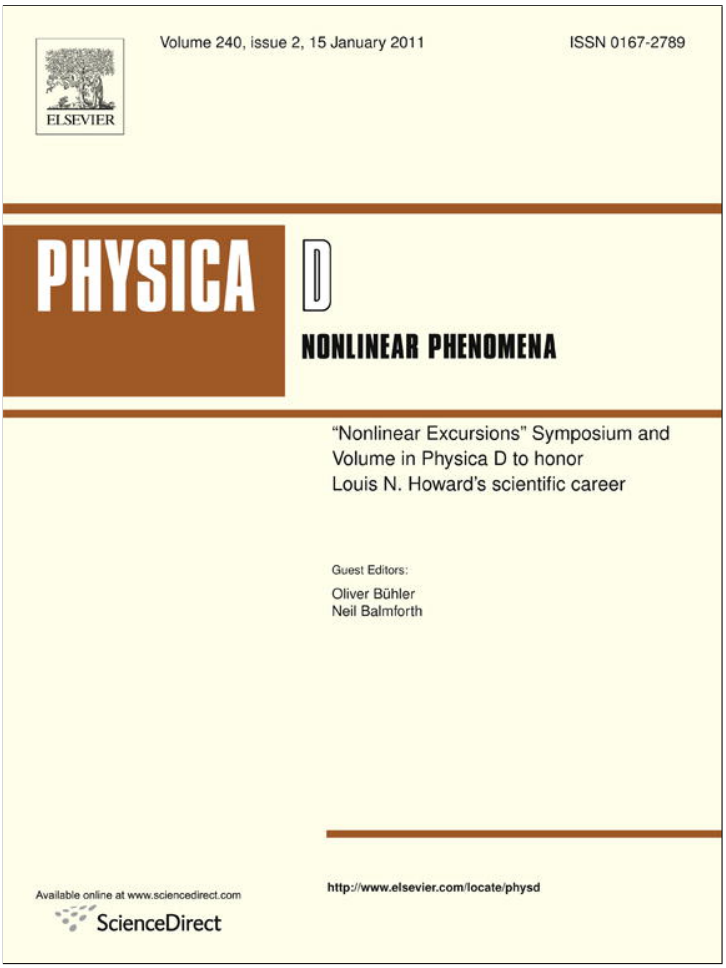

Physica D 240 (2011) 101–113

Contents lists available at ScienceDirect

Physica D

journal homepage: www.elsevier.com/locate/physd

An interacting particle model with compact hierarchical structure

N.J. Balmforth [a,*], W.R. Young [b]

[a] Departments of Mathematics and Earth & Ocean Science, University of British Columbia, Vancouver, Canada

[b] Scripps Institution of Oceanography, University of California at San Diego, La Jolla, CA 92093–0230, USA

ARTICLE INFO

Article history:
Available online 8 October 2010

Keywords:
Reaction–diffusion
Clustering
Particle probability density function
Interacting particle system

ABSTRACT

We present a model interacting particle system with a population of fixed size in which particles wander randomly in space, and pairs interact at a rate determined by a reaction kernel with finite range. The pairwise interaction randomly selects one of the particles (the victim) and instantly transfers it to the position of the other (the killer), thus maintaining the total number. The special feature of the model is that it possesses a closed hierarchical structure in which the statistical moments of the governing master equation lead to closed equations for the reduced distribution functions (the concentration, pair correlation function, and so on). In one spatial dimension, we show that persistent spatial correlations (clusters) arise in this model and we characterize the dynamics in terms of analytical properties of the pair correlation function. As the range of the reaction kernel is increased, the dynamics varies from an ensemble of largely independent random walkers at small range to tightly bound clusters with longer-range reaction kernels.

1. Introduction

The most challenging problems in statistical physics and population biology involve a large number of spatially distributed interacting individuals. Although these models are completely described by a deterministic master equation, this formulation does not usually provide a practical means of extracting exact statistical information. The problem is that most models with spatially localized interactions between individuals do not lead to a closed hierarchy of reduced distribution functions: in the analogue of the Bogoliubov–Born–Green–Kirkwood–Yvon (BBGKY) hierarchy, the evolution equation for the single-particle density function (the concentration) involves the pair correlation function, and the evolution equation for the pair correlation function involves the triplet function, and so on. This closure problem complicates models of biological populations and reaction kinetics starting with Doi [1,2] and continuing to recent work [3–10]. In this article, we formulate a model that has no such drawback: the BBGKY hierarchy closes and there is a compact many-body formulation.

The individual-based model studied here is an assembly of N organisms ("bugs") moving through continuous space and time via diffusion. Birth and death occur simultaneously when one bug kills another, and a new individual is instantly born at the same position as the killer.[1] The two momentarily coincident bugs then separate on independent random walks, as illustrated in Fig. 1 for one space dimension. In common with many other models, the spatial correlation introduced by the birth events leads naturally to persistent fluctuations in the local population density [11–14]. Houchmandzadeh has recently shown that these reproductive pair correlations result in clustering in an experimental ecosystem [15]. But simultaneous birth and death ensures that the total population is fixed at N.

The birth–death interaction occurs only between pairs of bugs that are sufficiently close. This ingredient is implemented with Doi's "reaction kernel" $\nu(r)$, where r is the pair separation. Specifically, the death rate of bug q, located at $\boldsymbol{x}_q$, is

$$\text{death rate of bug } q = \sum_{\substack{p=1 \\ p\neq q}}^{N} \nu(|\boldsymbol{x}_p - \boldsymbol{x}_q|). \tag{1}$$

The notion of spatially local interaction is incorporated by using a nonnegative $\nu(r)$ that decreases over a characteristic distance a (the range of the interaction) from a global maximum $\nu(0)$. We typically use the simple "top-hat" model:

$$\nu(r) = \nu_0 \begin{cases} 1, & \text{if } r < a; \\ 0, & \text{if } r > a. \end{cases} \tag{2}$$

* Corresponding author.
E-mail addresses: njb@math.ubc.ca (N.J. Balmforth), wryoung@ucsd.edu (W.R. Young).

0167-2789/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.physd.2010.09.016

[1] This formulation is very special: if the dead bug is reborn at the position of any of the other $N-1$ bugs, rather than only at the position of the killer bug, then an alternative formulation results that is far more complicated and has no closed hierarchy.

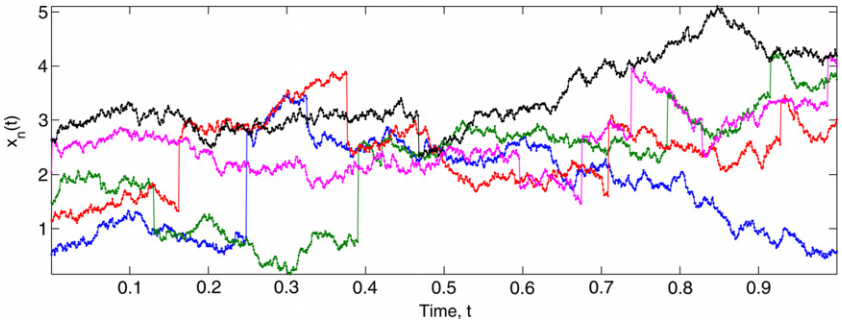

Fig. 1. Illustration of the model, showing the world-lines of $N = 5$ bugs. This simulation uses the top-hat reaction kernel in (2) with $\nu_0 = \kappa = 1$ and $a = 2$. The birth–death events are evident as the instantaneous translations resulting in momentarily coincident pairs. (To better illustrate the dynamics, this simulation uses a fixed time step, rather than the split-step event-driven scheme used for all our other simulations.)

The parameter ν_0 controls the overall death rate. Note that the sum above is also the *birth rate of bug q* because birth and death exactly cancel for every bug.

The model is similar to some earlier formulations, including the classic Ohta–Kimura, Moran and Fleming–Viot models in genetics and probability theory (see [16]). Almost identical fixed-N models have been reinvented in statistical physics [17,18,9]. The system studied here differs significantly from all these predecessors because the interaction encoded in $\nu(r)$ has finite range a. The earlier fixed-N models can be viewed as the special case of the current formulation in the limit that the interaction becomes infinite range ($a \to \infty$), and each bug interacts with all the other $N - 1$ bugs, no matter how distant. Our current model advances beyond these earlier theories, taking a step in the direction of greater realism by allowing the interaction to have finite range.

For $a \to \infty$, it is known that the N bugs quickly form a single large cluster which then walks randomly over the domain: a "wandering distribution" in Moran's [19] language. Fig. 2 illustrates (again in one dimension) how a finite range makes the dynamics of the model significantly richer: with $a < \infty$, multiple clusters of bugs persist indefinitely. The clusters nucleate from structureless initial conditions by collecting together all the bugs that are initially within range. Although the first clusters to appear are thereby isolated from one another, they subsequently wander and bump into each other. These collisions sometimes result in a merger into a single cluster, but on other occasions the clusters rebound almost like inelastic particles. Moreover, individual clusters occasionally break up into smaller units. Despite the complexity evident in these Monte Carlo simulations, we show here that basic statistical properties of this process can be found analytically. We thereby extract details of the long-time dynamics, and establish the controls set by the interaction length a. Thus the process illustrated in Figs. 1 and 2 is a canonical interacting particle model of cluster formation, with an important control parameter corresponding to the interaction range a. Analytic results can be obtained and used to understand some aspects of the dynamics. Though the model itself applies in spaces of arbitrary dimension, for simplicity, we focus on a single spatial dimension here.

2. Mathematical formulation

2.1. A review of the N-particle probability density function formalism

We begin by reviewing some mathematics required to formulate the master equation of the model described above (see, e.g., [20]). At time t, the population in a single realization is specified by the positions of all N bugs, which we gather together into the vector,

$$\boldsymbol{X} \equiv [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_N]\,. \tag{3}$$

The N-particle probability density function over this space is a nonnegative function, $F_N(\boldsymbol{X}, t)$, defined such that

$$F_N(\boldsymbol{X}, t)\mathrm{d}\boldsymbol{X} = \text{Prob}\{\text{a bug in } [\boldsymbol{x}_1, \boldsymbol{x}_1 + \mathrm{d}\boldsymbol{x}_1], \text{ another in } [\boldsymbol{x}_2, \boldsymbol{x}_2 + \mathrm{d}\boldsymbol{x}_2] \text{ etc.}\}, \tag{4}$$

and normalized with

$$\int F_N(\boldsymbol{X}, t)\mathrm{d}\boldsymbol{X} = 1. \tag{5}$$

Note that the dimensions of F are length^{-dN}, where d is the dimension of the space. The bugs are indistinguishable, so we can freely exchange $\boldsymbol{x}_p$ and $\boldsymbol{x}_q$; this is the permutation symmetry of F:

$$F_N(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p, \ldots, \boldsymbol{x}_q, \ldots \boldsymbol{x}_N, t) = F_N(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_q, \ldots, \boldsymbol{x}_p, \ldots \boldsymbol{x}_N, t). \tag{6}$$

As a simple example of an N-particle probability density function, suppose that the particles are dropped with uniform probability into a one-dimensional interval of length L; then $F_N = L^{-N}$.

Given $F_N(\boldsymbol{X}, t)$, we may compute the expectation of any function $A(\boldsymbol{X}, t)$ as the N-fold integral,

$$\langle A \rangle \equiv \int A(\boldsymbol{X}, t) F_N(\boldsymbol{X}, t)\mathrm{d}\boldsymbol{X}. \tag{7}$$

An important special case that guides us for defining the local concentration and pair function is given by choosing indicator functions of a particular subinterval of the total domain, $\chi(\boldsymbol{x}_p)$ (equalling unity if the pth particle lies within that subdomain and zero otherwise). More specifically, if $A(\boldsymbol{X}) = \sum_{n=1}^{N} \chi(\boldsymbol{x}_n)$, then $\langle A \rangle$ is the expected number of bugs in the subdomain, and using the permutation symmetry (6),

$$\left\langle \sum_{n=1}^{N} \chi(\boldsymbol{x}_n) \right\rangle = \int \chi(\boldsymbol{x}_1) C(\boldsymbol{x}_1, t)\mathrm{d}\boldsymbol{x}_1, \tag{8}$$

where

$$C(\boldsymbol{x}_1, t) \equiv N \int F_N(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_N, t)\mathrm{d}\boldsymbol{x}_2 \cdots \boldsymbol{x}_N \tag{9}$$

is the concentration. Likewise,

$$\left\langle \left(\sum_{n=1}^{N} \chi(\boldsymbol{x}_n) \right)^2 \right\rangle = \left\langle \sum_{n=1}^{N} \chi(\boldsymbol{x}_n) \right\rangle + \iint G(\boldsymbol{x}_1, \boldsymbol{x}_2, t)\chi(\boldsymbol{x}_1)\chi(\boldsymbol{x}_2)\mathrm{d}\boldsymbol{x}_1\mathrm{d}\boldsymbol{x}_2, \tag{10}$$

where the pair function is

$$G(\boldsymbol{x}_1, \boldsymbol{x}_2, t) = N(N-1) \int F_N(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \ldots \boldsymbol{x}_N, t)\mathrm{d}\boldsymbol{x}_3 \cdots \boldsymbol{x}_N. \tag{11}$$

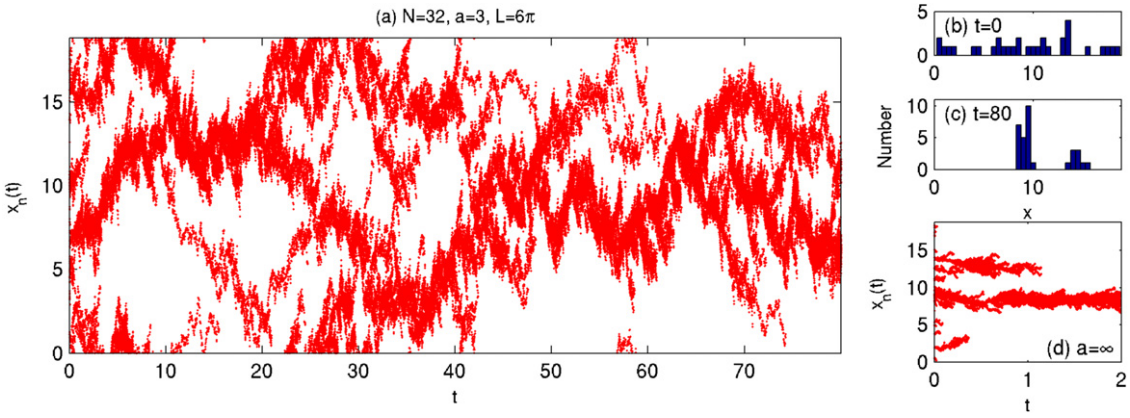

Fig. 2. The left panel shows results of a one-dimensional Monte Carlo simulation with $a = 3$ in (2) and $\kappa = \nu_0 = 1$. There are $N = 32$ bugs diffusing round a circle with perimeter $L = 6\pi$. The upper two panels on the right show histograms of the (random) initial and final positions. The lowest panel on the right shows that, with infinite range, effectively $a \geq L/2$, a single wandering cluster rapidly and permanently collects all the bugs.

Two interpretations of the pair function are helpful. First, $G(\boldsymbol{x}_1, \boldsymbol{x}_2, t)/C^2$ is the probability of having bugs at $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, relative to that which would obtain if the bugs were independently distributed. Second, $G(\boldsymbol{x}_1, \boldsymbol{x}_2, t)\mathrm{d}\boldsymbol{x}_1/C(\boldsymbol{x}_2)$ is the conditional probability of having an individual at $\boldsymbol{x}_1$, given that there is an individual at $\boldsymbol{x}_2$. The functions $C(\boldsymbol{x}_1, t)$ and $G(\boldsymbol{x}_1, \boldsymbol{x}_2, t)$ are just the first two reduced distribution functions in a hierarchy of N functions. We focus purely on C and G in the current study.

2.2. Particle interactions and the master equation

The N-particle probability density function $F_N(\boldsymbol{X}, t)$ satisfies a "master equation" corresponding to the fixed-N and finite-a model proposed in the introduction. The bugs wander independently around the domain until two of them interact, leading to a death; the dead bug is then instantly reborn at the position of the bug that killed it. This scheme automatically conserves the total particle number N, and is characterized by the total death rate obtained by summing (1) from $q = 1$ to N:

$$\text{Total death rate} = 2 \sum_{1\leq p<q\leq N} \nu(\boldsymbol{x}_p - \boldsymbol{x}_q). \tag{12}$$

By considering how the possible states of the system evolve, we arrive at the master equation,

$$(\partial_t - \kappa\Delta)F_N = 2 \sum_{1\leq p<q\leq N} \left[\delta(\boldsymbol{x}_p - \boldsymbol{x}_q) \int \nu(\boldsymbol{x}_p - \boldsymbol{x}'_q) \times F_N(\boldsymbol{x}_1, \ldots \boldsymbol{x}'_q, \ldots \boldsymbol{x}_N)\mathrm{d}\boldsymbol{x}'_q - \nu(\boldsymbol{x}_p - \boldsymbol{x}_q)F_N\right], \tag{13}$$

where

$$\Delta \equiv \nabla_1^2 + \nabla_2^2 + \cdots + \nabla_N^2 \tag{14}$$

is the sum of all N spatial Laplacians. The terms driving evolution in (13) account for spatial diffusion by random walking with diffusivity κ, and the birth and death process. For the latter, the total death rate in (12) determines the rate at which configurations disappear, giving the final summed term in (13); the terms of that sum represent the rate at which the pth bug kills its qth neighbour. Simultaneous births, however, reinject new configurations in which the interacting pairs are spatially coincident, generating the source terms containing $\delta(\boldsymbol{x}_p - \boldsymbol{x}_q)$; the strength of each source is dictated by the rate at which the pth bug kills the qth neighbour, which demands an integral over all possible positions of the latter before it was killed (the permutation symmetry is also exploited to reduce the number of pairings in the sum).

A crucial and dramatic feature of the master Eq. (13) is that, if one integrates over $\boldsymbol{x}_N$, the Nth Laplacian disappears (provided that the boundary conditions introduce no flux of new configurations), and a number of the birth and death terms cancel from the right, leaving a closed equation for the reduced density, $\int F_N \ \mathrm{d}\boldsymbol{x}_N$. Moreover, that equation is identical to the master equation for F_{N-1}. In other words,

$$F_{N-1}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N-1}, t) = \int F_N(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N-1}, \boldsymbol{x}_N, t)\mathrm{d}\boldsymbol{x}_N. \tag{15}$$

Thus by successively integrating over $\boldsymbol{x}_N, \boldsymbol{x}_{N-1}$ down to $\boldsymbol{x}_3$, we obtain the $N = 2$ model. But the $N - 2$-fold integral of F_N also provides the pair function G in (11). Hence, apart from the normalization,

$$N(N-1) = \iint G(\boldsymbol{x}_1, \boldsymbol{x}_2, t)\mathrm{d}\boldsymbol{x}_1\mathrm{d}\boldsymbol{x}_2, \tag{16}$$

the pair function satisfies the same equation as F_2:

$$(\partial_t - \kappa\Delta)G = 2\delta(\boldsymbol{x}_1 - \boldsymbol{x}_2) \int \nu(\boldsymbol{x}_1 - \boldsymbol{x}'_2)G(\boldsymbol{x}_1, \boldsymbol{x}'_2)\mathrm{d}\boldsymbol{x}'_2 - 2\nu(\boldsymbol{x}_2 - \boldsymbol{x}_1)G(\boldsymbol{x}_1, \boldsymbol{x}_2). \tag{17}$$

Taking a further integral of (17) over $\boldsymbol{x}_2$, and noting that

$$(N-1)C(\boldsymbol{x}_1, t) = \int G(\boldsymbol{x}_1, \boldsymbol{x}_2, t)\mathrm{d}\boldsymbol{x}_2, \tag{18}$$

we see that the concentration satisfies the $N = 1$ version of the model, which is simply the diffusion equation,

$$(\partial_t - \kappa\Delta)C = 0. \tag{19}$$

The distinguishing feature of the model is that G does not appear in (19), and the triplet function does not occur in (17). In other words, the hierarchy of equations for the reduced distribution functions is closed at each level, and, up to normalization, F_N for every N has the same concentration field and pair function (provided that the initial conditions are equivalent). This reductive feature is shared by the earlier infinite-range models, as well as some related theories of coalescing Brownian motion (see, e.g., [21]). Our quantitative analysis of the population dynamics is based on the two functions C and G.

2.3. Scaling and collective coordinates

The interaction kernels we have in mind possess an overall strength, ν_0, and a characteristic range, a. For example, we typically use the "top-hat" model in (2). Given ν_0 and the diffusivity, κ, we reduce the number of parameters by using nondimensional variables:

$$\hat{\boldsymbol{x}}_j = \sqrt{\frac{\nu_0}{\kappa}}\boldsymbol{x}_j, \quad \text{and} \quad \hat{t} = \nu_0 t. \tag{20}$$

This choice scales the diffusivity and pairwise death rate to unity, whilst keeping the number of particles, N, the nondimensional interaction range, $\hat{a} = a\sqrt{\nu_0/\kappa}$, and the characteristic size of the domain, $L\sqrt{\nu_0/\kappa}$, as dimensionless control parameters. After introducing this scaling, we drop the hat decoration from the dimensionless variables, and the only change to the governing equations is then the replacement of κ and ν_0 with unity.

To solve (17), it is also sometimes convenient to change variables to the "collective coordinates":

$$\boldsymbol{x} \equiv \frac{1}{2}(\boldsymbol{x}_1 + \boldsymbol{x}_2), \qquad \boldsymbol{y} \equiv \boldsymbol{x}_1 - \boldsymbol{x}_2, \\ G(\boldsymbol{x}_1, \boldsymbol{x}_2, t) = \mathcal{G}(\boldsymbol{x}, \boldsymbol{y}, t). \tag{21}$$

In terms of these variables, the pair equation is

$$\left(\partial_t - \frac{1}{2}\nabla_{\boldsymbol{x}}^2 - 2\nabla_{\boldsymbol{y}}^2\right)\mathcal{G} = 2\delta(\boldsymbol{y})\int \nu(\boldsymbol{y}')\mathcal{G}\left(\boldsymbol{x} - \frac{1}{2}\boldsymbol{y}', \boldsymbol{y}', t\right)\mathrm{d}\boldsymbol{y}' - 2\nu(\boldsymbol{y})\mathcal{G}. \tag{22}$$

Even simpler is the marginal density,

$$g(\boldsymbol{y}, t) \equiv \int \mathcal{G}(\boldsymbol{x}, \boldsymbol{y}, t)\mathrm{d}\boldsymbol{x}, \tag{23}$$

which satisfies

$$\frac{1}{2}g_t = \nabla_{\boldsymbol{y}}^2 g + \delta(\boldsymbol{y})\int \nu(\boldsymbol{y}')g(\boldsymbol{y}', t)\mathrm{d}\boldsymbol{y} - \nu(y)g. \tag{24}$$

Indeed, $g(\boldsymbol{y}, t)$ is our main tool to decipher the bug dynamics quantitatively.

3. One-dimensional dynamics in periodic domains

3.1. The initial-value problem

To complement our theoretical discussion of the model dynamics, we perform event-driven Monte Carlo simulations based on the algorithm of Gillespie [22]. The bugs are randomly scattered initially over a one-dimensional interval of length L with periodic boundary conditions (e.g., a circle with perimeter L), and then evolved using the scheme outlined in Appendix A. Details from a sample computation are shown in Fig. 3. As for the example shown earlier (Fig. 2), after a short transient, the random initial distribution of the bugs gives way to persistent spatial organization. The histogram of the final bug positions illustrates how the structure takes the form of loosely bound "clusters". To define this term more quantitatively, we count the number of intervals in the bug distribution with length greater than the interaction range a; these buffers divide the bugs into interacting spatial organizations, the clusters. Panel (c) of Fig. 3 shows how the number of cluster fluctuates irregularly in time. Note that, in a periodic domain, and with a sufficiently large N, there might be no empty intervals longer than a, which leads to a zero cluster count. Also shown in panel (d) is the fraction of the total number of pairs that are within range, which is another measure of the degree of spatial organization; if $L < Na$ and the bugs were uniformly distributed, this fraction would equal unity (and it would vanish if $L > Na$).

The spatial structuring is, at first sight, in contrast with what one expects from the fact that the bug concentration satisfies the diffusion equation (19). On the other hand, spatial correlations occur naturally because birth creates coincident pairs of bugs [12,11,14]. The concentration equation does not describe this reproductive clumping because C is an ensemble average over all possible realizations of the bug population. The ensemble has translation symmetry, and thus in the statistical steady state the concentration is uniform, and equal to N/L. The simplest statistic describing the structure evident in the realization shown in Fig. 3 is the pair function G.

Also shown in Fig. 3 are histograms of the bug positions and pair separations averaged over the duration of two longer simulations. These statistics provide empirical versions of the steady-state concentration, $C(x, t)$, the reduced pair function, and $g(y, t)$. The concentration field should, in principle, evolve to a constant level, but convergence is slow and the simulations still exhibit variations about this level by the end of the computations. The separation histogram in panel (f) demonstrates the spatial clustering, and is in agreement with the analytical solution derived below in (27).

As the interaction range increases, with the ratio a/L held fixed, so too does the degree of spatial correlation, largely because the bugs organize themselves through births and deaths without significant spatial movement. A second example, for $a = 8$ (shown in Fig. 4), displays the formation of more sharply defined clusters. The adjustment from the random initial condition gives way to a protracted phase of evolution in which the small clusters wander as coherent units, and thereby come into contact with their neighbours. The clusters then merge together to create larger units, so that the number of clusters decreases until eventually only a single cluster remains.

A summary of the observed cluster dynamics is provided in Fig. 5, which illustrates a series of initial-value computations with $N = 24$ bugs. Well-defined clusters emerge smoothly as we progress along the line $L = 40a$ in parameter space. Averages of the number of clusters and fraction of in-range pairs are shown against a in the lower panels of the same figure. The main panel also compares the location of the initial-value computations on the parameter plane with a simple estimate for the onset of pronounced clustering, determined as described below.

3.2. The one-dimensional steady-state pair function

In one dimension, and for the steady state, the pair equation (24) reduces to

$$g_{yy} - \nu(y)g + \delta(y)\int \nu(y')g(y')\mathrm{d}y' = 0, \tag{25}$$

which can be solved analytically for our top-hat reaction kernel (2).

If $L < 2a$, then the range of the interaction exceeds the domain size, each bug interacts with all of the other $N - 1$ bugs, and the model is equivalent to the existing infinite-range theories mentioned in the introduction. In this case, the solution of (25) is

$$g(y) = \frac{N(N-1)}{2L\sinh L/2}\cosh\left(\frac{L}{2} - |y|\right), \quad \text{for } |y| < L/2. \tag{26}$$

If $L > 2a$, on the other hand, a certain fraction of the total possible separations are out of range, and the solution of (25) is

$$g(y) = \frac{N(N-1)}{L(L + 2\sinh a - 2a)}\begin{cases}\cosh(a - |y|), & \text{if } |y| < a;\\ 1, & \text{if } a < |y| < L/2.\end{cases} \tag{27}$$

The solutions above are extended beyond the interval $-L/2 < y < L/2$, using the periodicity condition, $g(y + L) = g(L)$.

The solution in (27) implies that the probability that a separation $|y|$ is less than a is

$$\mathrm{Prob}\{|y| < a\} = \frac{2\sinh a}{L + 2\sinh a - 2a}, \tag{28}$$

which also equals the expected fraction of in-range pairs. Now, $\mathrm{Prob}\{|y| < a\} = 1/2$ when $L = L_{1/2}(a)$, where

$$L_{1/2}(a) \equiv 2\sinh a + 2a. \tag{29}$$

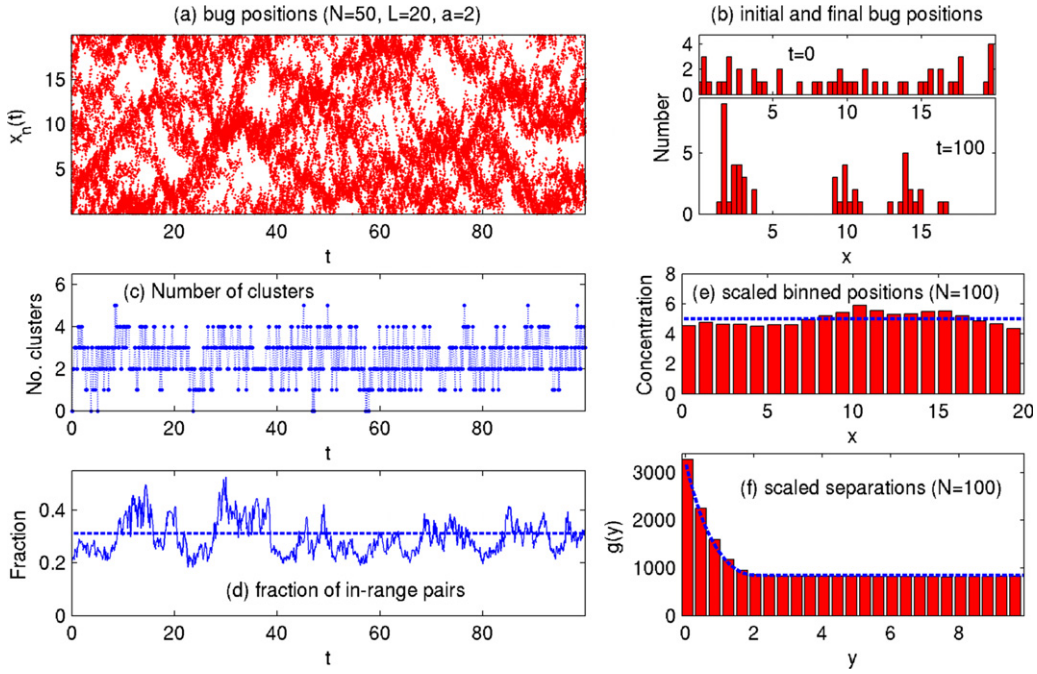

Fig. 3. Monte Carlo simulation with $N = 50$, $a = 2$, and $L = 20$. Panel (a) shows a time series of the bug positions; the initial and final histograms are shown in panel (b). Using the data in (a), in panel (c) we count the number of clusters by locating empty gaps of length greater than $a = 2$. Panel (d) shows the fraction of pair separations less than a (solid line), compared with the analytic prediction in (28) (the dashed line). Panels (e) and (f) show the empirical single-particle density and the density of pair separations; the dashed curves in panels (e) and (f) show the expected long-time distributions, $C = N/L$, and $g(y)$ from (27). To obtain the stable histograms in panels (e) and (f) we used simulations with $N = 100$ bugs, running up to $t = 10^3$ and averaged over two realizations.

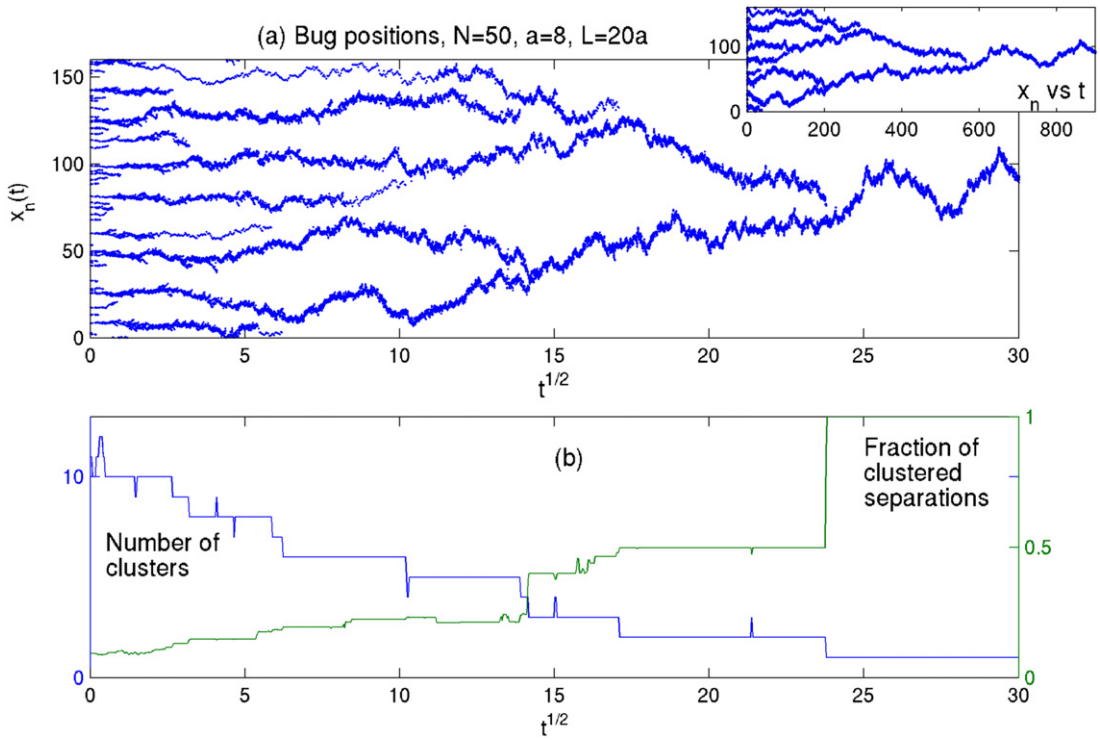

Fig. 4. Sample initial-value problem with $N = 50$, $a = 8$, and $L = 160$. Panel (a) shows the time series of bug positions, and in panel (b) we show the number of clusters and the fraction of "in-range" pair separations (the steady-state prediction (28) is 0.95). Both panels are plotted against $\sqrt{t}$ to highlight the short-time behaviour.

Thus, when $L < L_{1/2}(a)$, each pair separation is more likely to be in range than out of range; for $L > L_{1/2}(a)$, pair separations are likelier to be beyond the range of the reaction kernel. Thus a simple criterion to predict the onset of pronounced clustering is that $L < L_{1/2}(a)$. The solid curve in Fig. 5 is the function $L_{1/2}(a)$.

Because $L_{1/2}(a)$ increases exponentially with a, the expected fraction of in-range pairs in (28) rapidly approaches unity as one proceeds to the right of the curve in Fig. 5. Thus, the steady-state bug distribution in this region of parameter space must take the form of a single cluster for most of the time. Indeed, if all the bugs

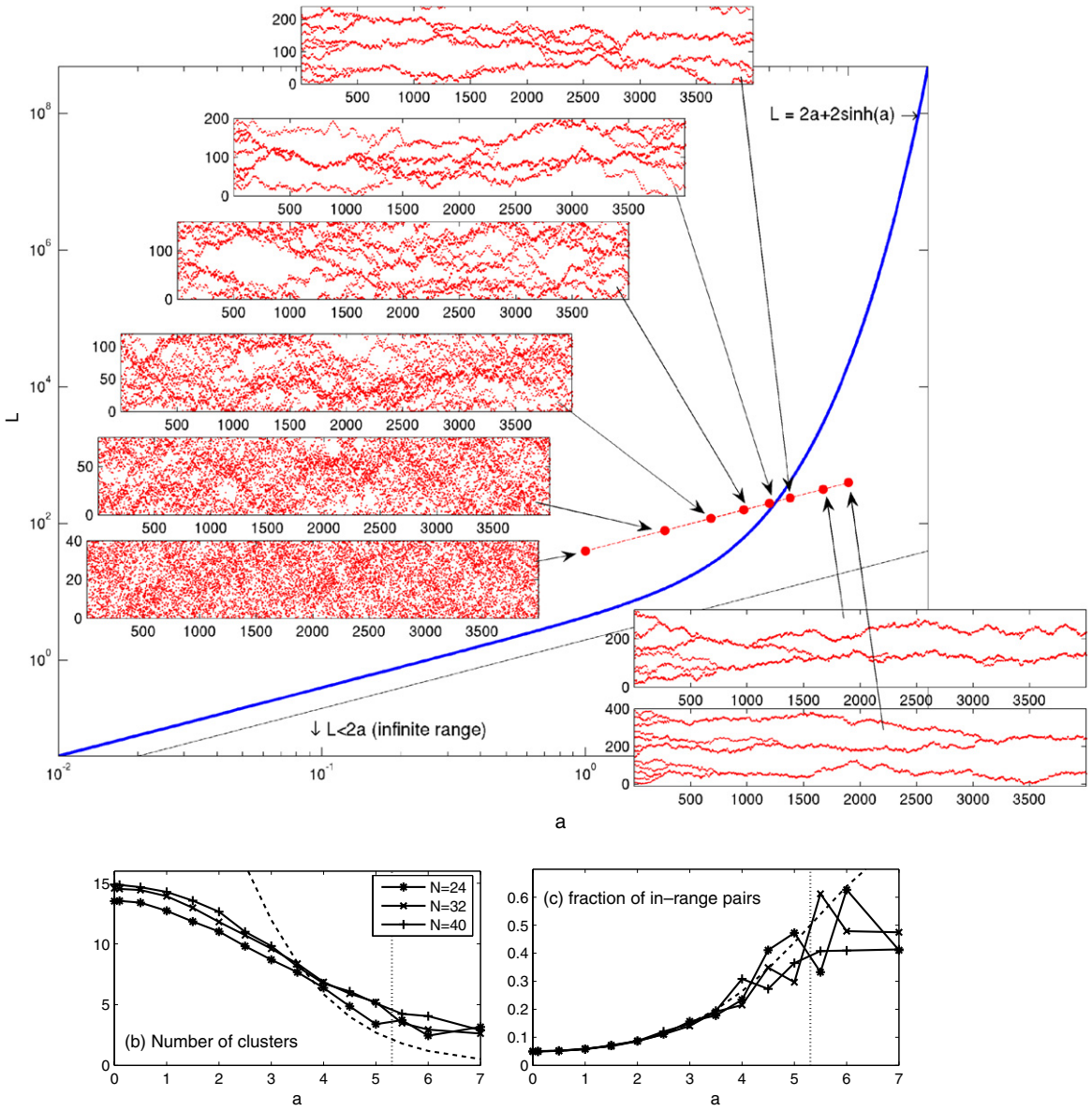

Fig. 5. Shown in the main panel is a plot of the (a, L)-parameter plane, showing the curve $L_{1/2}(a) \equiv 2\sinh a + 2a$, the line $L = 2a$ (below which the system effectively has infinite range), and the locations of several initial-value problems with $N = 24$ (lying along the line $L = 40a$). Time series of those computations are shown as insets. The lower panels show the number of clusters and fraction of in-range pairs, averaged over $1000 < t < 4000$, and at least one, and typically two, realizations for computations with varying N (as indicated) along $L = 40a$; the scatter in the simulation data at larger values of a reflects the failure of convergence to a steady state. The dashed lines in (b) and (c) show the crude estimate L/ℓ from (64), and the steady-state prediction of (28); the vertical dotted lines indicate $L = L_{1/2}(a)$.

are arranged into K clusters, then

$$\text{the number of in-range pairs} \le N\left(\frac{N}{K} - 1\right), \tag{30}$$

with the maximum achieved when each cluster contains N/K bugs (assuming that $N \gg 1$ so that the number in each cluster can be regarded as a continuous variable). Consequently, a steady state containing K clusters must satisfy

$$N\left(\frac{N}{K} - 1\right) \ge N(N-1)\,\text{Prob}\{|y| < a\} = N(N-1)\frac{2\sinh a}{L + 2\sinh a - 2a}. \tag{31}$$

The inequality above is equivalent to

$$K \le \frac{N(L + 2\sinh a - 2a)}{L + 2N\sinh a - 2a}, \tag{32}$$

which constrains the number of clusters that can persist for each set of parameters. But (32) cannot be satisfied for a given K if a is sufficiently large. In fact, with $K \ge 2$ and $N \gg 1$, we find that multiple clusters can only persist if $L > L_{1/2}(a)$. That is, the curve in Fig. 5 can also be regarded as the boundary below which only a single cluster can persist. The initial-value problems at larger a shown in this figure have not yet reached steady state, and have too many clusters, as highlighted by the final panel (see also Appendix A).

4. Single cluster dynamics in infinite domains

With finite range, the dynamics of a single, isolated cluster is the same whatever the spatial domain length, L. For example, the cluster at the end of the simulation in Fig. 4 is not affected by the finite size of the domain (this simulation could have been performed with any $L > 160$, as none of the bugs ever comes within an interaction distance of the domain's edge at

$x = L$). Likewise, the interaction between two colliding clusters is independent of the domain length provided that the cluster sizes are sufficiently less than L. Thus, in order to understand in more detail the dynamics of the Monte Carlo simulations presented above, we study clusters with $L \to \infty$, for which the pair function offers a selection of largely analytical and insightful results. The current section deals with a single cluster in order to understand the fragmentation dynamics seen in Figs. 2, 3 and 5; the next section is aimed at uncovering further details of cluster interaction.

4.1. Monte Carlo simulations

To consider the infinite-domain dynamics of a single cluster, we prepare the initial conditions by releasing all N bugs at the origin. The corresponding initial conditions for (17) and (19) are

$$C(x,0) = N\delta(x), \quad \text{and} \quad G(x_1,x_2,0) = N(N-1)\delta(x_1)\delta(x_2). \tag{33}$$

The diffusion equation for C then has a well-known Gaussian solution. The main property of this solution we use below is $\langle x_1^2\rangle = N^{-1}\int x^2 C\,\mathrm{d}x = 2t$, which says little about the internal dynamics of the cluster. To extract information about the cluster we must obtain $G(x_1,x_2,t)$, or more simply $g(y)$ in (23).

A suite of initial-value problems with varying a is shown in Fig. 6. In each case, the initial distribution spreads out rapidly. For smaller interaction length, the cluster breaks up as individuals escape, with the bug distribution evolving like an ensemble of independent random walkers (see the first panel for $a = 1$, which also shows the root-mean-square displacements, $\pm\sqrt{2t}$, expected for such walkers). With larger a, the diffusing bugs are unable to escape from their neighbours; the cluster stops expanding and maintains its integrity. As shown by the final example in the figure, the cluster then wanders coherently much like a single random walker; this analogy is made firmer below. Eventually clusters break up by splitting into subclusters (for the last realization with $a = 5$ in Fig. 6, the simulation must be continued for rather longer to observe such a fission).

To obtain a clearer picture of the break-up process, and in particular how often either individuals escape or the parent fragments, we performed further simulations, evolving the clusters until they first divide into two. The number of bugs in the smaller of the offspring is then recorded; histograms of this cluster size for 500 realizations are shown in Fig. 7 for varying a. These histograms are empirical measures of a "fragmentation" probability distribution for offspring size and are peaked at smaller numbers indicating that the original cluster is more likely to split into a bigger descendant with a smaller satellite. However, the chance of a more equable split increases with a.

The break-up of clusters in an infinite domain is foreshadowed by the pair equation (22), which does not have a steady-state solution if $L \to \infty$, with finite N and a, for the top-hat interaction kernel.[2] The interpretation of this result is now clear: a cluster cannot indefinitely maintain its integrity because individuals occasionally escape or the cluster splits into descendants, and the out-of-range bugs can then continue to walk randomly away from each other. Of course, if the domain were actually large but finite, the bugs could not escape each other forever, so the cluster would reform after a sufficiently long time, leading to the steady state in (27). These conclusions are partially dependent on the form of the interaction kernel, as we see next.

4.2. Bound clusters

We depart for the moment from considering only the top-hat interaction kernel to demonstrate that, if the decay of $\nu(y)$ is sufficiently slow as $|y| \to \infty$, then the bugs can never diffuse far enough from their neighbours to avoid interaction. In this case the N bugs are effectively permanently bound into a single cluster even if $L = \infty$. An extreme case is the infinite-range model, which leads to the steady solution,

$$g(y) = \frac{1}{2}N(N-1)\mathrm{e}^{-|y|}. \tag{34}$$

This solution is also obtained by taking the limit $L \to \infty$ in (26). We refer to compact, normalizable, $L = \infty$, steady solutions of (22), analogous to (34), as "bound clusters".

To construct a bound-cluster solution of (25) with a general reaction kernel, the differential equation

$$g_{yy} - \nu(|y|)g = 0 \tag{35}$$

must have a solution which decays to zero as $|y| \to \infty$. The decay must be rapid enough to ensure that $\int g\,\mathrm{d}y$ converges. Considering reaction kernels with power-law decay,

$$\nu(y) \sim |y|^{-2q}, \quad \text{as } |y| \to \infty, \tag{36}$$

it is straightforward to see that the condition for the existence of a bound-cluster solution is that $q < 1$. Specifically, if $q < 1$, then (35) has a rapidly decaying solution, namely $g \sim \exp(-|y|^{1-q})$.

The transitional case, $q = 1$, is illustrated by the reaction kernel in

$$g'' - \frac{g}{(1+|y|/a)^2} + \delta(y)\int \frac{g(y')\mathrm{d}y'}{(1+|y'|/a)^2} = 0, \tag{37}$$

where a is again the characteristic range of the algebraically decaying kernel. The bound-cluster solution of this problem is

$$g(y) = \frac{(\lambda-1)N(N-1)}{2a\,(1+|y|/a)^{\lambda}}, \tag{38}$$

where $\lambda(a) \equiv \left(\sqrt{1+4a^2}-1\right)/2$. For this solution to be normalizable, one requires $\lambda(a) > 1$, or $a > \sqrt{2}$; i.e., bound clusters exist only if $a > \sqrt{2}$.

4.3. Cluster statistics

Returning to the case of a top-hat reaction kernel, we use three statistical measures: the expected fraction of in-range pairs,

$$J(t) \equiv \int_{-a}^{a} g(y,t)\mathrm{d}y, \tag{39}$$

and the centre of population and the variance,

$$\bar{x} \equiv N^{-1}\sum_{n=1}^{N} x_n, \qquad \rho^2 \equiv N^{-1}\sum_{p=1}^{N}(x_n-\bar{x})^2. \tag{40}$$

The latter quantities are related to the sum of the squared displacements by the identity

$$\bar{x}^2 + \rho^2 = N^{-1}\sum_{n=1}^{N} x_n^2. \tag{41}$$

Using the average $\langle\cdots\rangle$ defined in (7) and the permutation symmetry, one can show that the average centre of population is independent of time, so $\langle\bar{x}\rangle = 0$ if the bugs all begin at the origin, and

$$\left\langle\bar{x}^2\right\rangle = N^{-2}\left[N\langle x_1^2\rangle + \iint x_1x_2G\mathrm{d}x_1\mathrm{d}x_2\right]. \tag{42}$$

Likewise, the average of (41) is

$$\left\langle\bar{x}^2\right\rangle + \left\langle\rho^2\right\rangle = \langle x_1^2\rangle. \tag{43}$$

[2] One can obtain a solution by taking $L \to \infty$ and $N \to \infty$ with N/L fixed in (27). However, we do not consider this "thermodynamic limit" further.

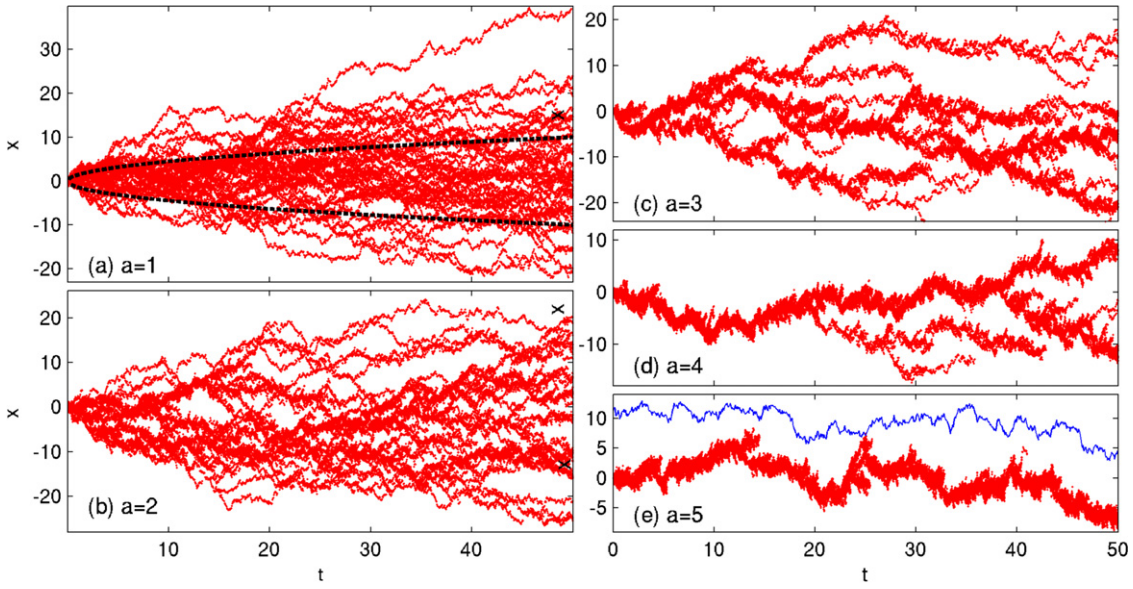

Fig. 6. A suite of initial-value problems in which 50 bugs leak out of a cluster with (a) $a = 1$, (b) $a = 2$, (c) $a = 3$, (d) $a = 4$, and (e) $a = 5$. In panel (a), the dashed lines show the root-mean-square displacement of a single random walker $\pm\sqrt{2t}$. Panel (e) also includes the path taken by a single random walker starting at $x = 10$.

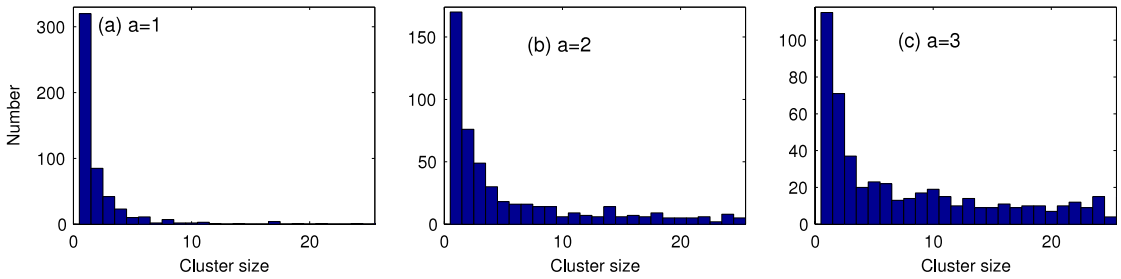

Fig. 7. Results of initial-value problems in which clusters (prepared initially by placing 50 bugs at the origin) are evolved until they first break up into two; the number in the smaller cluster is then recorded. For 500 realizations, this data is then collected into the histograms shown for $a = 1, 2$, and 3.

To obtain $\langle\bar{x}^2\rangle$ and $\langle\rho^2\rangle$, one needs $\iint x_1x_2G\mathrm{d}x_1\mathrm{d}x_2$ on the right of (42). Multiplying (17) by x_1x_2 and integrating,

$$\frac{\mathrm{d}}{\mathrm{d}t}\iint x_1x_2G\mathrm{d}x_1\mathrm{d}x_2 = \iint \nu(x_1-x_2)(x_1-x_2)^2G\mathrm{d}x_1\mathrm{d}x_2 = \int y^2\nu(y)\,g(y,t)\mathrm{d}y. \tag{44}$$

Thus J, $\langle x\rangle$ and $\langle\rho^2\rangle$ can be determined from the marginal density $g(y,t)$.

4.4. Infinite range

To recapitulate the much-studied case of infinite-range interaction, $\nu(y) = 1$, note that $\langle x_1^2\rangle = 2t$ and (44) is closed:

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}+2\right)\iint x_1x_2G\mathrm{d}x_1\mathrm{d}x_2 = 4N(N-1)t. \tag{45}$$

Solving this equation, and using the result in (42), one finds

$$\left\langle\bar{x}^2\right\rangle = 2N^{-1}t + \left(1-N^{-1}\right)\left(2t-1+\mathrm{e}^{-2t}\right). \tag{46}$$

Then from (43) the radius of the cluster is given by the square root of

$$\left\langle\rho^2\right\rangle = \left(1-N^{-1}\right)\left(1-\mathrm{e}^{-2t}\right). \tag{47}$$

Thus, when $t \ll 1$, the effective diffusivity of the centre of the cluster is N^{-1}, which is the centre-of-mass diffusivity of N independent random walkers. In this small-time regime, the cluster radius is $\sqrt{\langle\rho^2\rangle} \propto \sqrt{t}$. Once $t \gg 1$, the cluster stops expanding, with $\langle\rho^2\rangle \to 1 - N^{-1}$ and the diffusivity of the cluster centre equals that for a single bug. In other words, at small times the cluster centre behaves like the centre of N independent random walkers, but at large times the N bugs perform tightly correlated random walks, and the diffusivity of the cluster centre equals that of an individual.[3] This is illustrated in the final panel of Fig. 6, where $a = 5$ is equivalent to infinite range over the limited duration of the simulation.

4.5. Finite range

Turning back to the case of finite range (with the top-hat interaction kernel), in order to obtain the right-hand side of (44) we must solve the one-dimensional version of the pair equation (24) with the initial condition $g(y,0) = N(N-1)\delta(y)$. This initial-value problem succumbs to the Laplace transform; for some details see Appendix B. Our current efforts are directed at computing the cluster statistics, $J(t)$, $\langle\bar{x}^2\rangle$ and $\langle\rho^2\rangle$, for which the entire Laplace-transform solution is not needed. We find, for example, the Laplace transform of $J(t)$:

$$\hat{J}(s) = \frac{N(N-1)[(\cosh\eta a-1)\sqrt{s/2}+\eta\sinh\eta a]}{\sqrt{2s}+s(\sqrt{s/2}\cosh\eta a+\eta\sinh\eta a)}, \tag{48}$$

where $\eta = \sqrt{1+s/2}$.

[3] In dimensional terms, the effective diffusivity is κ/N for $\nu_0 t \ll 1$, and then it increases to κ for $\nu_0 t \gg 1$. The long-time cluster radius is $(1-N^{-1})\kappa/\nu_0$.

From the transform solution above, one can develop approximations for the statistics for both short and long times. For $t \ll 1$, the transform solution reduces to that for infinite range, and so the statistics are identical to those given earlier in Section 4.4: $J(t) \sim 1$, $\langle \bar{x}^2 \rangle \sim 2t/N$ and $\langle \rho^2 \rangle \sim 2t(1 - N^{-1})$. In the other limit, the long-time asymptotic behaviour of the Laplace transform is determined by the rightmost singularity in the complex-s plane, which is a branch point at $s = 0$. We therefore approximate (48) in the neighbourhood of $s = 0$ by

$$\hat{J}(s) \sim \frac{N(N-1)\sinh a}{\sqrt{2s} + s\sinh a}, \tag{49}$$

where the term $s \sinh a$ is retained in the denominator in order to take into account situations in which $\sinh a$ is large, so that for some range of intermediate times the contribution of $s \sinh a$ to the inverse transform is comparable to that of $\sqrt{2s}$. Inverting the transform in (49) leads to

$$J(t) \sim N(N-1)\mathrm{e}^{\gamma t}\mathrm{erfc}(\sqrt{\gamma t}), \tag{50}$$

$$\sim N(N-1)\begin{cases} 1 - 2(\gamma t/\pi)^{1/2} + \gamma t, & \text{if } 1 \ll t \ll \gamma^{-1}; \\ (\pi\gamma t)^{-1/2}, & \text{if } 1 \ll \gamma^{-1} \ll t, \end{cases} \tag{51}$$

where $\gamma \equiv 2/\sinh^2 a$. Thus, bugs escape from the reaction zone $|y| < a$ on the time scale γ^{-1}. This prediction is confirmed in Fig. 8, which shows time series of $J(t)$ for various values of a, computed by numerical inversion of $\hat{J}(s)$ using the algorithm of Hollenbeck [23], together with data from Monte Carlo simulations. The first panel also compares the numerical solution for $J(t)$ with the approximation in (50).

These results imply that, no matter how large a, on times of order γ^{-1}, clusters are likely to split. Thus, for example, the single cluster in Fig. 4 at $t = 1000$ is not permanent (given that the bugs are too far from the boundary to make the domain length relevant). However, with $a = 8$, $\gamma^{-1} > 10^6$, and it is difficult to illustrate the ultra-long time regime with Monte Carlo simulation.

The long-time approximation can further be used to calculate the wandering of the centre of population:

$$\langle \bar{x}^2 \rangle \sim \frac{2t}{N} + \left(1 - \frac{1}{N}\right)(\sinh a - a)\sinh a \times \left[\mathrm{e}^{\gamma t}\mathrm{erfc}(\sqrt{\gamma t}) - 1 + 2\sqrt{\frac{\gamma t}{\pi}}\right], \tag{52}$$

$$\sim \begin{cases} 2t, & \text{if } 1 \ll t \ll \gamma^{-1} \text{ and } \sinh a \gg a; \\ 2N^{-1}t, & \text{if } 1 \ll \gamma^{-1} \ll t. \end{cases} \tag{53}$$

The first limit in (53) is equivalent to the case of infinite range; the diffusivity of the cluster equals the single-particle diffusivity, and the cluster does not spread ($\langle \rho^2 \rangle = 2t - \langle \bar{x}^2 \rangle$ becomes constant). On the other hand, with a fixed and $t \to \infty$, we obtain the second limit in (53), implying that the cluster's centre has diffusivity N^{-1}, which is characteristic of N independent random walkers. Simultaneously, the cluster inexorably expands, with $\langle \rho^2 \rangle \propto t$. The dynamics is illustrated in Fig. 9, which once again compares the statistics computed from the Laplace-transform solution with Monte Carlo simulations. Note that the tightly bound clusters at larger a diffuse like a single random walker for substantial durations of time, as suggested earlier in Fig. 6.

To summarize, if $\sinh a \gg 1$, then there are three evolutionary stages: in terms of dimensional variables,

$$\langle \bar{x}^2 \rangle \sim 2\kappa t \begin{cases} N^{-1}, & \text{if } t \ll \nu_0^{-1} \ll \tau; \\ 1, & \text{if } \nu_0^{-1} \ll t \ll \tau; \\ N^{-1}, & \text{if } \nu_0^{-1} \ll \tau \ll t. \end{cases} \tag{54}$$

Above, the cluster break-up time scale, τ, corresponds to γ^{-1}, and in dimensional variables is

$$\tau \equiv \frac{\sinh^2(a\sqrt{\nu_0/\kappa})}{\nu_0}. \tag{55}$$

Throughout the process, the radius of the cluster is given by $\langle \rho^2 \rangle = 2\kappa t - \langle \bar{x}^2 \rangle$.

5. Cluster interaction

5.1. A Laplace-transform solution

To study cluster interaction, suppose that at $t = 0$ we have a cluster of N_- bugs at $-b/2$, and a second cluster of N_+ bugs at $b/2$. Using collective coordinates, the initial condition for the pair equation is

$$\mathcal{G}(x, y, 0) = N_+(N_+ - 1)\delta\left(x - \frac{1}{2}b\right)\delta(y) + N_-(N_- - 1) \times \delta\left(x + \frac{1}{2}b\right)\delta(y) + N_+N_-\delta(x)[\delta(y - b) + \delta(y + b)], \tag{56}$$

which translates to the initial condition for the marginal density of

$$g(y, 0) = \mathcal{N}_1\delta(y) + \mathcal{N}_2\left[\delta(y - b) + \delta(y + b)\right], \tag{57}$$

where

$$\mathcal{N}_1 \equiv N_+(N_+ - 1) + N_-(N_- - 1) \quad \text{and} \quad \mathcal{N}_2 \equiv N_+N_-. \tag{58}$$

Details of the Laplace-transform solution to this initial-value problem are once more provided in Appendix B.

The expected fraction of in-range pairs now has the transform solution

$$\hat{J}(s) = \frac{\mathcal{N}_1[(\cosh\eta a - 1)\sqrt{s/2} + \eta\sinh\eta a] + 2\mathcal{N}_2\eta\mathrm{e}^{-(b-a)\sqrt{s/2}}\sinh\eta a}{\sqrt{2s} + s(\sqrt{s/2}\cosh\eta a + \eta\sinh\eta a)}, \tag{59}$$

again with $\eta = \sqrt{1 + s/2}$. The term proportional to $\mathcal{N}_1$ reproduces our earlier solution; the second term, with $\mathcal{N}_2$, captures the effect of the initially displaced and out-of-range cluster.

As before, we offer an analytical approximation based on the long-time limit with $s \ll 1$, but keeping terms of order $\sqrt{s}\sinh a$, and now also including the terms with $(b - a)\sqrt{s} = O(1)$:

$$\hat{J}(s) \sim \frac{\left(\mathcal{N}_1 + 2\mathcal{N}_2\mathrm{e}^{-(b-a)\sqrt{s/2}}\right)\sinh a}{\sqrt{2s} + s\sinh a}, \tag{60}$$

implying for large t that

$$J(t) \sim \mathcal{N}_1\mathrm{e}^{\gamma t}\mathrm{erfc}\left(\sqrt{\gamma t}\right) + 2\mathcal{N}_2\mathrm{e}^{\gamma t + (b-a)\sqrt{\gamma/2}} \times \mathrm{erfc}\left(\sqrt{\gamma t} + (b - a)/2\sqrt{2t}\right). \tag{61}$$

5.2. Simulations

Fig. 10 shows computations beginning with two clusters of equal size, $N_+ = N_- = N/2 = 10$, for $a = 3$ and $b = 10$. For this parameter setting, the dynamics is rich because cluster break-up and interaction can occur on similar timescales. In some of the simulations, such as panel (a), the break-up occurs before clusters interact with one another; in other simulations, such as panel (b), the clusters collide to form a larger unit before breaking up. In yet further realizations, such as panel (c), the fragments interact whilst the clusters are breaking up. These different scenarios are blended together in the ensemble-averaged number of in-range pairs

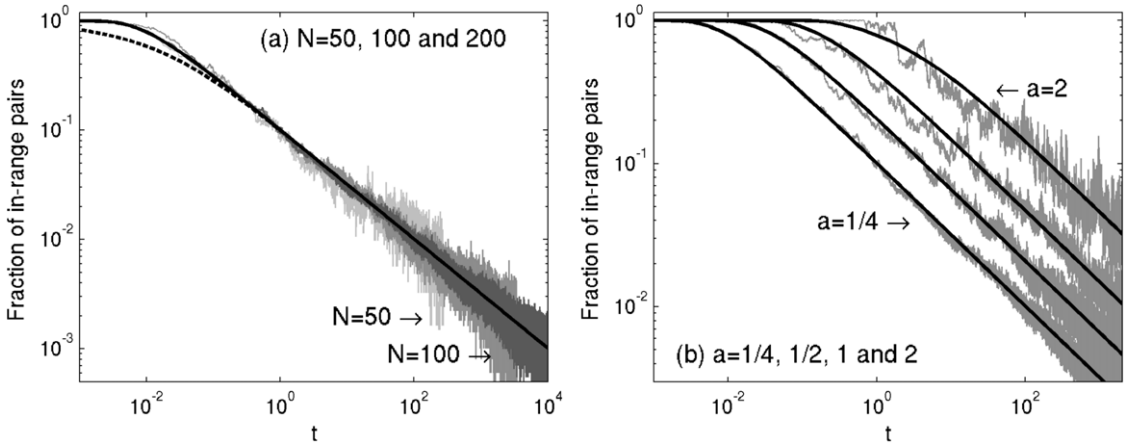

Fig. 8. The fraction of in-range pairs for simulations with (a) $a = 1/4$ and varying N. The data for $N = 50$ and 100 are shown in lighter shades of grey and end prematurely so that they can be more clearly distinguished. (b) $N = 100$ and varying a (as indicated). The solid lines indicate $J(t)$ computed by numerical inversion of the Laplace transform in (48) [23]; the dashed line in (a) shows the asymptotic approximation $e^{\gamma t}\mathrm{erfc}(\sqrt{\gamma t})$.

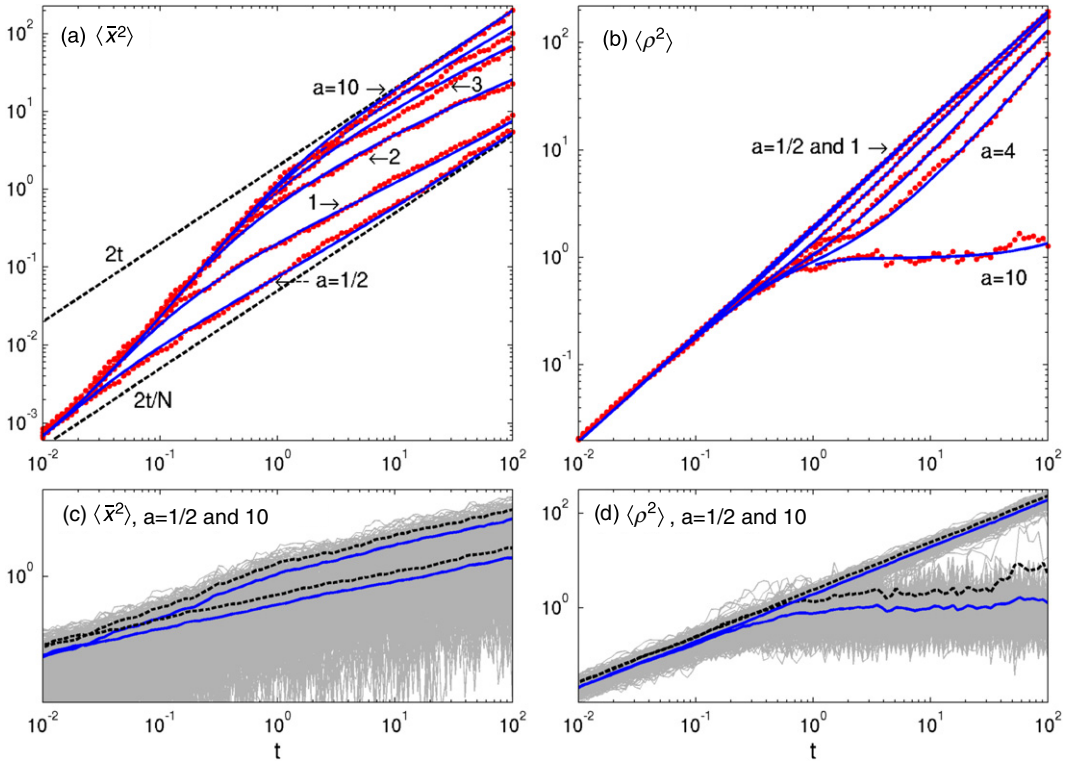

Fig. 9. (a) The mean-square displacement, $\langle \bar{x}^2 \rangle$, and in panel (b) the variance, $\langle \rho^2 \rangle$, computed numerically from the Laplace-transform solution of the initial-value problem for six different values of a (in an infinite domain). The dots show the results from simulations with $N = 40$ bugs, averaged over 200 realizations. To give an impression of the uncertainty in the simulation statistics, panels (c) and (d) show further results for $a = 1/2$ and 10: the light grey lines show the time series of all 200 realizations, the solid lines show $\langle \bar{x}^2 \rangle$ and $\langle \rho^2 \rangle$, and the dashed lines show those averages with addition of the standard deviation over the realizations.

shown in the right panel of Fig. 10. The corresponding statistics for the mean-square displacement and variance are shown in Fig. 11.

The local maximum in the expected number of in-range pairs evident in Fig. 10(d) arises from cluster collisions, an effect captured by the second error function in (61). The transient increase in $J(t)$ generated by this term begins for $t \sim (b-a)^2/8$, which is the characteristic collision time for the clusters. For $t \sim O\left(\frac{1}{4}(b-a)\sinh a\right)$, the collision-induced increase becomes balanced by the loss of in-range pairs incurred by cluster fragmentation, leading to the local maximum in $J(t)$.

A second example, with $a = 10$ and $b = 20$, is shown in Fig. 12. In this case, the break-up timescale is much longer, and the clusters maintain their identities throughout the simulations, except once they collide and merge together. Consequently, the number of in-range pairs increases monotonically throughout the simulation. The statistics of the mean-square displacement and variance are now dominated by the coherent wandering of the clusters and their collisions. To add further insight, we take the extra limit $\gamma t \ll 1$ in (61):

$$J \sim \mathcal{N}_1 + 2\mathcal{N}_2 \,\mathrm{erfc}\left(\frac{b-a}{2\sqrt{2t}}\right). \tag{62}$$

Note that all clusters eventually collide according to this formula, with $J \to \mathcal{N}_1 + 2\mathcal{N}_2 \equiv N(N-1)$ for $t \gg (b-a)^2/8$, as in the analogous problem of the absorption of a single random walker at a barrier placed initially at a distance $b - a$. The $\gamma t \ll 1$

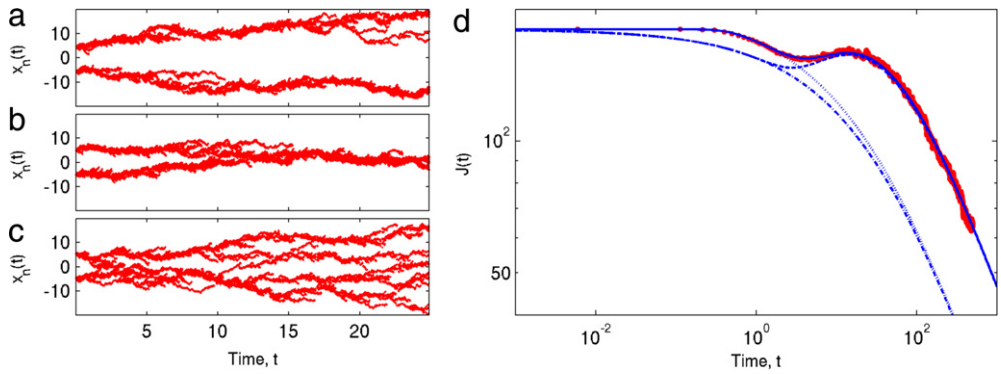

Fig. 10. Cluster interaction simulations for $a = 3$, $b = 10$, and $N = 20$, beginning with two clusters of equal size ($N_+ = N_- = 10$). Panels (a) through (c) show three different realizations, to illustrate the interaction dynamics. Panel (d) shows the expectation of the number of in-range pairs, $J(t)$, calculated either from the numerical inversion of the Laplace transform solution (59) (solid curve) or the approximation in (61) (dashed curve); the dotted and dashed–dotted curves without the local maximum near $t = 10$ show the expected number as though there were no cluster interaction (i.e. with $\mathcal{N}_2 = 0$). The dots in (d) show the number of in-range pairs averaged over 400 realizations.

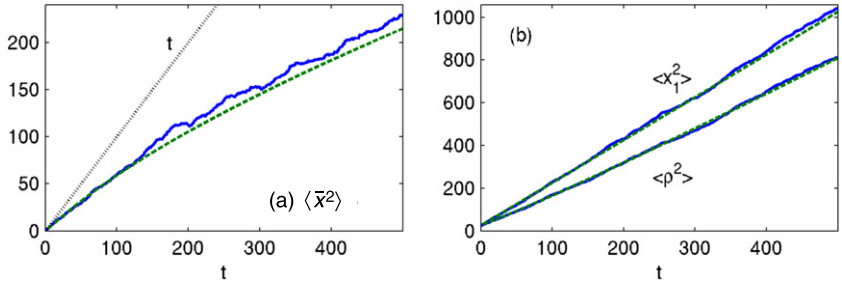

Fig. 11. Statistics of 400 realizations of the cluster interaction problem with $a = 3$, $b = 10$ and $N = 20$. Panel (a) shows the mean-square displacement, with the dashed line showing the expectation from the Laplace-transform solution (the dotted line is t). Panel (b) shows the statistics of $\langle x_1^2 \rangle$ and $\langle \rho^2 \rangle$, again with the expectations from the Laplace-transform solution.

approximation also leads to the results

$$\langle \bar{x}^2 \rangle \sim \begin{cases} t, & t \ll (b-a)^2/8, \\ 2t, & (b-a)^2/8 \ll t \ll \gamma^{-1}. \end{cases} \tag{63}$$

In other words, the statistics are initially dictated by the independent wandering of the two clusters (and $J \sim \mathcal{N}_1 = 180$ for the example in Fig. 12), which disperse from their original positions like two random walkers. The subsequent cluster collisions leave a decreasingly small number of realizations with two independent clusters, so that the statistics are ultimately dominated by the dynamics of a single cluster (with $J \sim N(N-1) = 380$). Fig. 12 illustrates this behaviour, showing the gradual switchover from two-cluster wandering to single-cluster diffusion for times of order $(b-a)^2/8 = 50$. Ultimately, for $t \sim O(\gamma^{-1})$, these approximations fail due to cluster fragmentation. On this ultra-long time scale we return to the scalings presented in Section 4.

6. Discussion and conclusion

In this article, we have presented a model of an interacting particle system which has the unusual feature that the statistical moments of the governing master equation generate closed equations for the reduced distribution functions. The relatively simple equations for the concentration and pair function can be solved to understand the dynamics illustrated by Monte Carlo simulations of the model. Some care is needed in order to decipher the information contained in the concentration and the pair function because those quantities are ensemble averages that respect the symmetries of the system (such as particle indistinguishability and translational invariance). In particular, the concentration satisfies the diffusion equation and tells us very little. Information regarding spatial structure, and specifically clustering, must be extracted from the pair function. Although we have thereby been able to characterize much of the dynamics of clusters, a deeper understanding (of, for example, the likelihood of multiple cluster states, or the fragmentation statistics in Fig. 7) requires us to proceed to the triplet function and beyond. Nevertheless, we can at least make some qualitative statements.

Simulations show that, in large domains with many particles, and higher values of a, the dynamics devolves to wandering clusters. Collisions between clusters can result in mergers that reduce the number of clusters. However, clusters can also fragment into smaller entities. These two processes must ultimately come into balance to establish a statistical equilibrium, so that clusters are separated by a typical distance ℓ which is determined by a competition between cluster break-up and collision. Equating the time scale for fragmentation, τ in (55), with the collision time scale, ℓ^2/κ, implies that the typical spacing between clusters is

$$\ell \equiv \sqrt{\kappa\tau} = \sqrt{\frac{\kappa}{\nu_0}} \sinh\left(a\sqrt{\frac{\nu_0}{\kappa}}\right). \tag{64}$$

Hence, if the domain has length L, we expect that there are roughly L/ℓ clusters, and each cluster contains on average $N\ell/L$ bugs. The estimate L/ℓ is included in Fig. 5, where it gives a rough guide to the number of clusters except for small interaction lengths. Of course, $N\ell/L \gg 1$, in order that the clusters are significant multiparticle organizations. Thus, finally, we can say that a multicluster state requires that both L/ℓ and $N\ell/L$ are large. In terms of external parameters, these two conditions require that

$$L \gg \sqrt{\frac{\kappa}{\nu_0}} \sinh\left(a\sqrt{\frac{\nu_0}{\kappa}}\right) \quad \text{and also}$$
$$N\sqrt{\frac{\kappa}{\nu_0}} \sinh\left(a\sqrt{\frac{\nu_0}{\kappa}}\right) \gg L. \tag{65}$$

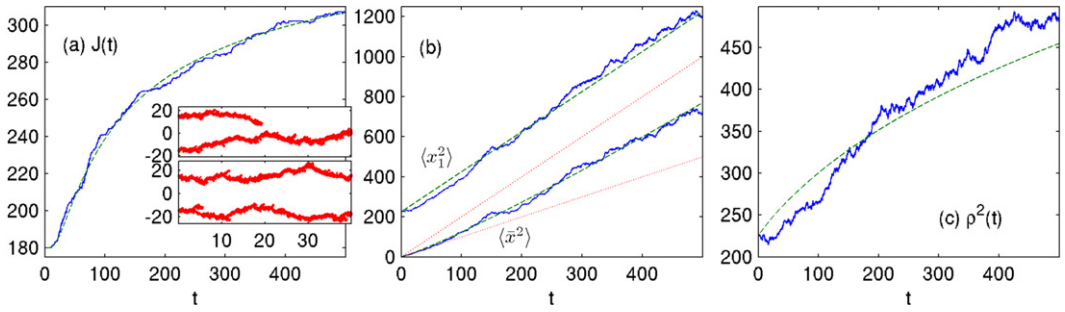

Fig. 12. Statistics of 200 realizations of the cluster interaction problem with $a = 10$, $b = 30$, and $N = 20$. Panel (a) shows the average number of in-range pairs, plus the Laplace-transform solution for $J(t)$, computed numerically [23]; the insets show two realizations, one with clusters that remain independent and the other with a collision. Panel (b) shows $\langle x_1^2\rangle$ and the mean-square displacement (the dotted lines show t and $2t$). Panel (c) shows the variance.

Acknowledgements

This work was partly performed at the Woods Hole Oceanographic Institution's Geophysical Fluid Dynamics Summer Program, which is supported by the National Science Foundation and the Office of Naval Research. WRY is supported by NSF OCE07-26320.

Appendix A. Details of the numerical algorithm

We summarize the split-step integration scheme used for the Monte Carlo simulations as follows [22]. Given the bug positions at time t, we compute the total death rate (12). The time to the next birth–death event, Δt, is then determined by random selection from an exponential probability density parameterized by that death rate. Based on the contribution of each pair of bugs to the death-rate sum in (12), we then draw another random number to select the interacting pair; a coin toss then decides which member of the pair is the killer, and the victim is then moved to the killer's position. Finally, the integration step is completed by allowing each bug to diffuse in space by taking an independent random step for a time Δt. The steps are selected from a Gaussian density with variance $2\kappa\Delta t$. Note that if all the bugs move out of range of one another then the death rate in (12) is zero, and consequently this algorithm results in $\Delta t = \infty$. This terminates the computation.

A limitation on the accuracy of the algorithm arises when the population becomes too dilute, so that the time step Δt is no longer small. In this situation, the random steps in space taken by the walkers can become appreciable, with the unfortunate consequence that the death rate changes significantly during Δt because bugs move in or out of range. Such errors are incurred when the number of bugs N is not large, the domain length L is relatively large, and the interaction range a is small (so that bugs do not cluster together).

A monitor of the accuracy of the algorithm is provided by the size of the time step, Δt. For example, for the simulations in Figs. 2–4, the typical time step was of order 10^{-3}, with steps occasionally becoming as large as 2.5×10^{-2}. For the simulations in Fig. 5, with fewer bugs, the typical and maximal time steps ranged from 0.03 and 0.6 at the smaller values of a (where the particles are most widely spaced), down to 0.005 and 0.05 for largest a (where tight clustering maintains small separations). The implied errors are unlikely to be excessive, and we attempted to further minimize their effect by computing multiple realizations or increasing the number of bugs whenever computational time permitted (see Fig. 5(b) and (c)).

Note that the oscillations in the data of Fig. 5(c) are not artifacts of errors in the integration algorithm. For the larger interaction ranges, the bugs cluster together tightly, and the time steps remain small. Instead, the oscillations arise because the population has insufficient time during the computation to reach its statistically steady state, and as a result, in some realizations, the bugs form multiple clusters that remain separated for the duration of the computation. At longer times, beyond our computational resources, these clusters would eventually merge to increase the number of in-range pairs. The convergence to the steady state can be explored by decomposing perturbations into normal modes; for large domains, one can establish that the decay rates of those modes are of order L^{-2}. Thus, relaxation occurs over times of order L^2, which is rather longer than the large-a computations in Fig. 5. The $O(L^2)$ relaxation can be interpreted alternatively as the characteristic time for clusters to wander across the domain and collide with one another; see Sections 4 and 5.

Appendix B. Details of the Laplace-transform solution

Using the Laplace transform,

$$\hat{g}(y,s) \equiv \int_0^\infty g(y,t)\mathrm{e}^{-st}\mathrm{d}t, \tag{B.1}$$

we can find the solution of (24) with the initial condition in (57). The solution in the range $0 < y < a$ is

$$\hat{g}(y,s) = A_1 \cosh \eta y + A_2 \sinh \eta y, \tag{B.2}$$

where $\eta \equiv \sqrt{1+s/2}$. Also, in (B.2),

$$A_1 \equiv -\frac{1}{2\eta}\left[\hat{J} + \frac{1}{2}\mathcal{N}_1\right],$$
$$A_2 \equiv \frac{1}{4\eta \sinh \eta s}\left[(s + 2\cosh \eta a)\hat{J} + \mathcal{N}_1(\cosh \eta a - 1)\right], \tag{B.3}$$

and $\hat{J}(s)$ is given in (59). For the single-cluster problem in Section 4, $b = 0$, $\mathcal{N}_1 = N(N-1)$, and $\mathcal{N}_2 = 0$; for the two-cluster solution in Section 5, these parameters are defined in (58).

Given this solution, we compute

$$\langle x_1 x_2\rangle = \mathcal{L}^{-1}\left\{\frac{1}{s}\int_{-a}^{a}\hat{g}(y,s)y^2\mathrm{d}y\right\} - \frac{b^2}{4(N-1)}, \tag{B.4}$$

where $\mathcal{L}^{-1}\{\cdots\}$ denotes the inverse Laplace transform. Then,

$$\langle \bar{x}^2\rangle = N^{-1}\langle x_1^2\rangle + (1-N^{-1})\langle x_1 x_2\rangle, \tag{B.5}$$

and

$$\langle \rho^2\rangle = (1-N^{-1})(\langle x_1^2\rangle - \langle x_1 x_2\rangle) \tag{B.6}$$

(given that $\langle x_1^2\rangle = b^2/4 + 2t$ and $\langle x_1 x_2\rangle = -b^2/4(N-1)$ at $t = 0$).

References

[1] M. Doi, Second quantization representation for classical many-particle system, J. Phys. A: Math. Gen. 9 (1976) 1465–1477.
[2] M. Doi, Stochastic theory of diffusion-controlled reaction, J. Phys. A: Math. Gen. 9 (1976) 1479–1495.
[3] D.A. Birch, W.R. Young, A master equation for a spatial population model with pair interactions, Theor. Popul. Biol. 70 (2006) 26–42.

[4] B.M. Bolker, S.W. Pacala, Using moment equations to understand stochastically driven spatial pattern formation in ecological systems, Theor. Popul. Biol. 52 (1997) 179–197.
[5] B.M. Bolker, S.W. Pacala, Spatial moment equations for plant competition: understanding spatial strategies and the advantages of short dispersal, Am. Nat. 153 (1999) 575–602.
[6] U. Dieckmann, R. Law, J.A.J. Metz, The Geometry of Ecological Interactions: Simplifying Spatial Complexity, Cambridge University Press, Cambridge, UK, 2000, 564+xiv pp.
[7] A.M. Etheridge, Survival and extinction in a locally regulated population, Ann. Appl. Probab. 14 (2004) 188–214.
[8] E. Hernández-García, C. López, Clustering, advection and patterns in a model of population dynamics with neighborhood-dependent rates, Phys. Rev. E 70 (2004) 016216.
[9] E. Hernández-García, C. López, Birth, death and diffusion of interacting particles, J. Phys.: Condens. Matter 17 (2005) S4263–S4274.
[10] R. Law, D.J. Murrel, U. Dieckmann, Population growth in space and time: spatial logistic equations, Ecology 84 (2003) 252–262.
[11] R. Adler, 1997, Superprocesses and plankton dynamics, in: Monte Carlo Simulation in Oceanography, Proceedings of the 'Aha Huliko'a Hawaiian Winter Workshop, University of Hawaii at Manoa.
[12] J. Felsenstein, A pain in the torus: some difficulties with models of isolation by distance, Am. Nat. 109 (1975) 359–368.
[13] B. Houchmandzadeh, Clustering of diffusing organisms, Phys. Rev. E 66 (2002) 052902.
[14] W.R. Young, A.J. Roberts, G. Stuhne, Reproductive pair correlations and the clustering of organisms, Nature 412 (2001) 328–331.
[15] B. Houchmandzadeh, Neutral clustering in a simple experimental ecological community, Phys. Rev. Lett. 101 (2008) 078103.
[16] A.M. Etheridge, An Introduction to Superprocesses, in: University Lecture Series, vol. 20, American Mathematical Society, Providence, 2000, 187+xii pp.
[17] Y.-C. Zhang, M. Serva, Polikarpov, Diffusion reproduction processes, J. Stat. Phys. 58 (1990) 849–861.
[18] M. Meyer, S. Havlin, A. Bunde, Clustering of independently diffusing individuals by birth and death process, Phys. Rev. E 54 (1996) 5567.
[19] P.A.P. Moran, Wandering distributions and the electrophoretic profile, Theor. Popul. Biol. 8 (1975) 318–330.
[20] N.G. van Kampen, Stochastic Processes in Physics and Chemistry, second ed., North-Holland, Amsterdam, 1997, 465+xiv pp.
[21] D. Ben-Avraham, S. Havlin, Diffusion and Reactions in Fractals and Disordered Systems, Cambridge University Press, UK, 2000, 316+xiv pp.
[22] D.T. Gillespie, A general method for numerically simulating the stochastic time evolution of coupled chemical reactions, J. Comput. Phys. 22 (1976) 403–434.
[23] K.J. Hollenbeck, 1998, INVLAP.M: a Matlab function for numerical inversion of Laplace transforms by the de Hoog algorithm. http://www.isva.dtu.dk/staff/karl/invlap.htm.