}{c_I^2} \sum_{n \neq l} \frac{\alpha_{n,m}}{z} \rho^{(3)}(I_l, 0_m, I_n)$$

$$+ \frac{2 P_p^I}{c_I} \frac{\alpha_{l,m}}{z} \lambda_I(|\mathbf{l} - \mathbf{m}|) - 2 P_p^I F_{II}(|\mathbf{l} - \mathbf{m}|) \lambda_I(\Delta), \quad (2.15)$$

where

$$\lambda_I(s) = 1 - \sum_{J=1}^N c_J F_{IJ}(s), \quad (2.16)$$

with $s = |\mathbf{l} - \mathbf{m}|$ being the radial distance between sites l and m. At this point we close our system using the KSA, as defined in Eq. (2.1). A key difference arises at this point between the one-species and N-species problems. In the case with one species, the three-point expressions in the movement terms were eliminated using conservation statements of the following form:

$$\rho^{(2)}(I_l, K_m) = \sum_{J=0}^N \rho^{(3)}(I_l, K_m, J_n). \quad (2.17)$$

However, in the N-species case, using these conservation equations will no longer eliminate $\rho^{(3)}$ from the movement term. This is due to the additional terms in the conservation equations that arise from having more than one species. Additionally, when the $\rho^{(3)}$ terms in the conservation equations are expressed using the KSA, the conservation equations no longer hold exactly, as we are using an approximation for each $\rho^{(3)}$ term, leading to additional discrepancies in our estimation of the averaged discrete behavior. Thus, we do not rewrite our movement terms using the conservation equations and instead use the KSA directly on the $\rho^{(3)}$ terms in Eq. (2.15). This leads to the movement terms of our equations having a different form to the one species case. Closing with the KSA leads to

$$
\frac{dF_{II}(|\mathbf{l}-\mathbf{m}|)}{dt} = \frac{2P_m^I}{1-\sum_{J=1}^N c_J} \sum_{n\neq l} \frac{\alpha_{n,m}}{z} \lambda_I(\Delta) \Bigg[F_{II}(|\mathbf{l}-\mathbf{n}|) - F_{II}(|\mathbf{l}-\mathbf{m}|)
$$
$$
- F_{II}(|\mathbf{l}-\mathbf{n}|) \sum_{J=1}^N c_J F_{IJ}(|\mathbf{l}-\mathbf{m}|) + F_{II}(|\mathbf{l}-\mathbf{m}|) \sum_{J=1}^N c_J F_{IJ}(|\mathbf{l}-\mathbf{n}|) \Bigg]
$$
$$
+ \frac{2P_p^I}{1-\sum_{J=1}^N c_J} \sum_{n\neq l} \frac{\alpha_{n,m}}{z} \lambda_I(|\mathbf{l}-\mathbf{m}|)\lambda_I(\Delta) F_{II}(|\mathbf{l}-\mathbf{n}|) + \frac{2P_p^I}{c_I} \frac{\alpha_{l,m}}{z} \lambda_I(\Delta) - 2P_p^I \lambda_I(\Delta) F_{II}(|\mathbf{l}-\mathbf{m}|). \quad (2.18)
$$

We now consider the evolution of the cross-correlation equations, which are applicable to agents of different species separated by a distance of $|\mathbf{l}-\mathbf{m}|$. We again start with evolution of the two-point distribution function, which we can again simplify due to translational invariance, to arrive at

$$
\frac{d\rho^{(2)}(I_l,K_m)}{dt} = P_m^I \sum_{n\neq l} \frac{\alpha_{n,m}}{z} [\rho^{(3)}(K_l,0_m,I_n) - \rho^{(3)}(K_l,I_m,0_n)] + P_m^K \sum_{n\neq l} \frac{\alpha_{n,m}}{z} [\rho^{(3)}(I_l,0_m,K_n) - \rho^{(3)}(I_l,K_m,0_n)]
$$
$$
+ P_p^I \sum_{n\neq l} \frac{\alpha_{n,m}}{z} \rho^{(3)}(K_l,0_m,I_n) + P_p^K \sum_{n\neq l} \frac{\alpha_{n,m}}{z} \rho^{(3)}(I_l,0_m,K_n) - P_d^I \rho^{(2)}(I_l,K_m) - P_d^K \rho^{(2)}(I_l,K_m). \quad (2.19)
$$

From the definition of the correlation function, we obtain

$$
\frac{d\rho^{(2)}(I_l,K_m)}{dt} = c_I c_K \frac{dF_{IK}(|\mathbf{l}-\mathbf{m}|)}{dt} + \left(c_I \frac{dc_K}{dt} + c_K \frac{dc_I}{dt} \right) F_{IK}(|\mathbf{l}-\mathbf{m}|). \quad (2.20)
$$

We relate Eqs. (2.19) and (2.20) in the same way as before to obtain

$$
\frac{dF_{IK}(|\mathbf{l}-\mathbf{m}|)}{dt} = \frac{P_m^I}{c_I c_K} \sum_{n\neq l} \frac{\alpha_{n,m}}{z} [\rho^{(3)}(K_l,0_m,I_n) - \rho^{(3)}(K_l,I_m,0_n)] + \frac{P_m^K}{c_I c_K} \sum_{n\neq l} \frac{\alpha_{n,m}}{z} [\rho^{(3)}(I_l,0_m,K_n) - \rho^{(3)}(I_l,K_m,0_n)]
$$
$$
+ \frac{P_p^I}{c_I c_K} \sum_{n\neq l} \frac{\alpha_{n,m}}{z} \rho^{(3)}(K_l,0_m,I_n) + \frac{P_p^K}{c_I c_K} \sum_{n\neq l} \frac{\alpha_{n,m}}{z} \rho^{(3)}(I_l,0_m,K_n)
$$
$$
- P_p^I F_{IK}(|\mathbf{l}-\mathbf{m}|)\lambda_I(\Delta) - P_p^K F_{IK}(|\mathbf{l}-\mathbf{m}|)\lambda_K(\Delta). \quad (2.21)
$$

Again, we use the KSA to close our equations, without rewriting our movement terms using the conservation equations, obtaining

$$
\frac{dF_{IK}(|\mathbf{l}-\mathbf{m}|)}{dt} = \frac{P_m^I}{1-\sum_{J=1}^N c_J} \sum_{n\neq l} \frac{\alpha_{n,m}}{z} \lambda_I(\Delta) \Bigg[F_{IK}(|\mathbf{l}-\mathbf{n}|) - F_{IK}(|\mathbf{l}-\mathbf{m}|)
$$
$$
- F_{IK}(|\mathbf{l}-\mathbf{n}|) \sum_{J=1}^N c_J F_{KJ}(|\mathbf{l}-\mathbf{m}|) + F_{IK}(|\mathbf{l}-\mathbf{m}|) \sum_{J=1}^N c_J F_{KJ}(|\mathbf{l}-\mathbf{n}|) \Bigg]
$$
$$
+ \frac{P_m^K}{1-\sum_{J=1}^N c_J} \sum_{n\neq l} \frac{\alpha_{n,m}}{z} \lambda_K(\Delta) \Bigg[F_{IK}(|\mathbf{l}-\mathbf{n}|) - F_{IK}(|\mathbf{l}-\mathbf{m}|)
$$
$$
- F_{IK}(|\mathbf{l}-\mathbf{n}|) \sum_{J=1}^N c_J F_{IJ}(|\mathbf{l}-\mathbf{m}|) + F_{IK}(|\mathbf{l}-\mathbf{m}|) \sum_{J=1}^N c_J F_{IJ}(|\mathbf{l}-\mathbf{n}|) \Bigg]
$$
$$
+ \frac{P_p^I}{1-\sum_{J=1}^N c_J} \sum_{n\neq l} \frac{\alpha_{n,m}}{z} \lambda_K(|\mathbf{l}-\mathbf{m}|)\lambda_I(\Delta) F_{IK}(|\mathbf{l}-\mathbf{n}|)
$$
$$
+ \frac{P_p^K}{1-\sum_{J=1}^N c_J} \sum_{n\neq l} \frac{\alpha_{n,m}}{z} \lambda_I(|\mathbf{l}-\mathbf{m}|)\lambda_K(\Delta) F_{IK}(|\mathbf{l}-\mathbf{n}|) - P_p^I \lambda_I(\Delta) F_{IK}(|\mathbf{l}-\mathbf{m}|) - P_p^K \lambda_K(\Delta) F_{IK}(|\mathbf{l}-\mathbf{m}|).
$$
$$
(2.22)
$$

B. Deriving the correlation PDEs

As in previous work [27], we now derive PDE descriptions of the correlation functions. We demonstrate the method in two dimensions (2D), although the same principles apply in three dimensions (3D), as seen in Ref. [27]. We begin with the PDE for the auto-correlation functions. The PDE will be defined for the spatial domain $\Delta < s < \infty$. We expect lattice sites that are far apart to be independent, thus leading to the boundary condition

$$F_{II}(s \to \infty) = 1. \tag{2.23}$$

To obtain the boundary condition at $s = \Delta$ we use Eq. (2.18), which gives

$$\frac{dF_{II}(\Delta)}{dt} = \frac{P_m^I}{2\left(1 - \sum_{J=1}^{N} c_J\right)} \lambda_I(\Delta) \left[-3F_{II}(\Delta) + \{2F_{II}(\sqrt{2}\Delta) + F_{II}(2\Delta)\}\lambda_I(\Delta) + F_{II}(\Delta) \sum_{J=1}^{N} c_J\{2F_{IJ}(\sqrt{2}\Delta) + F_{IJ}(2\Delta)\} \right]$$

$$+ \frac{P_p^I}{2\left(1 - \sum_{J=1}^{N} c_J\right)} \lambda_I(\Delta)^2 [2F_{II}(\sqrt{2}\Delta) + F_{II}(2\Delta)] + \frac{P_p^I}{2c_I}\lambda_I(\Delta) - 2P_p^I \lambda_I(\Delta)F_{II}(\Delta). \tag{2.24}$$

To obtain the PDE, we rewrite Eq. (2.18) using Cartesian coordinates and then perform a Taylor expansion of the correlation function around (x,y). Finally, we take $\Delta \to 0$, explicitly retaining our boundary terms and return to radial coordinates, obtaining the following 2D PDE:

$$\frac{\partial F_{II}(s)}{\partial t} = \frac{P_m^I}{2\left(1 - \sum_{J=1}^{N} c_J\right)} \lambda_I(\Delta) \left[\nabla^2 F_{II}\lambda_I(s) + F_{II}(s) \sum_{J=1}^{N} c_J \nabla^2 F_{IJ}(s) \right]$$

$$+ \frac{2P_p^I}{1 - \sum_{J=1}^{N} c_J} \lambda_I(\Delta)\lambda_I(s)F_{II}(s) - 2P_p^I \lambda_I(\Delta)F_{II}(s), \qquad \Delta < s < \infty, \tag{2.25}$$

where ∇^2 is the polar laplacian with no angular dependence.

We do the same for the cross-correlation functions to obtain the following PDE:

$$\frac{\partial F_{IK}(s)}{\partial t} = \frac{P_m^I}{4\left(1 - \sum_{J=1}^{N} c_J\right)} \lambda_I(\Delta) \left[\nabla^2 F_{IK}(s)\lambda_K(s) + F_{IK}(s) \sum_{J=1}^{N} c_J \nabla^2 F_{KJ}(s) \right]$$

$$+ \frac{P_m^K}{4\left(1 - \sum_{J=1}^{N} c_J\right)} \lambda_K(\Delta) \left[\nabla^2 F_{IK}(s)\lambda_I(s) + F_{IK}(s) \sum_{J=1}^{N} c_J \nabla^2 F_{IJ}(s) \right]$$

$$+ \frac{P_p^I}{1 - \sum_{J=1}^{N} c_J} \lambda_I(\Delta)\lambda_K(s)F_{IK}(s) + \frac{P_p^K}{1 - \sum_{J=1}^{N} c_J} \lambda_K(\Delta)\lambda_I(s)F_{IK}(s)$$

$$- P_p^I \lambda_I(\Delta)F_{IK}(s) - P_p^K \lambda_K(\Delta)F_{IK}(s), \tag{2.26}$$

with boundary conditions

$$\frac{dF_{IK}(\Delta)}{dt} = \frac{P_m^I}{4\left(1 - \sum_{J=1}^{N} c_J\right)} \lambda_I(\Delta) \Bigg\{ -3F_{IK}(\Delta) + [2F_{IK}(\sqrt{2}\Delta) + F_{IK}(2\Delta)]\lambda_K(\Delta)$$

$$+ F_{IK}(\Delta) \sum_{J=1}^{N} c_J [2F_{KJ}(\sqrt{2}\Delta) + F_{KJ}(2\Delta)] \Bigg\}$$

$$+ \frac{P_m^K}{4\left(1 - \sum_{J=1}^{N} c_J\right)} \lambda_K(\Delta) \Bigg\{ -3F_{IK}(\Delta) + [2F_{IK}(\sqrt{2}\Delta) + F_{IK}(2\Delta)]\lambda_I(\Delta)$$

$$+ F_{IK}(\Delta) \sum_{J=1}^{N} c_J [2F_{IJ}(\sqrt{2}\Delta) + F_{IJ}(2\Delta)] \Bigg\}$$

$$+ \frac{P_p^I + P_p^K}{4\left(1 - \sum_{J=1}^{N} c_J\right)} \lambda_I(\Delta)\lambda_K(\Delta)[2F_{IK}(\sqrt{2}\Delta) + F_{IK}(2\Delta)] - P_p^I \lambda_I(\Delta)F_{IK}(\Delta) - P_p^K \lambda_K(\Delta)F_{IK}(\Delta), \tag{2.27}$$

and

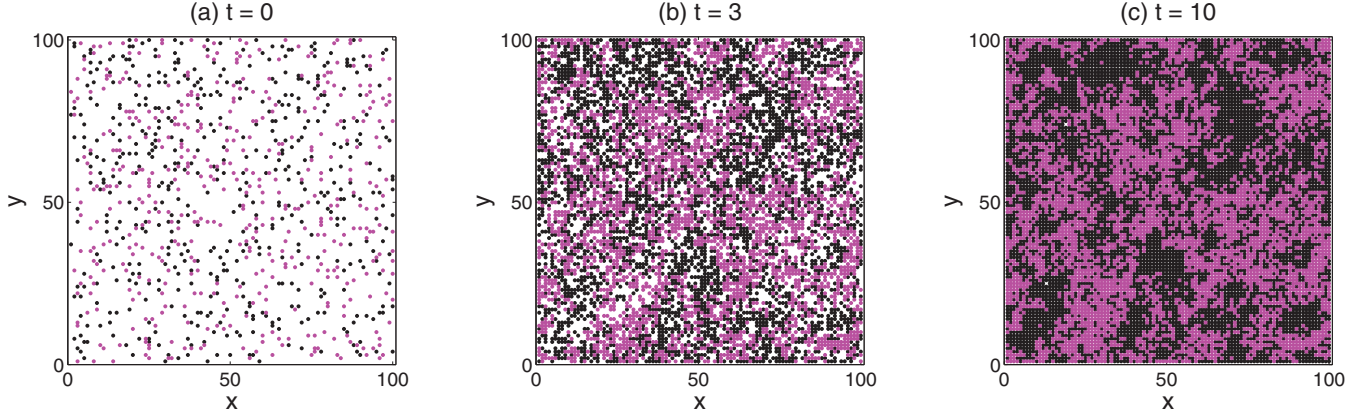$$F_{IK}(s \to \infty) = 1. \tag{2.28}$$

FIG. 1. (Color online) An individual discrete simulation showing the build-up of spatial correlations. Details of the numerical algorithm can be found in Sec. III. In this case, $P_m^A = 20$, $P_p^A = 1$, $P_m^B = 5$, $P_p^B = 1$. Species A is in black, while species B is magenta/gray.

III. RESULTS FROM THE TWO-SPECIES CASE

We demonstrate the utility of our method for the $N = 2$ case, whereby we have two species (A and B). This leads to a coupled system of two ODEs (for the evolution of the cell density of each species) and three PDEs (two for the evolution of the auto-correlation functions and one for that of the cross-correlation function), with each PDE having an ODE at the $s = \Delta$ boundary as well as a far-field boundary condition. We discretize our equations using the backwards Euler method and solve the resulting nonlinear equations using the tridiagonal matrix algorithm [34] with Picard iteration [35]. We compare these results to those obtained using the MFA [Eq. (2.11)] and averaged results of discrete simulations. We solve the system of ODEs for the MFA using a fourth-order Runge-Kutta method [34] with a timestep, δt, of 0.1 (smaller timesteps were tested to confirm this choice). To obtain the averaged discrete results, we use a modified Gillespie algorithm [14] and average over 100 realizations, for a 100×100 lattice. For

all results presented here, we use identical initial conditions, $c_A(0) = c_B(0) = 0.05$.

A. Systems undergoing proliferation and movement

We first investigate systems undergoing proliferation and movement but not death ($P_d = 0$). Results from a single realization of the discrete system can be seen in Fig. 1. We discuss here three cases where the MFA makes very different predictions from what is observed in the averaged discrete behavior. In all cases, our correlations PDE method provides excellent predictions of the averaged discrete behavior, as seen in Fig. 2. First, we investigate the case where the two species have the same proliferation rate but different movement rates [Fig. 2(a)]. The MFA predicts that the two species will behave in the exact same way, leading to each population occupying half of the lattice sites on average. However, the averaged individual realizations show that the species with the higher movement rate will generally reach a higher total agent density
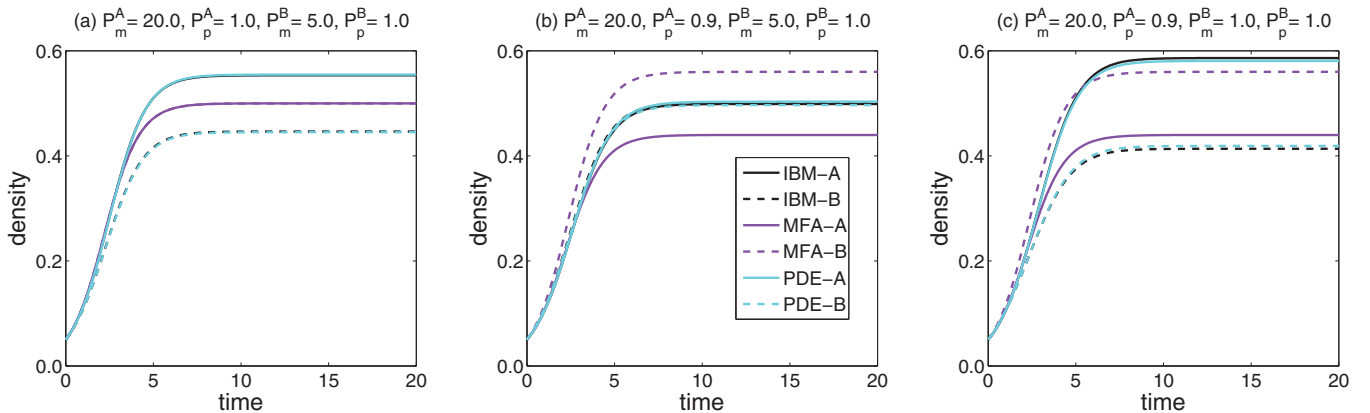

FIG. 2. (Color online) Our correlations PDE method [using Eqs. (2.23) to (2.28)] accurately predicts the averaged discrete behavior in situations where the MFA does not provide good agreement. The solid lines correspond to species A, while the dashed lines correspond to species B. The averaged discrete results are in black, the MFA in purple/dark gray and the correlations PDE in cyan/light gray. In all of these cases, the death rates are zero ($P_d^A = P_d^B = 0$). In (a), the two species have the same proliferation rates but different movement rates. We observe that the MFA predicts that both species will behave in exactly the same manner. However, the averaged discrete behavior shows that the species with a higher movement rate has a larger steady-state density. In (b), the MFA predicts different behavior for the two different species while the averaged discrete simulations predict almost identical behavior. In (c), the MFA predicts that species B will dominate at steady-state, while the averaged discrete simulations predict that species A will dominate.
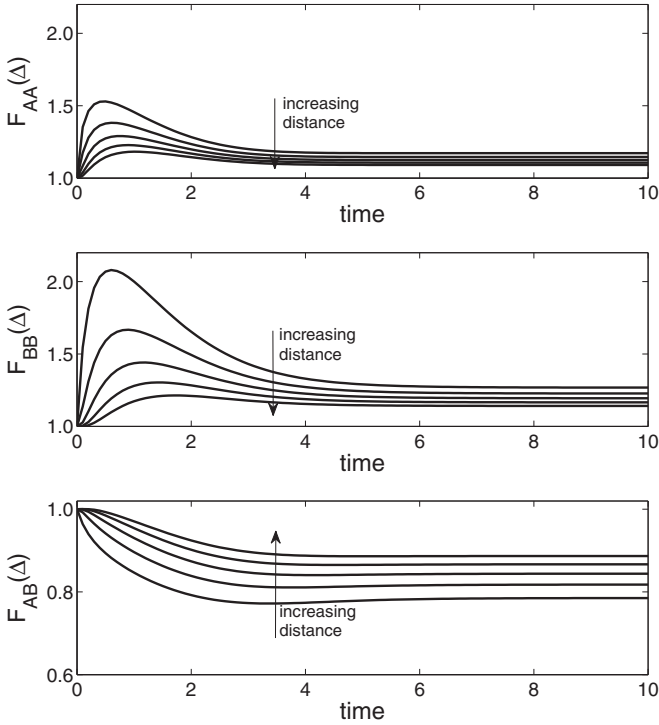
FIG. 3. The correlation functions F_{AA}, F_{BB}, and F_{AB} [from Eqs. (2.23) to (2.28)], for parameters $P_m^A = 20$, $P_p^A = 1$, $P_m^B = 5$, $P_p^B = 1$ have different properties. The autocorrelation functions (F_{AA} and F_{BB}) are always greater than or equal to unity, indicating that lattice sites are more likely to be occupied by a particular species if there are others nearby. A lower movement rate (as for species B) leads to a species having larger correlations, as clusters break up more slowly. The cross-correlation function, F_{AB}, is always less than or equal to one as the two species are inversely correlated; the presence of A in site l means it is less likely that B will occupy a nearby site. The curves in each plot are for distances from Δ to 3Δ in steps of 0.5Δ.

at the steady state. Our correlations PDE method provides a good approximation to the averaged individual behavior. Second, we look at a case whereby the MFA predicts different agent densities at steady state for the two species, but the averaged discrete data show that we expect the two species to have very similar agent densities at steady state [Fig. 2(b)]. We find that higher movement rates can compensate for decreased proliferation rates, a feature that is overlooked in the MFA. The last case we examine in detail is one in which the dominant species at steady state predicted by the MFA is not the same as that obtained from the averaged discrete behavior [Fig. 2(c)]. This shows that even if one species has a lower rate of proliferation, it is still able to maintain a higher long-term agent density, on average, if it has a high enough rate of movement compared to the other species. All of these examples demonstrate the disagreement with the averaged individual behavior as a result of assuming lattice site independence with the MFA. These errors can be successfully avoided by incorporating spatial correlations in lattice site occupancy into the model framework, as demonstrated in each of these cases by our correlations PDE model.

Additionally, we look at the behavior of the correlation functions predicted by our PDE model in Fig. 3. The autocorrelation functions, F_{AA} and F_{BB}, remain above unity, with similar behavior to that seen in previous work [14,27]. This indicates that each species is spatially correlated with itself: if one lattice site is occupied by a particular species, it is more likely that the same species will occupy a nearby lattice site. A greater maximum is observed for F_{BB}, for the values of s shown in Fig. 3, demonstrating that there are stronger correlations between species B than species A. This can be explained by the lower movement rate of species B, which leads to slower dispersal of clustered agents. In contrast to what has been seen previously, the cross correlation function, F_{AB}, remains below unity, thus the two species are inversely correlated with one another: if one site is occupied by species A, it is less likely that nearby sites will contain species B.
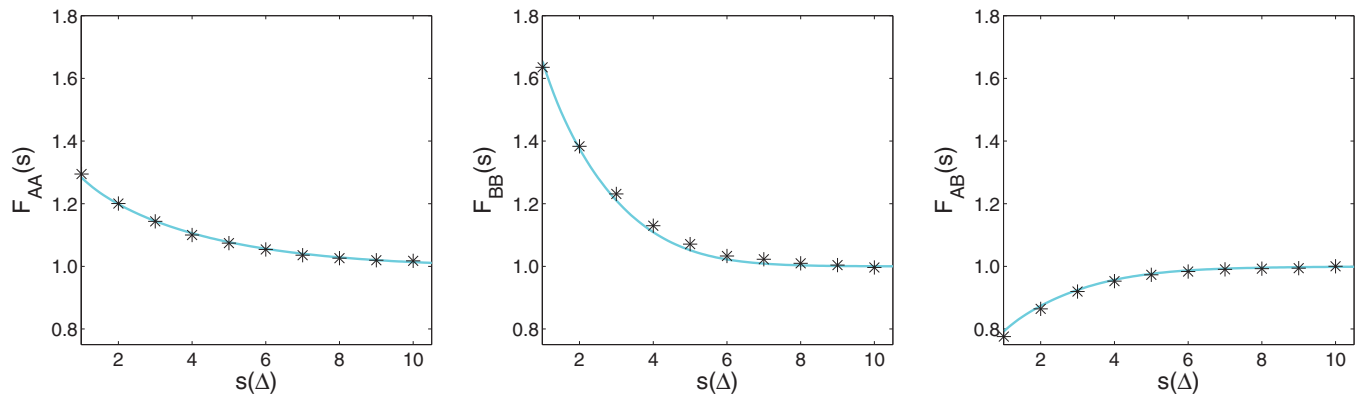

FIG. 4. (Color online) We examine the behavior of the correlations over space at a particular time ($t = 2$) for the case where $P_m^A = 20$, $P_p^A = 1$, $P_m^B = 5$, $P_p^B = 1$ for the correlations PDE method [solid line, using equations (2.23) to (2.28)]. We observe that the autocorrelation functions are above unity for smaller distances, decaying to unity as the distance between the lattice sites increases. The cross-correlation functions are below unity for smaller distances, again approaching unity as the distance between the lattice sites increases. We notice that F_{BB} is larger than F_{AA} as species B has higher levels of correlations due to its lower movement rate. Additionally, we include the averaged discrete results (asterisks), showing that the agreement between the correlations PDE method and the averaged discrete results is excellent.
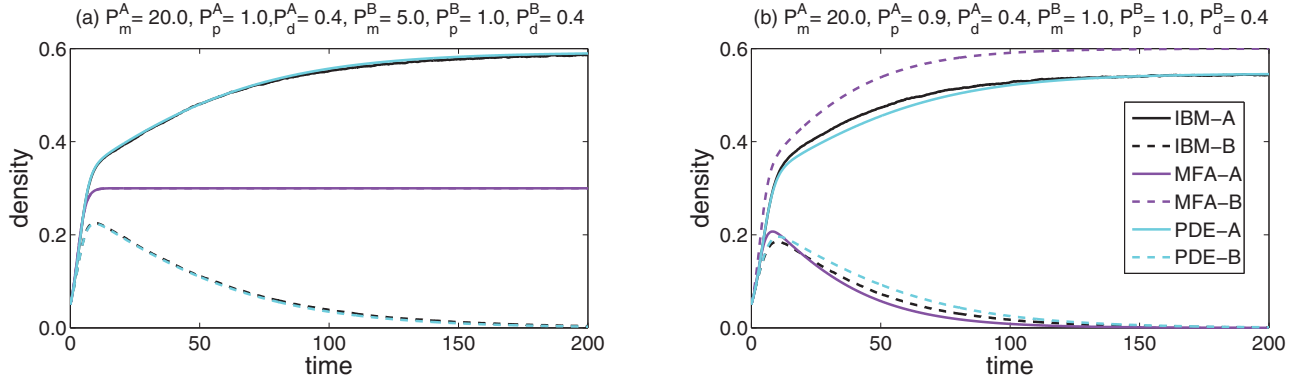
FIG. 5. (Color online) Including death in the system can lead to one species becoming extinct for the averaged discrete simulations and the correlations PDE method [using equations (2.23) to (2.28)]. The solid lines correspond to species A, while the dashed lines correspond to species B. The averaged discrete results are in black, the MFA in purple/dark gray and the correlations PDE in cyan/light gray. Here we see that the MFA does not agree with the averaged discrete behavior. In case (a), the MFA predicts that both species will coexist with the same steady-state, while the averaged discrete model and the correlations PDE method predict that species B will go extinct. In case (b), the MFA predicts that species A will go extinct, whereas the averaged discrete simulations predict that it is in fact species B that will go extinct. The correlations PDE method accurately predicts this behavior.

We also look at how the correlation functions behave over space, comparing the averaged discrete results with our correlations PDE model (Fig. 4). For the autocorrelation functions, F_{AA} and F_{BB}, we notice that the correlation functions decrease for increasing distances. They eventually reach a steady-state of unity, implying that lattice sites far apart are uncorrelated, which is what we anticipate. The correlation function for F_{BB} is larger than for F_{AA}. This is because, in this case, the movement rate for species B is smaller, therefore clusters of species B break up more slowly, leading to increased correlations between lattice sites. The cross-correlation function, F_{AB}, also has a value of unity for large distances due to lattice sites becoming independent as the distance between them increases. It is less than unity for smaller distances as site occupancy by the two different species is inversely correlated: if a site is occupied by species A, it is less likely that a nearby site will be occupied by species B.

B. Including death

Incorporating equal, nonzero death rates for each species into the $N = 2$ system usually leads, on average, to the species with the lower steady state without death now becoming extinct (Fig. 5), which is the absorbing state for the contact process [28]. For example, if we include equal nonzero death rates in the system in Fig. 2(a), the averaged discrete results show that we expect the species with the lower rate of movement to go extinct. This can be seen in Fig. 5(a). Additionally, an individual simulation with these parameters can be seen in Fig. 6, which shows the extinction of species B. Our correlations PDE method agrees with the averaged discrete results; however, the MFA predicts the species will coexist at the same levels. Moreover, if we include equal nonzero death rates in the system in Fig. 2(c), species B becomes extinct in line with the averaged discrete results. However, the MFA predicts that species A will go extinct, leaving species B. Our correlations PDE method accurately predicts that species B
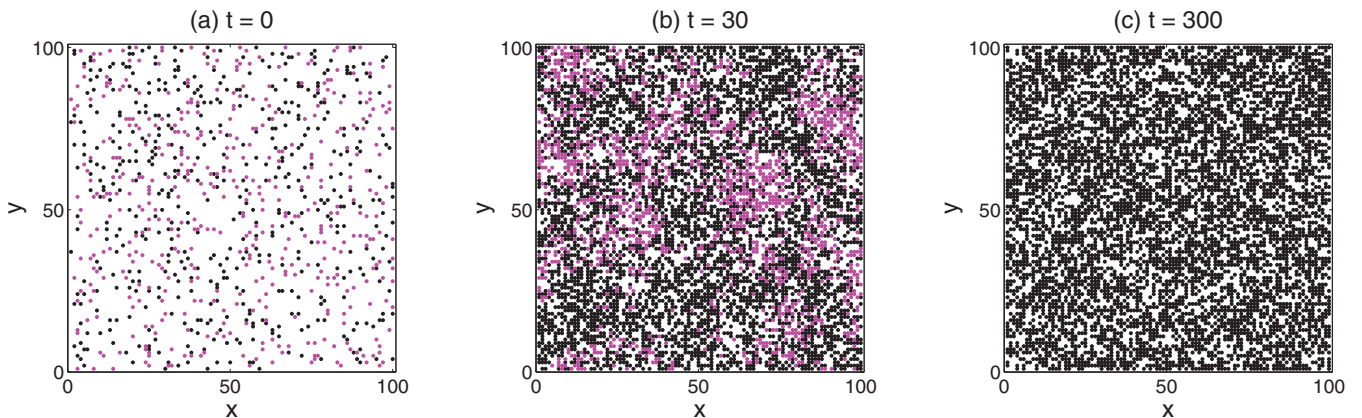

FIG. 6. (Color online) A single discrete simulation with nonzero death rates ($P_d^A = P_d^B = 0.4$) shows one of the species becoming extinct. Species A is in black, while species B is magenta/gray. In this case, $P_m^A = 20$, $P_p^A = 1$, $P_m^B = 5$, $P_p^B = 1$, corresponding to the averaged behavior seen in Fig. 5(a). While the two species have the same proliferation and death rates, the movement rate for species B is lower and leads to species B becoming extinct.

will go extinct, providing a good approximation to the averaged discrete results. This can be seen in Fig. 5(b).

IV. CONCLUSIONS AND FUTURE WORK

We have developed a method for incorporating PDE representations of spatial correlations into a volume-excluding model with multiple species undergoing proliferation, movement, and death on a regular lattice. We show the success of this method in accurately predicting the averaged discrete behavior for the two-species case, with and without agent death in the system. We note the success of this system in a number of cases where the MFA fails at predicting the averaged discrete behavior. Even though the species are not directly interacting, the presence of one species can affect the survival of another.

Currently, this correlations PDE method is only suitable for homogeneous initial conditions, as the assumption of translational invariance does not hold with inhomogeneous initial conditions. We would like to build on previous work [15] to deal with nonuniform initial conditions, such as would be necessary for an invading species. Additionally, it would be useful to incorporate direct interactions between the species. For example, the species may be competing for resources, or we could incorporate predation into the system.

The results shown here highlight the dangers of naively using the MFA to model situations where stochastic factors lead to spatial effects being important. Using the MFA we may predict co-existence of two species, when one of them may become extinct according to our averaged discrete results. Moreover, the MFA may even predict that a different species will become extinct in comparison to the averaged discrete results. This could have severe consequences if used to make predictions on the behavior of a system, as the predictions may be completely inaccurate. On the other hand, simply by including PDEs describing lattice site occupancy correlations, we are able to accurately predict the averaged discrete behavior, leading to a far more reliable framework for problems of this kind.

[1] T. A. Denison and Y. H. Bae, J. Control Release **164**, 187 (2012).

[2] M. Gerlinger, A. J. Rowan, S. Horswell, J. Larkin, D. Endesfelder, E. Gronroos, P. Martinez, N. Matthews, A. Stewart, P. Tarpey, I. Varela, B. Phillimore, S. Begum, N. Q. McDonald, A. Butler, D. Jones, K. Raine, C. Latimer, C. R. Santos, M. Nohadani, A. C. Eklund, B. Spencer-Dene, G. Clark, L. Pickering, G. Stamp, M. Gore, Z. Szallasi, J. Downward, P. A. Futreal, and C. Swanton, New Engl. J. Med. **366**, 883 (2012).

[3] A. Marusyk, V. Almendro, and K. Polyak, Nat. Rev. Cancer **12**, 323 (2012).

[4] J. Gurnell, *The Natural History of Squirrels* (Christopher Helm, London, 1987).

[5] K. MacKinnon, Mamm. Rev. **8**, 185 (1978).

[6] F. A. Brauer and C. Castillo-Chavez, *Mathematical Models in Population Biology and Epidemiology* (Springer, Berlin, 2001).

[7] J. D. Murray, *Mathematical Biology I: An Introduction* (Springer, Berlin, 2002).

[8] S. D. Webb, M. J. Keeling, and M. Boots, J. Theor. Biol. **249**, 140 (2007).

[9] M. J. Simpson, B. J. Binder, P. Haridas, B. K. Wood, K. K. Treloar, D. L. S. McElwain, and R. E. Baker, Bull. Math. Biol. **75**, 871 (2013).

[10] M. A. Lewis, J. Math. Biol. **41**, 430 (2000).

[11] G. B. Ermentrout and L. Edelstein-Keshet, J. Theor. Biol. **160**, 97 (1993).

[12] P. Hogeweg, Appl. Math. Comp. **27**, 81 (1988).

[13] J. Moreira and A. Deutsch, Adv. Complex Syst. **05**, 247 (2002).

[14] R. E. Baker and M. J. Simpson, Phys. Rev. E **82**, 041905 (2010).

[15] M. J. Simpson and R. E. Baker, Phys. Rev. E **83**, 051922 (2011).

[16] G. Ascolani, M. Badoual, and C. Deroulers, Phys Rev. E **87**, 012792 (2013).

[17] U. Dieckmann and R. Law, *Relaxation Projections and the Method of Moments* (Cambridge University Press, Cambridge, 2000), Chap. 21, pp. 412–457.

[18] S. P. Ellner, J. Theor. Biol. **210**, 435 (2001).

[19] R. Law, D. J. Murrell, and U. Dieckmann, Ecology **84**, 252 (2003).

[20] D. J. Murrell and R. Law, J. Anim. Ecol. **69**, 471 (2000).

[21] D. J. Murrell, U. Dieckmann, and R. Law, J. Theor. Biol. **229**, 421 (2004).

[22] C. E. Dangerfield, J. V. Ross, and M. J. Keeling, J. R. Soc. Interface **6**, 761 (2009).

[23] K. J. Sharkey, C. Fernandez, K. L. Morgan, E. Peeler, M. Thrush, J. F. Turnbull, and G. B. Bowers, J. Math. Biol. **53**, 61 (2006).

[24] K. J. Sharkey, Theor. Pop. Biol. **79**, 115 (2011).

[25] F. Barraquand and D. J. Murrell, Methods Ecol. Evol. **4**, 276 (2013).

[26] M. Raghib, N. A. Hill, and U. Dieckmann, J. Math. Biol. **62**, 605 (2011).

[27] D. C. Markham, M. J. Simpson, and R. E. Baker, Phys. Rev. E **87**, 062702 (2013).

[28] T. M. Liggett, *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes* (Springer, Berlin, 1999).

[29] J. T. Cox and R. B. Schinazi, Ann. Probab. **3**, 853 (2009).

[30] J. G. Kirkwood, J. Chem. Phys. **3**, 300 (1935).

[31] J. G. Kirkwood and E. M. Boggs, J. Chem. Phys. **10**, 394 (1942).

[32] J. Mai, V. N. Kuzovkov, and W. von Niessen, J. Chem. Phys. **98**, 10017 (1993).

[33] J. Mai, V. N. Kuzovkov, and W. von Niessen, Physica A **203**, 298 (1994).

[34] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, Cambridge, 2007).

[35] C. Zheng and G. D. Bennett, *Applied Contaminant Transport Modeling* (Wiley-Blackwell, London, 2002).