

This is a repository copy of *Living in groups: spatial-moment dynamics with neighbour-biased movements*.

White Rose Research Online URL for this paper:  
<https://eprints.whiterose.ac.uk/153151/>

Version: Accepted Version

---

**Article:**

Binny, Rachelle N, Law, Richard [orcid.org/0000-0002-5550-3567](https://orcid.org/0000-0002-5550-3567) and Plank, Michael (2019) Living in groups: spatial-moment dynamics with neighbour-biased movements. *Ecological Modelling*. 108825. ISSN 0304-3800

<https://doi.org/10.1016/j.ecolmodel.2019.108825>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

1 Living in groups: spatial-moment dynamics  
2 with neighbour-biased movements

3 Rachelle N. Binny<sup>1,2</sup>, Richard Law<sup>3</sup>, Michael J. Plank<sup>2,4,\*</sup>

4 1. Manaaki Whenua Landcare Research, Lincoln 7640, New Zealand.

5 2. Te Pūnaha Matatini, New Zealand.

6 3. Department of Biology, University of York, York YO10 5DD, United  
7 Kingdom.

8 4. School of Mathematics and Statistics, University of Canterbury, Christchurch  
9 8140, New Zealand.

10 \* Corresponding author email: michael.plank@canterbury.ac.nz

11 **Abstract**

12 Herd formation in animal populations, for example to escape a  
13 predator or coordinate feeding, is a widespread phenomenon. Under-  
14 standing which interactions between individual animals are impor-  
15 tant for generating such emergent self-organisation has been a key  
16 focus of ecological and mathematical research. Here we show the re-  
17 lationship between the algorithmic rules of herd-forming agents, and  
18 the mathematical structure of the corresponding spatial-moment dy-  
19 namics. This entails scaling up from the rules of individual, herd-  
20 generating behaviour to the macroscopic dynamics of herd struc-  
21 ture. The model employs a mechanism for neighbour-dependent,  
22 directionally-biased movement to explore how individual interac-  
23 tions generate aggregation and repulsion in groups of animals. Our  
24 results show that a combination of mutually attractive and repulsive  
25 interactions with different spatial scales is sufficient to lead to the  
26 stable formation of groups with a characteristic size.

27 Keywords: collective behaviour; herd formation; moment closure ap-  
28 proximation; neighbourhood interactions; spatial point process.

## 29 1 Introduction

30 The self-organisation of animals into herds, and the use of individual-based  
31 models to learn about the rules underlying this process, is a core subject in  
32 behavioural ecology (Krause et al., 2002). Herd formation is most often con-  
33 sidered in terms of movements of individuals, biased by their interactions at  
34 small spatial scales. However, these movements can affect the dynamics of  
35 populations and communities at larger spatial scales. In his seminal work,  
36 “Geometry for the selfish herd”, Hamilton (1971) proposed that aggrega-  
37 tion of animals into groups or herds, could be driven by the ‘selfish’ desire  
38 of an animal to reduce its predation risk by manoeuvring to positions that  
39 would place other population members closer to the predator. Underlying  
40 this idea was the concept of an animal’s *domain of danger*, a region of space  
41 containing all points nearer to that individual than to any other individual.  
42 The larger an animal’s domain of danger, the greater its risk of predation,  
43 and Hamilton therefore theorised that aggregation arose simply due to each  
44 animal undergoing movements towards its nearest neighbour, to reduce the  
45 size of its domain of danger. Stemming from this original idea, James et al.  
46 (2004) considered a model with greater biological realism, by incorporating  
47 a *limited domain of danger*, representing either a limited detection range  
48 or attack range of predators, that could be applied to animal groups of  
49 finite size. Further work by Reluga and Viscido (2005) pointed out that  
50 rules for generating realistic selfish herds need interactions beyond an in-  
51 dividual’s nearest neighbours, and showed how predation-based selection  
52 could increase the influence of distant neighbours. Other models explored  
53 animal aggregation behaviour by introducing sensory zones of individuals,  
54 for example zones of repulsion or attraction that drive animals towards or  
55 away from neighbouring individuals, giving rise to higher order structure in  
56 the population (Couzin et al., 2002; Wood and Ackland, 2007; Bode, 2011;  
57 Herbert-Read et al., 2011). One such model, proposed by Lukeman et al.  
58 (2010), used imagery data to infer individual zones of repulsion-alignment-  
59 attraction to describe self-aggregation in surf scoter flocks.

60 In addition to individual-based models, other common modelling ap-  
61 proaches for herd formation involve the use of mathematical equations of  
62 motion for individuals or populations. For example, “Lagrangian” equa-  
63 tions of motion describe individuals’ trajectories in terms of forces and  
64 velocities. “Eulerian” continuum equations (i.e. partial differential equa-  
65 tions), based on a diffusion approximation of random motion, are also  
66 widely employed to describe the evolution (in time and space) of mean-field  
67 density for swarms (Parrish and Edelstein-Keshet, 1999). The key problem  
68 with mean-field models is that they consider only the first spatial moment  
69 (the average density of individuals) and invoke an assumption that all in-

70 individuals interact in proportion to this average density (i.e. equivalent to  
71 assuming a well-mixed population or that all interactions are long-ranged),  
72 thereby ignoring any spatial structure in a population. This can give mis-  
73 leading results for systems where spatial structure is an important driver  
74 of the population dynamics (Law et al., 2003).

75 Models for the dynamics of spatial moments deal explicitly with local  
76 spatial structure, and avoid the limitations of mean-field models by us-  
77 ing higher-order spatial moments. The second spatial moment, i.e. the  
78 density of pairs of individuals as a function of their spatial separation,  
79 carries information on local spatial structure, and there is now a substan-  
80 tial body of theory for spatial-moment dynamics up to second order for  
81 birth-death-movement processes (Bolker and Pacala, 1997; Dieckmann and  
82 Law, 2000; Murrell and Law, 2003). This theory has been extended to  
83 consider multiple interacting species (Plank and Law, 2015), for example  
84 in predator-prey systems (Murrell, 2005; Barraquand and Murrell, 2013).  
85 A formal mathematical derivation that allows construction of a dynami-  
86 cal system for the second spatial moment in the presence of directionally-  
87 biased movement has been given by (Middleton et al., 2014; Binny et al.,  
88 2015, 2016a) and extended to include birth and death processes (Binny  
89 et al., 2016b). This mechanism for neighbour-dependent directional bias  
90 has been shown to be a strong driver of spatial structure, such as aggrega-  
91 tion, in motile cell populations (Binny, 2016). The directionally-biased  
92 movement modelling framework has been extended to multiple species by  
93 Surendran et al. (2018b) in the context of cell-obstacle interactions and by  
94 Surendran et al. (2018a) to chase-escape dynamics. However, directional  
95 movement of animals, as they respond to cues from their neighbourhoods,  
96 have not previously been part of this framework (but see Murrell and Law  
97 (2000) for nondirectional, environment-dependent movement).

98 Spatial moment dynamics are capable of providing mechanistic under-  
99 standing of the effects of individual interactions that repeated simulations  
100 of individual-based models alone cannot. Although it is not typically pos-  
101 sible to obtain closed-form solutions for the spatial moments, which must  
102 be approximated numerically, the structure of the equations can provide  
103 analytical insights into the relationships between model parameters and  
104 solutions. For example, spatial moment approximations have revealed:  
105 how and why spatial structure affects population carrying capacity (Law  
106 et al., 2003); new mechanisms for coexistence (Murrell and Law, 2003); the  
107 relative importance of different drivers of spatial structure (Binny et al.,  
108 2016b); and an analytical equivalence between mean population density  
109 and interaction range (Binny, 2016). Although straightforward to simu-  
110 late in principle, individual-based models are stochastic processes with a  
111 very high dimensional state space and are not amenable to analytical ap-

112 proaches except in special cases (Blath et al., 2007). In addition, although  
113 individual-based models are relatively efficient to simulate for small pop-  
114 ulations, the computational cost for models with interactions among in-  
115 dividuals increases faster than linearly with population size (Binny et al.,  
116 2016b). In contrast, the computational cost of solving a spatial moment  
117 dynamics approximation is insensitive to population size (Surendran et al.,  
118 2018b) so this represents an efficient alternative to individual-based models  
119 for large or growing populations.

120 The purpose of this paper is two-fold. First, we employ new mathemati-  
121 cal theory recently developed in the context of collective cell behaviour, that  
122 allows scaling up from directionally-biased agent movements to macroscopic  
123 dynamics (Binny et al., 2016a; Surendran et al., 2018b), and demonstrate  
124 how it can be applied in the ecological setting of herd formation in animals.  
125 The key mathematical expressions encoded in the rules of the individual-  
126 based model become clear in doing this. Secondly, we show that the spatial  
127 properties of herd formation are captured by the macroscopic dynamics,  
128 through appropriate choice of interaction kernels for directionally-biased  
129 movement. **This provides a foundation to bring biased movement**  
130 **into the earlier models of spatial-moment dynamics that focus on**  
131 **births, deaths and unbiased movement (Plank and Law, 2015).**  
132 **The framework will enable herd development to be studied in**  
133 **the broader context of population and community dynamics. To**  
134 **facilitate this future work, the mathematical derivations are given**  
135 **in a multi-species setting.**

## 136 2 Stochastic, individual-based model

137 Spatial-moment dynamics of birth, death and growth processes have been  
138 dealt with previously (Bolker and Pacala, 1997; Dieckmann and Law, 2000;  
139 Murrell and Law, 2003; Adams et al., 2013). Therefore here we con-  
140 sider only movement of individuals of fixed types. We first consider an  
141 individual-based model for motile agents. For generality, we allow individ-  
142 uals to be of an arbitrary number of types, indexed  $i \in \{1, \dots, i_{\max}\}$ . These  
143 could be species allowing, for instance, spatial interactions of predators and  
144 herd-living prey (the indexing can be ignored if all individuals are of the  
145 same type). Processes take place in a continuous two-dimensional space,  
146 which is large compared with the scale over which individuals interact and  
147 move; a point in the space is given by the vector  $x = (x_1, x_2)$  of Cartesian  
148 coordinates.

149 **2.1 Model for biased movement**

150 The population comprises a fixed number  $n$  of individuals numbered  $p =$   
 151  $1, \dots, n$ , and the state at time  $t$  is characterised by their types and locations  
 152  $(i_p, x_p)$ . Individual  $p$  moves in a series of discrete steps, which occur at a  
 153 rate  $M_p$  that may depend on its neighbourhood. This is a Poisson process  
 154 over time, so the probability of movement in a short period  $\delta t$  is  $M_p \delta t +$   
 155  $O(\delta t^2)$ . Movement events are assumed to occur as instantaneous jumps  
 156 (i.e. a position jump process). As soon as a movement takes place, the  
 157 state of the system is changed, potentially leading to a change in  $M_p$  as  
 158 well.

159 We allow both an intrinsic and a neighbourhood contribution to the  
 160 movement rate, given by

161 
$$M_p = m_{i_p} + \sum_{q \neq p} w_{i_p i_q}(x_p, x_q). \quad (1)$$

162 Here  $m_i$  is the intrinsic component of the movement rate for type  $i$ , and  
 163  $w_{i_p i_q}(x_p, x_q)$  is an extra contribution to the movement rate caused by a  
 164 neighbouring individual  $q$  of type  $i_q$  at location  $x_q$ . The contribution may  
 165 depend on the location and type of both  $p$  and  $q$ . The weight typically  
 166 attenuates with distance from  $p$  to  $q$  and could depend on whether individ-  
 167 ual  $q$  is the same species or, say, a predator species. The overall effect of  
 168 neighbours is obtained by summing over all  $q$ , excluding individual  $p$  itself.

169 When individual  $p$  moves from  $x_p$ , it jumps to another location  $u_p =$   
 170  $x_p + \xi$  where  $\xi$  is a random variable in  $\mathbb{R}^2$  with a bivariate probability  
 171 density function (PDF) of the form

172 
$$\hat{\mu}_p(\xi) = f_{i_p}(|\xi|) \hat{g}_p(\arg(\xi)), \quad (2)$$

173 where  $\arg(\xi) \in [0, 2\pi)$  denotes the direction of the vector  $\xi$ . The PDF in  
 174 Eq. (2) is separated into two independent parts for the distance moved  $|\xi|$   
 175 and the direction of movement  $\arg(\xi)$ . For simplicity, we assume that  $f_i(|\xi|)$   
 176 is neighbourhood-independent (though it may depend on the individual's  
 177 type  $i$ ) and given by the Gaussian function with mode  $r_i$  and variance  $s_i^2$ :

178 
$$f_i(|\xi|) = C_i e^{-\frac{(|\xi| - r_i)^2}{2s_i^2}}, \quad 0 \leq |\xi| \leq r_{i,\max}, \quad (3)$$

179 where  $C_i$  is a normalisation constant. In contrast to the distance moved,  
 180 the direction of movement does depend on the neighbourhood of individual  
 181  $p$ , and is the core mechanism underpinning herd development here. The  
 182 neighbourhood dependence takes the form of a bias vector  $\hat{\eta}_p$  for individual  
 183  $p$ , defined below, that provides the parameters for a circular probability  
 184 distribution for the direction of movement.

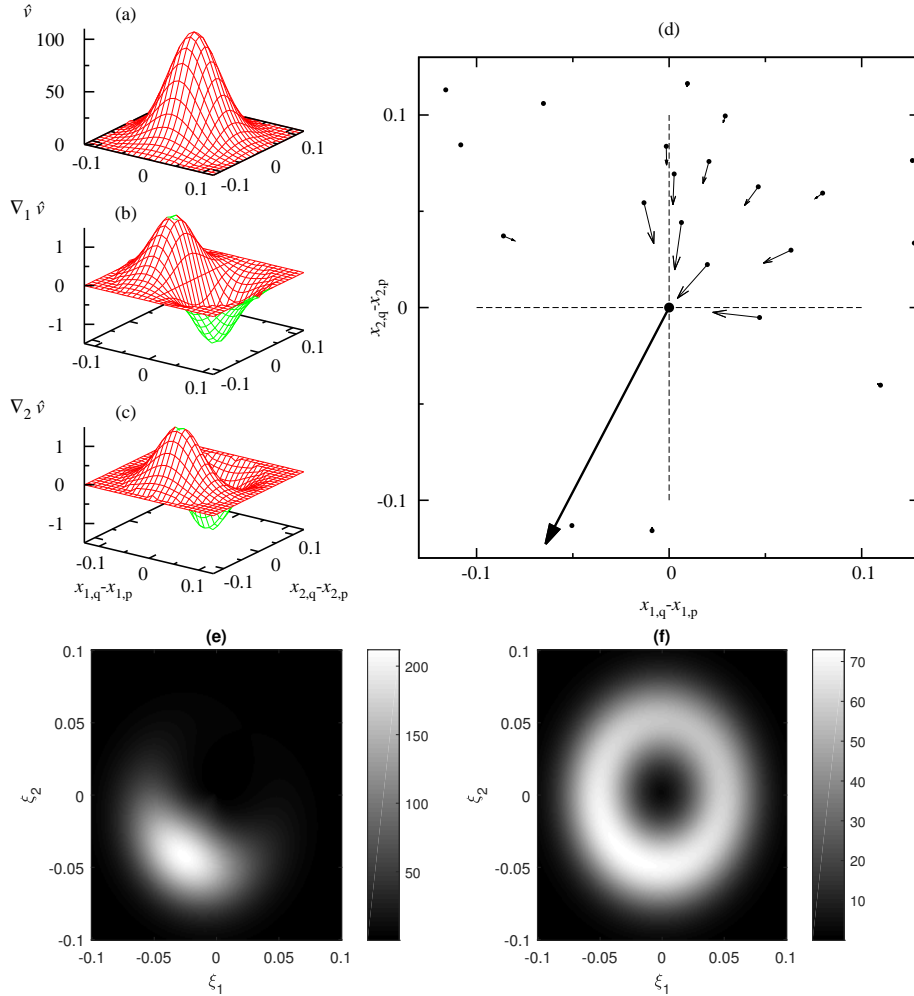


Figure 1: Schematic diagram showing how the bias vector and the movement distribution of a focal individual are constructed. (a) A bias kernel  $v$ , from which the gradient vector  $\nabla v$ , whose  $x_1, x_2$  components are plotted in (b) and (c), is obtained. (d) Contribution of neighbouring individuals (light arrows) to the bias vector of the focal individual at the origin (bold arrow). Note that the light arrows are not the bias experienced by the neighbouring individuals, but their contribution to the bias of the focal individual. The direction of the bias vector determines the preferred direction and its magnitude determines how tightly peaked the distribution is around the preferred distribution. Note the bias vector does not determine the new location of the focal individual. (e, f) Bivariate probability density function Eq. (2) for the movement vector  $\xi$  of the focal individual in the case of strong bias ( $\beta = 0.15$ ) and weak bias ( $\beta = 0.01$ ) respectively. Movement distance is distributed according to Eq. (3) with  $r = 0.05$ ,  $s = 0.02$ ,  $r_{\max} = r + 3s$ .

185 The bias vector is obtained from the gradient vector of a bias kernel  
 186 function that carries the key biological information. As an example, we de-  
 187 scribe the construction of a bias vector for a single focal individual located  
 188 at the origin in Fig. 1. This starts with a bias kernel function  $v_{i_p i_q}(x_q - x_p)$ ,  
 189 here a standard Gaussian function of the distance  $x_q - x_p$  between two indi-  
 190 viduals (Fig. 1a), potentially dependent on both the focal individual's type  
 191  $i_p$  and the neighbouring individual's type  $i_q$ . The kernel gives a gradient  
 192 vector  $\nabla v_{i_p i_q}(x_q - x_p)$ , i.e. the partial derivatives of  $v_{i_p i_q}$  in the two spatial  
 193 dimensions (Fig. 1b, c). The contribution of neighbouring individual  $q$  of  
 194 type  $i_q$  and location  $x_q$  to the bias vector of the focal individual  $p$  is the gra-  
 195 dient vector evaluated at  $x_q - x_p$  (light arrows on neighbouring individuals  
 196 in Fig. 1d). A neighbour vector that points towards the origin corresponds  
 197 to a repulsive effect of the neighbour on the focal individual (an outward  
 198 arrow would be an attractive effect). Summing all neighbour vectors gives  
 199 the bias vector for the focal individual (bold arrow on the focal individual  
 200 in Fig. 1d):

$$\hat{\eta}_p = \beta_{i_p} \sum_{q \neq p} \nabla v_{i_p i_q}(x_q - x_p), \quad (4)$$

202 where  $\beta_{i_p}$  is a parameter scaling the overall strength of bias. In the example  
 203 (Fig. 1d), the neighbourhood gives the focal individual a preferred direction  
 204 of movement away from the cluster of individuals on its upper right-hand  
 205 side. Note that changing the sign of the bias kernel in Fig. 1a would reverse  
 206 the direction of all arrows in Fig. 1d and hence produce an attractive rather  
 207 than a repulsive bias.

208 Once the bias vector  $\hat{\eta}_p$  for individual  $p$  is computed, its direction of  
 209 movement  $\theta$  is drawn from the von Mises distribution (independent of the  
 210 distance moved) with preferred direction  $\arg(\hat{\eta}_p)$  and concentration  $|\hat{\eta}_p|$ .  
 211 This distribution has probability density function

$$\hat{g}_p(\theta) = g(\theta, \hat{\eta}_p) = \frac{\exp(|\hat{\eta}_p| \cos(\theta - \arg(\hat{\eta}_p)))}{2\pi I_0(|\hat{\eta}_p|)}, \quad (5)$$

213 where  $I_0$  is the modified Bessel function of the first kind and zero order.  
 214 If the magnitude of the bias vector is large, the von Mises distribution is  
 215 tightly peaked around  $\arg(\hat{\eta}_p)$ , meaning the individual is highly likely to  
 216 move in a direction close to the preferred direction (Fig. 1e). This situation  
 217 would arise if the focal individual has multiple near neighbours exerting  
 218 bias in similar directions (as in the example in Fig. 1d). Conversely, if the  
 219 magnitude of the bias vector is small, the von Mises distribution is more  
 220 broadly distributed (Fig. 1f). In the limit where the bias vector has zero  
 221 magnitude, the von Mises distribution is a uniform distribution on  $[0, 2\pi)$ ,  
 222 meaning the focal individual is equally likely to move in any direction. This  
 223 situation would arise if the focal individual has no near neighbours, or has



224 neighbours that are symmetrically positioned on opposite sides such that  
 225 their contributions to the bias vector cancel one another out.

## 226 **2.2 Implementation**

227 We initialised realizations of the stochastic individual-based process with  
 228 a fixed population of  $n = 200$  individuals of a single type. The individuals  
 229 were distributed in a unit arena as a spatial Poisson process at the start  
 230 of each realization; in other words, each individual's location was chosen  
 231 uniformly at random and independently of all other individuals. Distances  
 232 are given relative to the unit of the arena. We used periodic boundary  
 233 conditions, and updated the state of the system in continuous time using  
 234 the Gillespie algorithm (Gillespie, 1977). For simplicity, we assumed the  
 235 movement rate to be independent of neighbourhood by setting  $w_{i_p i_q} = 0$   
 236 for all  $p$  and  $q$  in Eq. (1), leaving in place only an effect of neighbours on  
 237 the direction of intrinsic movements.

238 Eqs. (2)–(5) define the bivariate movement distribution of a focal in-  
 239 dividual  $p$ . Vectors  $\xi$  from this bivariate distribution were obtained by  
 240 independently generating the distance and direction of movement. The  
 241 probability that the distance moved  $|\xi|$  by an individual of type  $i$  lies in  
 242 the infinitesimal interval  $[r, r + dr]$  is  $r f_i(r) dr$ . Hence, movement distance  
 243 of an individual of type  $i$  has PDF

$$244 \quad h_i(r) = r f_i(r).$$

245 Random numbers from this distribution were generated via the following  
 246 rejection sampling algorithm:

- 247 1. Generate a normally distributed random number  $R \sim N(r_i, s_i^2)$
- 248 2. If  $R$  lies outside the interval  $[0, r_{i,\max}]$ , go to step 1. This results in a  
 249 sample from the distribution with PDF  $f_i(r)$  specified by Eq. (3).
- 250 3. Accept  $R$  with probability  $P(R) = R/r_{i,\max}$ , otherwise go to step  
 251 1. This results in a sample from the distribution with PDF  $h_i(r)$  as  
 252 required.

253 The direction of movement  $\theta$  was generated from the von Mises distribution  
 254 with PDF given by Eq. (5). This requires the bias vector  $\hat{\eta}_p$  for individual  
 255  $p$  to be calculated, according to Eq. (4).

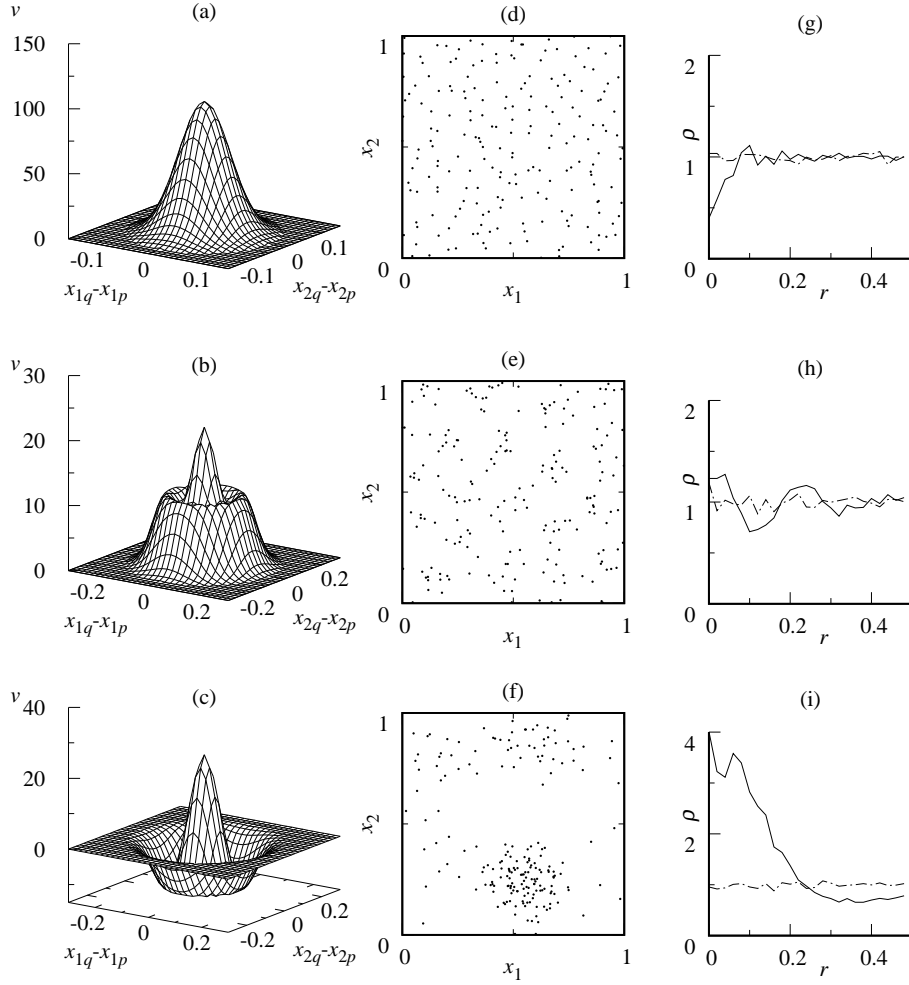


Figure 2: Territories and clusters developing from contrasting bias kernels. (a) A single positive Gaussian function Eq. (6) ( $\sigma_1 = 0.04$ ,  $N = 0.0099$ ) leads to formation of territories. (b) Adding a second Gaussian function, Eq. (7), that peaks at a distance  $\bar{r} = 0.12$  from the origin ( $\sigma_1 = \sigma_2 = 0.04$ ,  $N = 0.0477$ ,  $k_2 = 0.5$ ) leads to small clusters. (c) Subtracting a second Gaussian function, Eq. (7), that reaches its minimum at a distance  $\bar{r} = 0.12$  from the origin ( $\sigma_1 = \sigma_2 = 0.04$ ,  $N = 0.0401$ ,  $k_2 = -0.5$ ) leads to a single large cluster. Gaussian functions in the bias kernels were truncated at  $\pm 3$  standard deviations. Bias strength of the gradient vector  $\beta = 0.01$ . (d),(e),(f) Snapshots of locations of individuals at time  $t = 10$ ; the spatial patterns change continuously over time, starting from a spatial Poisson process. (g),(h),(i) Contrasting pair correlation functions  $\rho(r)$  of the spatial patterns develop by  $t = 10$  (continuous lines,  $\delta r = 0.02$ ); the dash-dot lines show  $\rho(r)$  at time  $t = 0$ . Neighbourhoods act only on the direction of movement here, not on the rate of movement. Movement distance is distributed according to Eq. (3) with  $r = 0.05$ ,  $s = 0.02$ ,  $r_{\max} = r + 3s$ . Movement rate  $m = 1$ .

256 **2.3 Biased-movement kernels and spatial structure**

The choice of a kernel for biased movement is a biological matter with far-reaching consequences. Fig. 2 gives three examples. The first is a single Gaussian function centred on the origin

$$(a): v_{i_p i_q}(r) = \frac{1}{N} e^{-r^2/2\sigma_1^2} \quad (6)$$

where  $r = |x_q - x_p|$  is the distance of neighbour  $q$  from focal individual  $p$ ,  $\sigma_1$  is a measure of the width of the function, and  $N$  is a normalisation constant. The second and third examples combine a Gaussian function centred on the origin with one offset from the origin by an amount  $\bar{r}$  and with width  $\sigma_2$ :

$$v_{i_p i_q}(r) = \frac{1}{N} (e^{-r^2/2\sigma_1^2} + k_2 e^{-(r-\bar{r})^2/2\sigma_2^2}) \quad (7)$$

257 the weight  $k_2$  of the outer function having different signs: (b)  $k_2 > 0$ , and  
 258 (c)  $k_2 < 0$ .

259 A kernel based on the single Gaussian function generates a gradient  
 260 vector that points towards the origin, creating a region of repulsion around  
 261 each individual. This means that individuals tend to move away from near  
 262 neighbours (Fig. 2a), leading to territory formation (Fig. 2d). A kernel  
 263 based on a double Gaussian function in which the outer Gaussian is positive  
 264 ( $k_2 > 0$ , Fig. 2b), generates three concentric rings: an inner ring where  
 265 the gradient vector points towards the origin, an intermediate ring where it  
 266 points away from the origin, and an outer ring where it points towards the  
 267 origin. This creates short-range repulsion, medium-range attraction and  
 268 long-range repulsion, leading individuals to form small clusters (Fig. 2e).  
 269 A kernel based on a double Gaussian function, in which the outer Gaussian  
 270 is negative ( $k_2 < 0$ , Fig. 2c), generates two concentric rings: an inner ring  
 271 where the gradient vector points towards the origin, and an outer ring  
 272 where it points away from the origin. This creates short-range repulsion  
 273 and long-range attraction, leading towards coalescence of the population  
 274 into a single mega-herd (Fig. 2f). The reverse order (attraction-repulsion)  
 275 would lead to collapse of individuals within groups to a single point, which  
 276 would not be not biologically reasonable.

277 Short-range repulsion (Fig. 2a, d) creates space around individuals, and  
 278 is a natural basis for territories, defended by individuals or groups, that  
 279 come about from scarcity of resources (Maher and Lott, 1995). Adding  
 280 longer-range attraction (Fig. 2c, f) allows for benefits of living in groups,  
 281 such as a reduced risk of predation, increased chance of detecting predators,  
 282 and less need for individual vigilance (Hamilton, 1971; Pulliam, 1973; El-  
 283 gar, 1989). With the short-range repulsion still in place, some space around

284 individuals remains and this can lead to remarkable spatial structure, such  
285 as that observed in king penguin colonies (Gerum et al., 2018). However,  
286 the combination of local repulsion and longer-range attraction can lead to  
287 very large groups forming (Olson et al., 2009). In practice, populations  
288 often break up into much smaller groups because of the costs of living to-  
289 gether, such as the need for synchronized behaviour (Gajamannage et al.,  
290 2017), levels of stress (Markham et al., 2015), possibly the spread of disease  
291 (Griffin and Nunn, 2012; Sah et al., 2017), and competition/cooperation  
292 between males (DuVal, 2007). Adding a further outer region of repulsion  
293 (Fig. 2b, e) allows break-ups to happen, the smaller groups being dis-  
294 tributed non-randomly over space, with spatial structure inside the groups  
295 themselves.

296 The spatial structures in Fig. 2 are clearly quite different, and this  
297 difference is summarised in their pair correlation functions (Fig. 2g,h,i). A  
298 pair correlation function  $\rho_{ij}(r)$  is a standard, second-order spatial statistic,  
299 based on the density of pairs of points of type  $i, j$  as a function of the  
300 distance  $r$  between them (Illian et al., 2008). In the absence of spatial  
301 structure at a distance  $r$ ,  $\rho_{ij}(r)$  takes a value 1; if there is an excess of  
302 pairs (clustering),  $\rho_{ij}(r) > 1$ ; if there is a lack of pairs (regular pattern),  
303  $\rho_{ij}(r) < 1$ . Thus the space that individuals create around themselves in  
304 Fig. 2d shows up as a lack of pairs at short distance in the pair correlation  
305 in Fig. 2g. The clusters that form in Fig. 2e appear as an excess of pairs at  
306 short distances in Fig. 2h, and a lack of pairs at slightly longer distances.  
307 The clusters themselves are not distributed at random across space, and  
308 leave an attenuating oscillatory signal in the pair correlation as distance  
309 increases. The location of the secondary peak in Fig. 2h at around  $r = 0.2$   
310 corresponds to the typical distance between adjacent clusters. The mega-  
311 herd developing in Fig. 2f appears as a large peak of pairs at short distances  
312 from the interaction of local repulsion and longer-distance attraction, with  
313 pairs becoming less frequent beyond the peak (Fig. 2i). The function does  
314 not tend to 1 at large distances, because the cluster is on the same spatial  
315 scale as the arena.

316 At a single point in time, repeated realizations of the stochastic processes  
317 from the same initial statistical distribution have different spatial configu-  
318 rations, but the same basic information is retained in the pair correlation  
319 functions. As time goes on, the spatial patterns change, and the pair cor-  
320 relation functions track the developing spatial structure. This tracking is  
321 evident in Fig. 2g,h,i. The realizations all started as Poisson processes lack-  
322 ing spatial structure, and with pair correlation functions close to 1 at all  
323 distances. But, by  $t = 10$ , the functions are quite distinct from one another,  
324 as shown in Fig. 2. The significance of the time-dependent pair correla-  
325 tion becomes important below, because a measure of this kind becomes the

326 state variable of the spatial-moment dynamics. In some ecological systems,  
 327 statistical stationarity may eventually be reached. But in others, such as  
 328 predator-prey systems, it is conceivable that the pair correlation functions  
 329 could develop periodic behaviour and continue to change indefinitely. The  
 330 long-term behaviour of the pair correlation function under a given choice  
 331 of bias kernel is not sensitive to the particular choice of initial conditions.

### 332 **3 Spatial-moment dynamics**

333 Here we show how the algorithmic rules of the individual-based stochastic  
 334 process can be described mathematically to give deterministic approxima-  
 335 tion in the form of a dynamical system for the second spatial moment.

#### 336 **3.1 Definition of spatial moments**

337 In defining the spatial moments, it helps to think of small regions of area  
 338  $h$ , so that the  $O(h^2)$  probability of containing more than one individual  
 339 is vanishingly small. Formally, the first spatial moment at time  $t$  is the  
 340 expected value of the density obtained from the stochastic process at time  
 341  $t$ , in the limit as  $h \rightarrow 0$ :

$$342 \quad Z_{1,i}(x) = \lim_{h \rightarrow 0} \frac{E[n_i(\delta x)]}{h}, \quad (8)$$

343 where  $n_i(\delta x)$  is the number of individuals of type  $i$  in the region  $\delta x$  centred  
 344 on  $x$ .

345 In the case of the second moment, we consider two regions of area  $h$ :  
 346  $\delta x$  centred on  $x$  containing  $n_i$  individuals of type  $i$ , and  $\delta y$  centred on  $y$   
 347 containing  $n_j$  individuals of type  $j$ . The second spatial moment at time  $t$  is  
 348 the expected value of the pair density from the stochastic process at time  
 349  $t$ , in the limit as  $h \rightarrow 0$  (Plank and Law, 2015):

$$350 \quad Z_{2,ij}(x, y) = \lim_{h \rightarrow 0} \frac{E[n_i(\delta x)n_j(\delta y) - \delta_{ij}n_i(\delta x \cap \delta y)]}{h^2}. \quad (9)$$

351 The second term in the numerator (with Kronecker delta  $\delta_{ij}$ ) is needed to  
 352 remove a pair that  $i$  in  $\delta x$  would otherwise create with itself. Below we  
 353 also use the third moment, the density of triplets  $Z_{3,ijk}(x, y, z)$ , defined in a  
 354 similar way after removing all non-distinct triplets (Plank and Law, 2015).

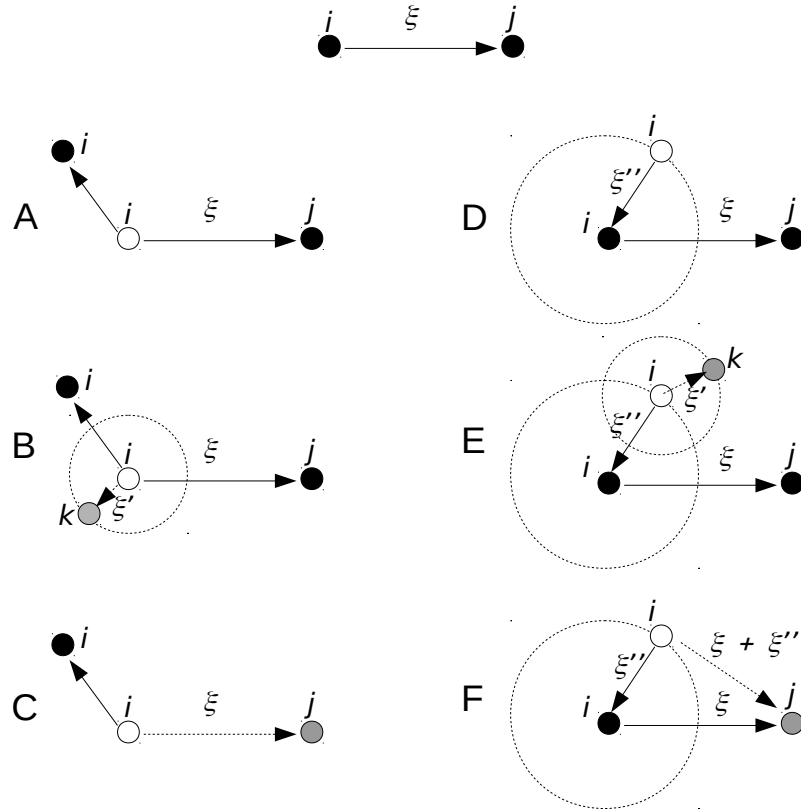


Figure 3: Geometry of the six flux terms A, ..., F in which movement of an individual of type  $i$  changes the pair density  $Z_{2,ij}(\xi)$  in a model of spatial-moment dynamics, numbered as described in the text. The object at the top is the  $ij$  pair: an individual of type  $j$  displaced by  $\xi$  from the focal individual of type  $i$ . Black-filled circles are locations of individuals after movement; empty circles are the positions from which they move; grey circles are neighbours that affect the movement; a dotted circle represents an integration over a neighbourhood; arrows are vectors. Geometries A, B, C in the first column destroy the pair; geometries D, E, F in the second column create the pair. A, ..., F are given as formal expressions (10), ..., (15) in the text.

## 3.2 Dynamics of the second moment

For simplicity, we consider dynamics in a homogeneous space, meaning that the statistics of the spatial point process in any subdomain are the same, regardless of the location of that subdomain. In this case, the first spatial moment  $Z_{1,i}$  is independent of spatial location  $x$ . Further, since the model consists only of movement and there is no birth or death, there is no change in first moment over time, so the first moment is simply a constant determined by the fixed population size. The second spatial moment  $Z_{2,ij}$  can be expressed as a function of the displacement vector between two individuals  $\xi = y - x$ , rather than as a function of their physical locations  $x$  and  $y$  (see Fig. 3 for geometric interpretation). Similarly, the third moment  $Z_{3,ijk}$  can be expressed in terms of two displacement vectors,  $\xi = y - x$  and  $\xi' = z - x$ .

Although the first moment is constant, the second moment does change over time as spatial structure develops, as was evident from the pair correlation functions in Fig. 2. The second moment and all higher moments are functions of time, but for clarity we omit the time argument below. The normalised second moment  $Z_{2,ij}(\xi)/(Z_{1,i}Z_{1,j})$  relates to the measure of spatial structure in Fig. 2g,h,i; it is the expected value of the pair correlation function  $\rho_{ij}(r)$  under isotropy. Thus, to follow the dynamics of the second moment is equivalent to following the behaviour of the average pair correlation function over time. In other words, the dynamics track the development of spatial structure over time. With  $Z_{2,ij}(\xi)$  as the state variable, we have a dynamical system describing changes in a function, as opposed to a dynamical system of a scalar quantity, the density of individuals (i.e. we have a partial as opposed to an ordinary differential equation). This is to be expected because the dynamical system has to carry information about the location of individuals relative to one another.

A formal derivation from the stochastic process (Binny et al., 2015, 2016a) leads to six terms affecting the rate of change in the second moment  $Z_{2,ij}(\xi)$  due to movement by the focal individual of type  $i$ , labelled (A)–(F) below and with geometries illustrated in Fig. 3. Symmetric terms corresponding to movement of the other individual (of type  $j$ ) in the pair are obtained by making the transformation  $\langle i, j, \xi \rightarrow j, i, -\xi \rangle$  to each of the terms below.

First are three negative terms that account for the ways in which an existing pair, consisting of a individual of type  $i$  separated from an individual of type  $j$  by a vector  $\xi$ , can be destroyed. Bias in the movement direction does not enter into these terms, because movement by the focal individual in any direction destroys the pair.

395 (A) Intrinsic rate of movement  $m_i$  of the focal individual:

$$396 \quad f_A = -Z_{2,ij}(\xi)m_i. \quad (10)$$

397 (B) Effect of the neighbourhood of the focal individual on its movement  
398 rate:

$$399 \quad f_B = -\sum_k \int Z_{3,ijk}(\xi, \xi')w_{ik}(\xi')d\xi'. \quad (11)$$

400 This incorporates the density of neighbours of type  $k$  displaced by  $\xi'$  from  
401 the focal individual (conditional on the presence of the individual of type  
402  $j$  displaced by  $\xi$  from the focal individual), given by the third moment  
403  $Z_{3,ijk}(\xi, \xi')$ . The kernel function  $w_{ik}(\xi')$  gives a weight to the effect of the  
404 neighbour on the movement rate of the focal individual. The overall effect  
405 of the neighbourhood is then obtained by integrating over all displacements  
406  $\xi'$  and summing over all types  $k$ .

407 (C) The other individual (of type  $j$ ) in the pair also affects the movement  
408 rate of the focal individual, with a contribution weighted by  $w_{ij}(\xi)$ :

$$409 \quad f_C = -Z_{2,ij}(\xi)w_{ij}(\xi). \quad (12)$$

410 Mirroring the negative terms are three positive terms that account for  
411 the ways in which a pair, consisting of an individual of type  $i$  separated  
412 from an individual of type  $j$  by a vector  $\xi$ , can be created. Since this can  
413 only occur via movement, this always starts with an  $ij$  pair separated by  
414 a different vector, denoted  $\xi + \xi''$ , followed by a movement by vector  $\xi''$ .  
415 These terms are more intricate than those in Eqs. (10)–(12) because they  
416 have to cover all possible starting locations for the focal individual and this  
417 needs to allow for bias in movement direction.

418 (D) Intrinsic movement rate of the focal individual, allowing for all starting  
419 points:

$$420 \quad f_D = m_i \int Z_{2,ij}(\xi + \xi'')\mu_{ij}(\xi'', \xi + \xi'')d\xi''. \quad (13)$$

421 Here, the term inside the integral is the probability of starting with an  
422  $ij$  pair separated by vector  $\xi + \xi''$ , followed by a movement by  $\xi''$  of the  
423 individual of type  $i$ , which happens with probability density  $\mu_{ij}(\xi'', \xi + \xi'')$   
424 (see below). This is then integrated over  $\xi''$  to allow for all possible starting  
425 locations.

426 (E) Effect of the neighbourhood of the focal individual on its movement  
427 rate, depending on its starting location:

$$428 \quad f_E = \int \mu_{ij}(\xi'', \xi + \xi'') \left( \sum_k \int Z_{3,ijk}(\xi + \xi'', \xi')w_{ik}(\xi')d\xi' \right) d\xi''. \quad (14)$$

429 This is similar in structure to (11), accounting for the influence on the focal  
430 individual's movement rate of a third individual of type  $k$  at displacement



431  $\xi''$ . The outer integral over  $\xi''$  allows for all possible starting locations for  
 432 the focal individual.

433 (F) The other individual (of type  $j$ ) in the pair also affects the movement  
 434 rate of the focal individual:

$$435 \quad f_F = \int Z_{2,ij}(\xi + \xi'') \mu_{ij}(\xi'', \xi + \xi'') w_{ij}(\xi + \xi'') d\xi''. \quad (15)$$

436 This is similar in structure to (13), but instead of the intrinsic movement  
 437 rate  $m_i$ , accounts for the contribution to the focal individual's movement  
 438 rate from the other individual (of type  $j$ ) in the pair. When the pair is  
 439 initially separated by vector  $\xi + \xi''$ , this contribution is  $w_{ij}(\xi + \xi'')$ . Again,  
 440 the integral over  $\xi''$  allows for all possible starting locations.

441 The key ecological information for movement bias is contained in  $\mu_{ij}(\xi'', \xi +$   
 442  $\xi'')$ , which is the probability density that the focal individual's movement  
 443 vector is  $\xi''$ , conditional on the presence of an individual of type  $j$  located at  
 444  $\xi + \xi''$  relative to the focal individual. This is the movement vector needed  
 445 to create the  $ij$  pair separated by  $\xi$  as required. As with the stochastic  
 446 model (Eq. (2)), this is composed of two independent parts:

$$447 \quad \mu_{ij}(\xi'', \xi + \xi'') = f_i(|\xi''|) g(\arg(\xi''), \eta_{ij}(\xi + \xi'')). \quad (16)$$

448 The first part  $f_i(|\xi''|)$  relates to the distance moved by an individual of  
 449 type  $i$ , which is independent of the neighbourhood and given by Eq. (3).  
 450 The second part  $g(\arg(\xi''), \eta_{ij}(\xi + \xi''))$  is the probability density of mov-  
 451 ing in direction  $\arg(\xi'')$ , which does depend on the neighbourhood. This  
 452 dependence is encapsulated in the expected bias vector  $\eta_{ij}(\xi + \xi'')$  for an  
 453 individual of type  $i$  separated from an individual of type  $j$  by a vector  
 454  $\xi + \xi''$ :

$$455 \quad \eta_{ij}(\xi + \xi'') = \beta_i \left( \sum_k \int \nabla v_{ik}(\xi') \frac{Z_{3,ijk}(\xi + \xi'', \xi')}{Z_{2,ij}(\xi + \xi'')} d\xi' + \nabla v_{ij}(\xi + \xi'') \right) \quad (17)$$

456 Here  $\nabla v_{ik}(\xi')$  is the gradient vector of the bias kernel  $v_{ik}(\xi')$ . Eq. (17)  
 457 integrates over the neighbourhood of the focal individual for neighbouring  
 458 individuals of type  $k$ , then sums over all types  $k$ , and adds the effect of  
 459 the partner individual of type  $j$  in the pair. The parameter  $\beta_i$  gives an  
 460 overall weight for the bias. The bias vector provides the parameters for a  
 461 circular probability distribution. To match the stochastic model, we use a  
 462 von Mises distribution with peak angle  $\arg(\eta_{ij})$  and concentration  $|\eta_{ij}|$ , to  
 463 obtain the probability density function of the angle  $\arg(\xi'')$ .

464 Summing expressions (10)–(15), gives the total rate of change of the pair  
 465 density  $Z_{2,ij}(\xi)$ :

$$466 \quad \frac{\partial}{\partial t} Z_{2,ij}(\xi, t) = f_A(\xi, t) + \dots + f_F(\xi, t) + \langle i, j, \xi \rightarrow j, i, -\xi \rangle, \quad (18)$$

467 where the matching symmetric terms for the partner individual in the  $ij$   
468 pair are given by the substitutions  $\langle i, j, \xi \rightarrow j, i, -\xi \rangle$  (Plank and Law, 2015).  
469 We give the function arguments in full to make clear the time dependence.  
470 This is a formal, exact description of how the movement rules at the level  
471 of the individual translate into a dynamical system of pair densities at the  
472 macroscopic level, after averaging over many realizations of the stochastic  
473 process, starting from the same statistical distribution.

### 474 3.3 Closure of the second-moment dynamics

475 The dynamical system is not yet closed, because it contains the third spatial  
476 moment, the density of triplets. To deal with this, a closure approxima-  
477 tion is needed to replace the third moment by a function of lower-order  
478 moments. Although not usually recognized, closures are ubiquitous in eco-  
479 logical theory: ignoring spatial structure completely implies a closure of the  
480 form  $Z_{2,ij}(\xi) = Z_{1,i}Z_{1,j}$ , giving a dynamical system for the first moment  
481 (average density), i.e. the law of mass action, or the so-called mean-field  
482 assumption. A formal theory of closures at second order is a matter for  
483 research (Raghib et al., 2011; Dieckmann and Law, 2000; Murrell et al.,  
484 2004). Here, we use the Kirkwood closure (Kirkwood, 1935):

$$485 \quad Z_{3,ijk}(\xi, \xi') = \frac{Z_{2,ij}(\xi)Z_{2,ik}(\xi')Z_{2,jk}(\xi' - \xi)}{Z_{1,i}Z_{1,j}Z_{1,k}} \quad (19)$$

486 as we have found the exact choice of closure makes little difference when  
487 the dynamics deal only with movement (i.e. without birth and death) (see  
488 for example Fig 6.3 in Binny (2016)).

### 489 3.4 Spatial-moment dynamics as an approximation 490 scheme

491 After closure, the dynamical system is no more than an approximation for  
492 the expected value of the second moment of the stochastic process, because  
493 it ignores spatial information carried by higher-order moments. How well  
494 does this approximation work? This is analogous to asking how well the  
495 mean-field assumption works as a description of population dynamics; the  
496 answer to that question is that the approximation is poor if neighbourhoods  
497 are important (Raghib et al., 2011). The second-order closure should be  
498 better because it does carry spatial information, but would still be expected  
499 to become poor as higher-order spatial structure becomes important.

500 Fig. 4 compares the spatial signal of the spatial-moment dynamics with  
501 that of the stochastic individual-based model from which the dynamical

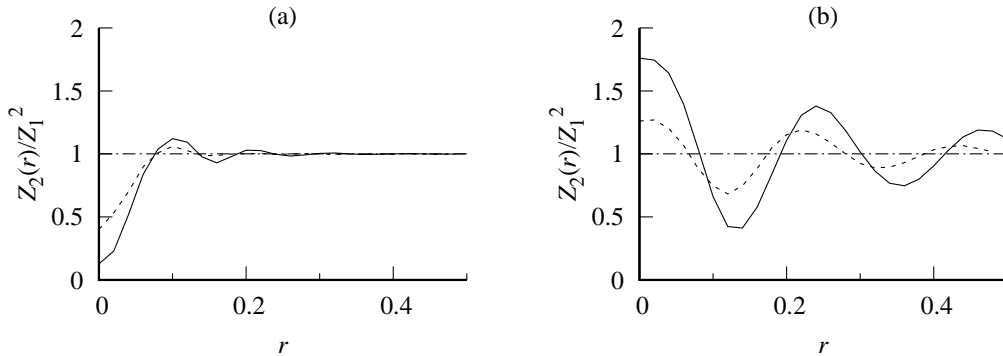


Figure 4: Solutions for the normalised pair density  $Z_2(r)/Z_1^2$  of the spatial-moment dynamics, Eqs (18)–(19), at time 10 (continuous lines), as a function of the distance  $r$  between the pair. These are approximations for the stochastic process of individual movement in Section 2, using parameter values that generated (a) territories in Fig. 2a, and (b) small clusters in Fig. 2b. For comparison, we also show the pair correlation functions (*sensu* Fig. 2g,h) averaged over 100 realizations of the stochastic process at time 10 (dashed lines). Initial conditions were spatial Poisson processes (dash-dot lines). Numerical integration was done by the Euler method, using Eq. (18)–(19), discretised as  $d\xi = 0.02$ ,  $dt = 0.05$ .

502 system (18), (19) was derived. For comparability with the stochastic re-  
503 sults, we assumed the movement rate to be independent of neighbourhood  
504 by setting  $w_{ij}(\cdot) = 0$  in Eqs (10)–(15), and leaving in place only an effect  
505 of neighbours on the direction of intrinsic movements. This means that  
506 the spatial-moment dynamics deal only with terms (10), (13) (geometries  
507 A and D in Fig. 3). We examined the dynamics for the bias kernels shown  
508 in Fig. 2a,b, as these generate structure at a small spatial scale. We would  
509 not expect to find a good approximation with the bias kernel in Fig. 2c,  
510 because spatial structure remains at large spatial scales. In other words,  
511 the pair correlation  $\rho(r)$  does not approach 1 as  $r$  increases in Fig. 2i.

512 Fig. 4 shows that the approximation scheme captures some basic signals  
513 of the stochastic, individual-based model. Fig.4 shows the characteristic  
514 regular structure arising from repulsive bias, manifested as a lack of pairs  
515 at short distance. Fig. 4 shows the distinct cluster formation as a result of  
516 short-range repulsion, medium-range attraction, and long-range repulsion.  
517 Although the quantitative match between the stochastic results and the  
518 spatial moments approximation is far from perfect, the key qualitative fea-  
519 tures of the emergent spatial structure are captured in the second moment.  
520 This illustrates two key points. First, it shows that the rules responsible for  
521 generating the spatial structure in the stochastic model are encapsulated  
522 by the dynamical system of spatial moments, despite the latter appear-  
523 ing to be completely different. Second it demonstrates that much of the  
524 information about spatial structure is carried just in the second spatial mo-

525 ment. In other words, there is some justification for closing the hierarchy  
526 at second order. The information shown in Fig. 4 would be lost completely  
527 in a mean-field model, which implicitly closes the system at the level of the  
528 first moment.

## 529 4 Discussion

530 This work draws on recent advances in spatial moment dynamics models  
531 of collective cell behaviour (Binny et al., 2016a; Surendran et al., 2018b)  
532 to address the issue of animal herding behaviour in ecology, and opens  
533 new research avenues in this setting. In particular, we have explored how  
534 using different forms of neighbourhood interaction kernels for directionally  
535 biased movement can give rise to formation of animal groups or herds.  
536 Individual-based models describing biased directional movement have been  
537 widely used in an ecological context (Codling and Hill, 2005; Benhamou,  
538 2006; Codling et al., 2007; Bode, 2011). However, this is the first time that a  
539 spatial moment dynamics model, capturing the outcomes of this directional  
540 bias at the macroscopic scale, has been used to describe animals living in  
541 groups. In doing this, we have shown the geometry of six flux terms that  
542 describe the exact relationship between the algorithmic individual-based  
543 model and the mathematical model (up to the second spatial moment).

544 Our results show that herd-like spatial structure can be generated solely  
545 from interactions among neighbouring individuals of the same species. In  
546 reality, this spatial structure can be strongly affected by interspecific in-  
547 teractions, such as the presence of predators. Future work will include  
548 explicitly applying the model framework developed here to systems with  
549 multiple interacting species. This has been done for cell–obstacle interac-  
550 tions (Surendran et al., 2018b) and chase–escape interactions (Surendran  
551 et al., 2018a), but these models use simple attractive or repulsive inter-  
552 actions, rather than the distance-dependent interactions that we employ  
553 here.

554 One advantage of spatial moment approximations over individual-based  
555 models is that the equations for the dynamics of spatial moments are de-  
556 terministic and only need to be solved once, rather than performing com-  
557 putationally intensive repeated simulations. They are also more tractable  
558 mathematically, permitting further analysis and exploration of parameter  
559 space. Computational power typically restricts simulation of individual-  
560 based models to systems with relatively low numbers of individuals, due to  
561 the requirements of tracking each individual’s movements and interactions  
562 with each of its neighbours over time. There are many such examples of

563 small-herd systems in ecology (see for example Table 1 in Reiczigel et al.  
564 (2008)). In contrast, the computational requirement for solving the spatial  
565 moments approximation is independent of population size. The methodol-  
566 ogy would lend itself to systems with much larger animal herds and offer  
567 insights that would otherwise require considerably greater computational  
568 resources to achieve through simulations alone.

569 Although the spatial-moment model shows the basic spatial structure,  
570 its fit to the stochastic model could clearly be improved. Attenuation of  
571 the spatial signal with increasing distance is rather slow in Fig 4b, which  
572 generates inaccuracies that can propagate to shorter distances. Also, at the  
573 shortest distances, the model overestimates the strength of spatial struc-  
574 ture; this may be because, after discretisation, spatial resolution becomes  
575 less good as  $r \rightarrow 0$ . Such issues could be dealt with by discretising over a  
576 larger space on a finer spatial grid, but this would have made computation  
577 unfeasible. In future work, a Fourier transform for the convolution inte-  
578 grals should be considered, as this could provide a major increase in speed  
579 of computation.

580 Previous models for animals living in herds have used the idea of zones  
581 of attraction and repulsion (Couzin et al., 2002; Bode, 2011). A zone of  
582 repulsion is also supported by data (Krause et al., 2002). Zones of repulsion  
583 and attraction have also been modelled in the cell behaviour literature,  
584 for example using the Lennard-Jones kernel (Jeon et al., 2010) and the  
585 Morse potential (Middleton et al., 2014; Matsiaka et al., 2017). Our model  
586 incorporates and builds on these ideas, including the possibility for multiple  
587 zones of attraction and repulsion with different spatial scales. Examples  
588 of the types of behaviour encapsulated by the bias kernels we have studied,  
589 and the resulting spatial structure, can be found in real animal populations.  
590 For example, Gerum et al. (2018) observed strong regular structure in king  
591 penguin (*Aptenodytes patagonicus*) colonies, caused by short-range nest  
592 site-protecting repulsive interactions between neighbours. Gajamannage  
593 et al. (2017) studied the formation of small clusters in cows (*Bos taurus*),  
594 generated by a balancing of costs to an individual of synchronisation (e.g.  
595 needing to concede to the timings of a large group, causing interrupted  
596 rest or grazing) with the benefits of reduced predation risk for larger, more  
597 defensible groups. Olson et al. (2009) observed the formation of a mega-  
598 herd in Mongolian gazelles (*Procapra gutturosa*), driven by habitat quality  
599 in a fragmented landscape.

600 Some animal behaviour models also have an orientation component to  
601 make individuals move in the same direction (Sumpter et al., 2008). This  
602 is more relevant for species where individuals in a group tend to be in  
603 continuous motion, such as shoaling fish or flocking birds. These situa-

604 tions require a velocity jump process (Codling et al., 2007, 2008), where  
605 reorientation events depend on the distance to and current orientation of  
606 other individuals in the neighbourhood (Agueh et al., 2011). In principle,  
607 the structure of such a population could be described by a second spatial  
608 moment in terms of the difference between the positions and orientations  
609 of two individuals in a pair, but this problem is currently untackled.

## 610 Acknowledgements

611 RNB’s PhD scholarship was funded by the Royal Society Te Apārangi  
612 Marsden fund (grant number 11-UOC-005). RNB and MJP were partly  
613 funded by Te Pūnaha Matatini. RL acknowledges funding from the Uni-  
614 versity of Canterbury Erskine Fellowship scheme. We thank Alex James  
615 for discussions on an earlier version of the model and D W Franks for  
616 discussions on factors affecting group size.

## 617 References

- 618 Adams, T. P., Holland, E. P., Law, R., Plank, M. J., and Raghil, M. (2013).  
619 On the growth of locally interacting plants: differential equations for the  
620 dynamics of spatial moments. *Ecology*, 94(12):2732–2743.
- 621 Agueh, M., Illner, R., and Richardson, A. (2011). Analysis and simulations  
622 of a refined flocking and swarming model of Cucker-Smale type. *Kinetic  
623 and Related Models*, 4(1):1–16.
- 624 Barraquand, F. and Murrell, D. J. (2013). Scaling up predator-prey dynam-  
625 ics using spatial moment equations. *Methods in Ecology and Evolution*,  
626 4:276–289.
- 627 Benhamou, S. (2006). Detecting an orientation component in animal paths  
628 when the preferred direction is individual-dependent. *Ecology*, 87(2):518–  
629 528.
- 630 Binny, R. N. (2016). *Spatial Moment Models for Collective Cell Behaviour*.  
631 PhD thesis, University of Canterbury, New Zealand.
- 632 Binny, R. N., Haridas, P., James, A., Law, R., Simpson, M. J., and Plank,  
633 M. J. (2016a). Spatial structure arising from neighbour-dependent bias  
634 in collective cell movement. *PeerJ*, 4:e1689.
- 635 Binny, R. N., James, A., and Plank, M. J. (2016b). Collective cell be-  
636 haviour with neighbour-dependent proliferation, death and directional  
637 bias. *Bulletin of Mathematical Biology*, 78:2277–2301.

- 638 Binny, R. N., Plank, M. J., and James, A. (2015). Spatial moment dy-  
639 namics for collective cell movement incorporating a neighbour-dependent  
640 directional bias. *Journal of the Royal Society Interface*, 12:20150228.
- 641 Blath, J., Etheridge, A., and Meredith, M. (2007). Coexistence in locally  
642 regulated competing populations and survival of branching annihilating  
643 random walk. *Annals of Applied Probability*, 17(5/6):1474–1507.
- 644 Bode, N. W. F. (2011). *Modelling collective motion in animals and the*  
645 *impact of underlying social networks*. PhD thesis, University of York,  
646 UK.
- 647 Bolker, B. and Pacala, S. W. (1997). Using moment equations to under-  
648 stand stochastically driven spatial pattern formation in ecological sys-  
649 tems. *Theoretical Population Biology*, 52:179–197.
- 650 Codling, E. and Hill, N. (2005). Sampling rate effects on measurements  
651 of correlated and biased random walks. *Journal of Theoretical Biology*,  
652 233(4):573–588.
- 653 Codling, E., Pitchford, J., and Simpson, S. (2007). Group navigation and  
654 the “many-wrongs principle” in models of animal movement. *Ecology*,  
655 88(7):1864–1870.
- 656 Codling, E. A., Plank, M. J., and Benhamou, S. (2008). Random walk  
657 models in biology. *Journal of the Royal society interface*, 5(25):813–834.
- 658 Couzin, I. D., Krause, J., James, R., Ruxton, G. D., and Franks, N. R.  
659 (2002). Collective memory and spatial sorting in animal groups. *Journal*  
660 *of Theoretical Biology*, 218(1):1 – 11.
- 661 Dieckmann, U. and Law, R. (2000). Relaxation projections and the method  
662 of moments. In Dieckmann, U., Law, R., and Metz, J. A. J., editors,  
663 *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*,  
664 chapter 21, pages 412–455. Cambridge University Press, Cambridge, UK.
- 665 DuVal, E. H. (2007). Adaptive advantages of cooperative courtship for sub-  
666 ordinate male lance-tailed manakins. *The American Naturalist*, 169:423–  
667 432.
- 668 Elgar, M. A. (1989). Predator vigilance and group size in mammals and  
669 birds: A critical review of the empirical evidence. *Biological Reviews*,  
670 64:13–33.
- 671 Gajamannage, K., Bollt, E. M., Porter, M. A., and Dawkins, M. S. (2017).  
672 Modeling the lowest-cost splitting of a herd of cows by optimizing a cost  
673 function. *Chaos: An Interdisciplinary Journal of Nonlinear Science*,  
674 27(6):063114.

- 675 Gerum, R., Richter, S., Fabry, B., Bohec, C. L., Bonadonna, F., Nesterova,  
676 A., and Zitterbart, D. P. (2018). Structural organisation and dynamics in  
677 king penguin colonies. *Journal of Physics D: Applied Physics*, 51 164004.
- 678 Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical  
679 reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361.
- 680 Griffin, R. H. and Nunn, C. L. (2012). Community structure and the spread  
681 of infectious disease in primate social networks. *Evolutionary Ecology*,  
682 26:779–800.
- 683 Hamilton, W. D. (1971). Geometry for the selfish herd. *Journal of Theo-*  
684 *retical Biology*, 31(2):295 – 311.
- 685 Herbert-Read, J. E., Perna, A., Mann, R. P., Schaerf, T. M., Sumpter,  
686 D. J. T., and Ward, A. J. W. (2011). Inferring the rules of interac-  
687 tion of shoaling fish. *Proceedings of the National Academy of Sciences*,  
688 108(46):18726–18731.
- 689 Illian, J., Penttinen, A., Stoyan, H., and D., S. (2008). *Statistical Analysis*  
690 *and Modelling of Spatial Point Patterns*. Wiley & Sons Chichester UK.
- 691 James, R., Bennett, P. J., and Krause, J. (2004). Geometry for mutualistic  
692 and selfish herds: the limited domain of danger. *Journal of Theoretical*  
693 *Biology*, 228:107–113.
- 694 Jeon, J., Quaranta, V., and Cummings, P. T. (2010). An off-lattice hybrid  
695 discrete-continuum model of tumor growth and invasion. *Biophysical*  
696 *Journal*, 98(1):37–47.
- 697 Kirkwood, J. G. (1935). Statistical mechanics of fluid mixtures. *The Jour-*  
698 *nal of Chemical Physics*, 3:300–313.
- 699 Krause, J., Ruxton, G. D., and Ruxton, G. D. (2002). *Living in groups*.  
700 Oxford University Press.
- 701 Law, R., Murrell, D. J., and Dieckmann, U. (2003). Population growth in  
702 space and time: spatial logistic equations. *Ecology*, 84(1):252–262.
- 703 Lukeman, R., Li, Y.-X., and Edelstein-Keshet, L. (2010). Inferring individ-  
704 ual rules from collective behavior. *Proceedings of the National Academy*  
705 *of Sciences of the United States of America*, 107:12576–80.
- 706 Maher, C. R. and Lott, D. F. (1995). Definitions of territoriality used in  
707 the study of variation in vertebrate spacing systems. *Animal Behaviour*,  
708 49:1581–1597.



- 709 Markham, A. C., Gesquiere, L. R., Alberts, S. C., and Altmann, J. (2015).  
710 Optimal group size in a highly social mammal. *Proceedings of the Na-*  
711 *tional Academy of Sciences*, 112:14882–14887.
- 712 Matsiaka, O. M., Penington, C. J., Baker, R. E., and Simpson, M. J.  
713 (2017). Continuum approximations for lattice-free multi-species models  
714 of collective cell migration. *Journal of Theoretical Biology*, 422:1–11.
- 715 Middleton, A. M., Fleck, C., and Grima, R. (2014). A continuum approxi-  
716 mation to an off-lattice individual-cell based model of cell migration and  
717 adhesion. *Journal of Theoretical Biology*, 359:220–232.
- 718 Murrell, D. J. (2005). Local spatial structure and predator-prey dynam-  
719 ics: counterintuitive effects of prey enrichment. *American Naturalist*,  
720 166:354367.
- 721 Murrell, D. J., Dieckmann, U., and Law, R. (2004). On moment closures for  
722 population dynamics in continuous space. *Journal of Theoretical Biology*,  
723 229:421–32.
- 724 Murrell, D. J. and Law, R. (2000). Beetles in fragmented woodlands: a  
725 formal framework for dynamics of movement in ecological landscapes.  
726 *Journal of Animal Ecology*, 69(3):471–483.
- 727 Murrell, D. J. and Law, R. (2003). Heteromyopia and the spatial coexis-  
728 tence of similar competitors. *Ecology Letter*, 6:48–59.
- 729 Olson, K. A., Mueller, T., Bolortsetseg, S., Leimgruber, P., Fagan, W. F.,  
730 and Fuller, T. K. (2009). A mega-herd of more than 200,000 Mongolian  
731 gazelles *Procapra gutturosa*: a consequence of habitat quality. *Oryx*,  
732 43(1):149–153.
- 733 Parrish, J. K. and Edelstein-Keshet, L. (1999). Complexity, pattern, and  
734 evolutionary trade-offs in animal aggregation. *Science*, 284(5411):99–101.
- 735 Plank, M. J. and Law, R. (2015). Spatial point processes and moment  
736 dynamics in the life sciences: a parsimonious derivation and some exten-  
737 sions. *Bulletin of Mathematical Biology*, 77:586–613.
- 738 Pulliam, H. R. (1973). On the advantages of flocking. *Journal of Theoretical*  
739 *Biology*, 38:419–422.
- 740 Raghil, M., Hill, N. A., and Dieckmann, U. (2011). A multiscale maxi-  
741 mum entropy moment closure for locally regulated space-time point pro-  
742 cess models of population dynamics. *Journal of Mathematical Biology*,  
743 62:605–53.

- 744 Reiczigel, J., Lang, Z., Rozsa, L., and Tóthmérész, B. (2008). Measures of  
745 sociality: two different views of group size. *Animal Behaviour*, 75:715–  
746 721.
- 747 Reluga, T. C. and Viscido, S. (2005). Simulated evolution of selfish herd  
748 behaviour. *Journal of Theoretical Biology*, 234:213225.
- 749 Sah, P., Leu, S. T., Cross, P. C., Hudson, P. J., and Bansal, S. (2017).  
750 Unraveling the disease consequences and mechanisms of modular struc-  
751 ture in animal social networks. *Proceedings of the National Academy of*  
752 *Sciences of the United States of America*, 114:4165–4170.
- 753 Sumpter, D., Buhl, J., Biro, D., and Couzin, I. (2008). Information transfer  
754 in moving animal groups. *Theory in biosciences*, 127(2):177–186.
- 755 Surendran, A., Plank, M. J., and Simpson, M. (2018a). Spatial struc-  
756 ture arising from chase-escape interactions with crowding. *bioRxiv*, page  
757 470799.
- 758 Surendran, A., Plank, M. J., and Simpson, M. J. (2018b). Spatial mo-  
759 ment description of birth–death–movement processes incorporating the  
760 effects of crowding and obstacles. *Bulletin of Mathematical Biology*,  
761 80(11):2828–2855.
- 762 Wood, A. J. and Ackland, G. J. (2007). Evolving the selfish herd: emer-  
763 gence of distinct aggregating strategies in an individual-based model.  
764 *Proceedings. Biological sciences*, 274:1637–42.