

Non-linear advection–diffusion equations approximate swarming but not schooling populations

Daniel Grünbaum^{a,*}, Karen Chan^a, Elizabeth Tobin^a, Michael T. Nishizaki^b

^a School of Oceanography, University of Washington, Box 35180, Seattle, WA 98195-7940, USA

^b Department of Biology, University of Washington, Seattle, WA 98195-1800, USA

ARTICLE INFO

Article history:

Received 1 January 2008

Received in revised form 2 June 2008

Accepted 4 June 2008

Available online 12 June 2008

Keywords:

Social behavior

Individual-based model

Partial differential equation

ABSTRACT

Advection–diffusion equations (ADEs) are concise and tractable mathematical descriptions of population distributions used widely to address spatial problems in applied and theoretical ecology. We assessed the potential of non-linear ADEs to approximate over very large time and space scales the spatial distributions resulting from social behaviors such as swarming and schooling, in which populations are subdivided into many groups of variable size, velocity and directional persistence. We developed a simple numerical scheme to estimate coefficients in non-linear ADEs from individual-based model (IBM) simulations. Alignment responses between neighbors within groups quantitatively and qualitatively affected how populations moved. Asocial and swarming populations, and schooling populations with weak alignment tendencies, were well approximated by non-linear ADEs. For these behaviors, numerical estimates such as ours could enhance realism and efficiency in ecosystem models of social organisms. Schooling populations with strong alignment were poorly approximated, because (in contradiction to assumptions underlying the ADE approach) effective diffusion and advection were not uniquely defined functions of local density. PDE forms other than ADEs are apparently required to approximate strongly aligning populations.

© 2008 Elsevier Inc. All rights reserved.

1. Introduction

An on-going development in mathematical biology has been the derivation of spatially explicit models that approximate population distribution changes across time and space. This development is driven by the central importance of spatial processes to many outstanding questions in basic and applied ecology. However, the models' utility depends firstly on whether they represent sufficient detail of movement and other biological characteristics to capture key mechanisms, and secondly on whether they are concise and efficient enough to be implemented in ecologically relevant situations. These often involve large population sizes operating over large time and space scales. The two performance metrics – detail and efficiency – are often at odds in model design and suggest fundamentally different modeling approaches. Details of behavior, physiological state, genetics and other characteristics are easily implemented in stochastic individual-based models (IBMs). However, IBMs tend not to be efficient for large populations: computational effort usually increases at least in proportion to the number of individuals and in some cases much faster. Large populations can often be described very efficiently using partial differential

equations (PDEs). However, PDEs do not explicitly represent many individual-level details.

These two modeling approaches can be put to best advantage if a direct and quantitative association can be made between a detailed IBM and a PDE that approximates it. In that case, the large time- and space-scale consequences of changes in biological details can be practically explored using a two-level approach, with a microscopic IBM model explicitly representing biological details and a macroscopic PDE efficiently approximating the IBM for large-scale calculations (see [32] for an informative general discussion of multi-scale coupling between microscopic and macroscopic models). This dual approach is greatly facilitated in some cases by a large mathematical literature on the derivation of PDEs for individual behaviors, usually in the form of advection–diffusion equations (ADEs) for biased random-walks, e.g., [2,27,26]. The library of individual-level movement algorithms for which there exists a ready ADE translation is large, but is concentrated in a small subset of biological behaviors: density-independent behaviors, in which individuals respond to their environment but not to each other, and in which responses to the environment are explicit statistical functions of space. Many other biologically important behaviors have been investigated in much less detail, notably behaviors in which responses to environments are mediated by internal state dynamics (e.g., physiological states such as hunger, sensory processes such as adaptation, and cognition) [8,13], and those driven

* Corresponding author.

E-mail address: grunbaum@ocean.washington.edu (D. Grünbaum).

by social responses and other direct interactions among individuals.

Social behaviors are usually modeled at the individual level with non-local responses, such as interaction zones. For example, IBMs of swarming behavior typically assume attraction zones, within which organisms move towards each other, and repulsion zones, within which organisms are repelled (see [24,26] for some early examples, [6] for some newer ones). Usually these responses are modulated by distance, the cumulative effect being longer-range attraction and shorter-range repulsion such that population density within groups does not increase without bound. Schooling, which differs from swarming by the relatively higher polarity (i.e., correlation among heading angles) of members within a social group, is usually modeled by the addition of alignment responses in which group members actively match velocities to their neighbors [1,19].

Most continuum models of social grouping are partial integro-differential equations, in which the integral terms reflect the non-local character of neighbor-neighbor responses within groups. Often these are in ADE form, but with advective velocity expressed in terms of convolutions over population density within a local neighborhood [20,21,31]. Most such models are not explicitly derived from individual-level behaviors, in part because estimating fluxes due to social responses requires assumptions about higher order statistics of individual positions within population distributions that may be difficult to justify or complex to implement [11,22]. However, the convolution functional form appears a plausible initial assumption, and provides useful theoretical insights into the dynamics of group formation and dispersal. The non-linear character of some social behaviors, in which populations move while tending to resist deviations above or below some ‘preferred’ density, is reminiscent of fluid flows. Several recent investigations have introduced potentially useful ideas from fluid mechanics to the investigation of social behaviors, notably decomposition of population fluxes into incompressible and potential motion components [30] and direct representation of pressure-like effects in auxiliary equations [4].

All of these approaches aim to explicitly represent the fine-scale distributions of individuals within and between each group. In practice, given the numerical demands of solving the relevant partial integro-differential equations, these approaches are suited to investigating the dynamics of single groups or interactions among small numbers of groups. They represent a computational savings over IBMs; however, many ecologically interesting applications involve many thousands of groups, spread over spatial domains that are very large compared to the extent of any single group. For these applications, the computational benefits of the IBM-to-PDE transition may not be enough to make implementation of social behavior models practical on ecological scales.

An alternative approach is provided by spatially implicit models that attempt to predict how size distributions of groups change over time (see the useful introduction in [24]). In this approach, assumptions are made about size-dependent rates of fusion and fission of groups, yielding a dynamical system for the number or frequency of groups of each feasible size [5,23,28]. There is debate about appropriate functional forms for fusion and fission rates, but typically fusion events are assumed to depend on encounter rates between groups, which are often approximately quadratic in frequency of groups. However, fission events often occur at an approximately constant probability per unit time, dependent on group size but not on the proximity of other groups. Hence, increasing population density results in a disproportionate increase in fusion rate relative to fission rate, causing a shift in the group size distribution towards larger groups.

An opportunity exists to combine the spatially explicit approaches used to represent population distributions with the spa-

tially implicit approaches that describe the subdivision of that population into groups within any given local area. The essence of this idea is to consider each group as a ‘meta-individual’, with movement characteristics that depend on its size and its local environment. Though interactions among individuals within groups are non-local and non-linear, these non-local responses operate only within a finite interaction distance. In many situations of ecological interest, groups spend most of their time separated from other groups by more than this distance, and during this time their movements are functionally independent. While they move independently, groups execute biased random walks. If biased random walk behavior occupies a sufficiently large fraction of their time, and fusion-fission events a sufficiently small fraction, the spatial distribution of groups of each size might be estimated as for other random walks with an ADE. So too could the distribution of the whole population, with appropriate weighting of size-dependent ADEs and accounting for fusion-fission dynamics.

If random walk statistics of groups vary with group size, as most simulation results suggest they do, then shifts in the group size distribution correspond to changes in the random walk characteristics of the population of groups. It appears a plausible hypothesis that for a given population density, individuals have a characteristic distribution across group sizes, each of which makes a characteristic contribution to population-level advection and diffusion, and that the cumulative advection and diffusion therefore varies in a characteristic way with population density. This reasoning suggests a non-linear advection-diffusion equation of the form

$$\frac{\partial P}{\partial t} = \nabla(D(P)\nabla P - \chi(P)\nabla S P) \quad (1)$$

where the coefficients $D(P)$ and $\chi(P)$ are functions of population density, and ∇S represents directional environmental cues. Flierl et al. [10] used a similar heuristic argument to suggest, based on size-dependent geometry and movement characteristics of social groups in an idealized model of schooling behavior, that the effective diffusion coefficient was a decreasing function of population size. Okubo et al. [25] proposed an analytical procedure for rigorously deriving density-dependent diffusion coefficients, but implementing this scheme is difficult.

Previous ecosystem-level models of social organisms such as tuna [15] and capelin [17] have assumed ADE forms, but without explicit functional dependence on population density. The overall goal of this paper is explore the potential of equations like (1) to endow these models with improved descriptions of social movements, without requiring a fundamental shift in numerical methods. A polynomial density-dependence is a convenient, though not unique, functional form in which to explore the potential of this approach. Therefore, our specific hypothesis is that IBMs of combined social and gradient-climbing behaviors can be approximated at very large space and time scales by (1) with coefficients that are polynomial functions of population density,

$$\begin{aligned} D(P) &= D_0 + D_1 P + D_2 P^2 + \dots + D_m P^m, \\ \chi(P) &= \chi_0 + \chi_1 P + \chi_2 P^2 + \dots + \chi_m P^m. \end{aligned} \quad (2)$$

Lacking a practical analytical alternative, we consider in this paper whether simple empirically driven numerical approaches – essentially curve fits of our assumed functional forms – can yield useful approximations to spatially explicit IBM models of complex social behavior. If so, immediate progress might be made addressing biological processes that are of practical importance but presently intractable, and insights might be gained into which analytical simplifications can be justified. If not, the failure of our quick and easy numerical approach may indicate that a more laborious and complex analytical approach would also likely be unsuccessful.

If (1) and (2) are adequate descriptors of complex IBM dynamics, a small number of simulation runs are sufficient to estimate the relevant coefficients. The objectives of this paper are: to carry out a set of IBM simulations with individuals exhibiting a range of social and gradient-climbing behaviors to serve as test cases (Section 2); to implement an efficient procedure to calculate best-fitting parameters of equations of the form (1),(2) (Section 3); to apply this procedure to the IBM results and assess how well the resulting PDEs approximate the IBM population distributions (Section 4); and, finally, to consider what the results imply for PDE approximation approaches for social behaviors (Section 5).

2. Individual-based simulations

The IBM used to generate tests cases in this paper combines a taxis-type gradient climbing behavior with attraction-repulsion and alignment responses to neighbors. [12] used a similar model to show that social interactions can enhance gradient-climbing in noisy environments. Individuals move at constant speed, here normalized to unity, within the $x-y$ plane. At regular small time intervals, δt , individuals increment their direction by a small angle change $\Delta\theta$, given by

$$\Delta\theta = \Delta\theta_{\text{tax}} + \Delta\theta_{\text{attr}} + \Delta\theta_{\text{rep}} + \Delta\theta_{\text{align}}. \quad (3)$$

Here, $\Delta\theta_{\text{tax}}$ is an angle increment due to taxis, which for the i th individual (with x -position x_i and heading angle θ_i) is

$$\Delta\theta_{\text{tax}} = \max(0, \Delta\theta_1 + g(t, x)\Delta\theta_2 \cos \theta_i)N_{i,t}. \quad (4)$$

$\Delta\theta_1$ is a heading-independent angle increment, and $\Delta\theta_2$ is the magnitude of response to a directional signal $g(t, x)$ that depends on heading angle, θ_i . This angle-dependence results in larger increments away from some directions, and hence a reduced probability of individuals moving in those directions. For positive values of $\Delta\theta_2$, this results in tactic movement in the positive x -direction when $g > 0$ and in the negative x -direction when $g < 0$. The directional signal may be, but does not have to be, the gradient of an attractant distribution, $S(t, x)$, in which case $g = \partial S / \partial x$. $N_{i,t}$ is chosen from a standard normal distribution at each time step for each individual.

$\Delta\theta_{\text{attr}}$ is the social attraction response. At each time step, each individual censuses the number of neighbors within radius r_{attr} to its right or left. If the total number of neighbors is below a target density μ_{attr} , then $\Delta\theta_{\text{attr}} = \pm\alpha$, the sign being chosen so that the increment is towards the side with more neighbors. Otherwise, $\Delta\theta_{\text{attr}} = 0$. The repulsion response is analogous: if the total number of neighbors within a radius r_{rep} exceeds μ_{rep} , individuals turn with increment $\Delta\theta_{\text{rep}} = \pm\beta$, the sign chosen so that the increment is away from the side with more neighbors. The alignment response is $\Delta\theta_{\text{align}} = \pm\gamma$, with the sign such that angle increment is towards matching the heading of a randomly chosen neighbor within a radius r_{align} . Over many steps, individuals with this alignment response tend to approach the mean heading of their neighbors.

Four parameter sets were used in the estimation procedures: an *asocial* behavior, in which there are no social interactions ($\alpha = \beta = \gamma = 0$); a *swarm* behavior, that included attraction and repulsion but not alignment ($\alpha = 4, \beta = 10, \gamma = 0$); a *weak* alignment behavior that included attraction and repulsion with a slight tendency to align with neighbors ($\alpha = 4, \beta = 10, \gamma = 1$); and a *strong* alignment behavior that included attraction and repulsion with a strong tendency to align with neighbors ($\alpha = 4, \beta = 10, \gamma = 3$). Note that for all the social behaviors, repulsion responses dominate both attraction and alignment, and attraction exceeds alignment. The interaction distances for all the social cases were $r_{\text{rep}} = 1$ and $r_{\text{attr}} = r_{\text{align}} = 1.5$. In all simulations, periodic boundary conditions were used in both x - and y -directions. Environmental variation, when present, was only in the

x -direction. Individuals were initially distributed with uniform random distributions in θ and y , and with probability density functions in x as described below.

3. Estimation of advection–diffusion coefficients

The coefficients in (1) and (2) can be efficiently estimated from IBM simulations by considering the spatially uniform case, $P(t, x) = \bar{P}$ and $g(t, x) = \bar{g}(t)$. In this case, the coefficients $D(\bar{P})$ and $\chi(\bar{P})$ are constant across the simulation domain. Within this uniform distribution, a subset of ‘marked’ individuals that at time 0 is in the interval $[x_1, x_2]$ subsequently has the distribution

$$\hat{P}(t, x) = \frac{\bar{P}}{2} \left[\text{erf} \left(\frac{x - x_1 - U(\bar{P})t}{\sqrt{4D(\bar{P})t}} \right) - \text{erf} \left(\frac{x - x_2 - U(\bar{P})t}{\sqrt{4D(\bar{P})t}} \right) \right] \quad (5)$$

where $U(\bar{P}) = \chi(\bar{P})\bar{g}(t)$. The expected number of individuals in the interval $[x_3, x_4]$ at time t is then

$$\hat{Q}(t) = \int_{x_3}^{x_4} \hat{P}(t, x) dx. \quad (6)$$

If individuals are assumed to move independently (i.e., if most marked individuals are in different groups from each other), then the probability of observing k individuals move from $[x_1, x_2]$ to $[x_3, x_4]$ in an IBM simulation over time t is the Poisson probability

$$\text{Prob}(k|\hat{Q}) = \frac{\hat{Q}^k e^{-\hat{Q}}}{k!}. \quad (7)$$

By subdividing the domain into vertical sectors with width ΔL and recording movements of individuals between sectors at time intervals Δt , each spatially uniform simulation provides many realizations of the random movements in which to evaluate the cumulative probability from (7) for assumed values of $D(\bar{P})$ and $\chi(\bar{P})$. Minimization of the negative log-likelihood of cumulative probability across a series of \bar{P} values provides a convenient means of estimating coefficients for $D(\bar{P})$ and $\chi(\bar{P})$, assuming the functional forms (2).

The IBM simulations used to estimate these coefficients employed 2^{16} individuals moving within a spatial domain of $L_x = 324$ in the x -direction. Domain sizes in the y -direction were $L_y = 2592, 1296, 648, 486, 324,$ and 216 to obtain a range of population densities, $\bar{P} = 0.0780, 0.1561, 0.3121, 0.4682, 0.6243,$ and 0.9364 . Individuals in tight hexagonal packing at the repulsion distance, $r_{\text{rep}} = 1$, have a population density of $P_{\text{max}} = 2/\sqrt{3} = 1.1547$; thus these simulations span most of the range between zero and maximum relevant population densities. Simulations were run with time steps $\delta t = 0.025$ for $t_{\text{max}} = 600$ time units, with the attractant increasing linearly in time as $\bar{g}(t) = g_0 t / t_{\text{max}}, g_0 = 0.125$. Parameter estimates based on sub-sampled data suggest these simulations were adequate to achieve convergence (Fig. 1).

Likelihood-maximizing parameters with $m = 2$ for the asocial, swarm, weak align and strong align behaviors are shown in Table 1, together with 95% likelihood-based confidence regions for these parameters. The corresponding non-linear advection and diffusion coefficients are plotted in Figs. 2 and 3. Comparisons of population fluxes predicted by (1) and (2) using best-fit coefficients to fluxes observed in the spatially uniform simulations (Fig. 4) suggests that the non-linear ADEs can reproduce the simulations to which they were fitted, for each of the asocial, swarm, weak align and strong align behaviors.

Key features of these results, with possible interpretations, are as follows: in the asocial case, the estimation procedure correctly determines the absence of significant density-dependence in the

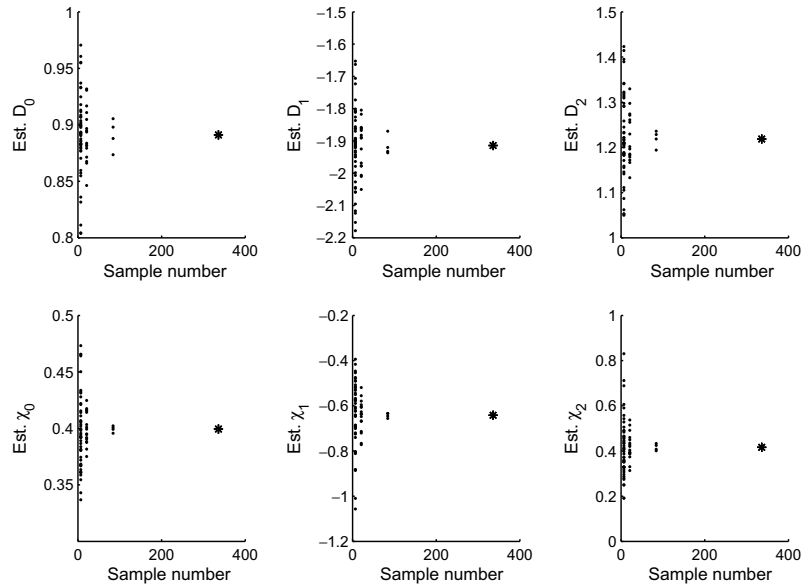


Fig. 1. Variability of estimated coefficients in (2) as a function of dataset size for the weakly aligning social behavior. The entire simulation dataset consisted of 336 samples, each of which represented a realization of the stochastic process in (5)–(7). Asterisks represent parameter estimates based on all 336 samples (Table 1). Points represent estimates from subsets of the samples (without replacement), i.e., 7 estimates from 48 samples each, 21 estimates from 16 samples each, and 4 estimates from 84 samples each. These plots suggest that estimates based on the complete 336 sample dataset have converged close to their true values. Convergence for the other behaviors were similar to that for the weakly aligning behavior, except for the strongly aligning behavior in which convergence was faster.

Table 1

Non-linear diffusion and taxis coefficients in Eqs. (1) and (2) with $m = 2$, estimated from IBM simulations of the four focal behaviors: asocial ($\alpha = \beta = \gamma = 0$); swarm ($\alpha = 4, \beta = 10, \gamma = 0$); weak alignment ($\alpha = 4, \beta = 10, \gamma = 1$) and strong alignment ($\alpha = 4, \beta = 10, \gamma = 3$)

Behavior	D_0	D_1	D_2
Asocial	1.6041 (1.5896,1.6187)	-0.0478 (-0.0758,-0.0195)	0.0238 (-0.012,0.0602)
Swarm	0.4957 (0.4926,0.499)	-0.9273 (-0.9322,-0.9222)	0.5932 (0.5872,0.5994)
Weak align	0.8911 (0.8874,0.8949)	-1.9143 (-1.9198,-1.9086)	1.2188 (1.2121,1.2257)
Strong align	3.2129 (3.2034,3.2226)	-6.6818 (-6.6944,-6.6689)	3.8585 (3.844,3.8735)
Strong adjusted	7.2262	-6.5716	3.7216
Behavior	χ_0	χ_1	χ_2
Asocial	0.5514 (0.5478,0.5549)	0.0017 (-0.00529,0.0087)	-0.0033 (-0.0122,0.00565)
Swarm	0.2203 (0.2182,0.2223)	-0.3700 (-0.3737,-0.3663)	0.2703 (0.2656,0.275)
Weak align	0.3994 (0.3972,0.4017)	-0.6417 (-0.6547,-0.6376)	0.4163 (0.4113,0.4213)
Strong align	0.5170 (0.5131,0.5208)	1.7363 (1.7293,1.7433)	-1.6664 (-1.6749,-1.6578)
Strong adjusted	0.5170	1.7363	-1.6664

Numbers in parentheses indicate the likelihood-based 95% confidence interval [18]. These parameters were obtained with $\Delta L = 81$ and $\Delta t = 40$, for $40 \leq t \leq 600$. Also shown are the ‘adjusted’ coefficients used for additional comparisons in Fig. 9. Adjusted coefficients are not maximum-likelihood estimates, so confidence regions for them are not reported. See text for additional details.

advection and diffusion coefficients. Swarming behavior, in which individuals aggregate into groups but do not actively align, leads to a dramatic decrease in effective diffusion and taxis compared to the asocial behavior. An interpretation is that an unpolarized group in some sense averages the random fluctuations of its members. It is then expected that individuals diffuse less if they constrain their movement to remain within the group. Increasing population density from near zero to approximately one-half the tight packing density P_{\max} decreases the effective diffusion rate. This may be because the group size distribution shifts from mostly small to mostly large groups, with larger groups exerting a stronger filter on random movements. It is less clear why the directional component is lowered by social behavior. One interpretation is that frequent heading adjustments to maintain appropriate degrees of crowding disrupt directional responses to the environment.

Consistent with this interpretation, addition of weak alignment responses significantly increases the effective diffusion and taxis coefficients (nearly doubling them, in this case). Alignment behaviors dampen heading adjustments due to over- and under-crowd-

ing, thereby increasing group velocity, directional persistence and sensitivity to directional cues. These changes are expected to increase diffusion and advection. However, D and χ are still decreasing functions of \bar{P} , probably reflecting the tendency of larger groups to have lower polarity and smaller velocities than smaller groups.

An interesting transition occurs when alignment increases from weak to strong. Stronger alignment results in dramatically larger diffusion and taxis coefficients, and the diffusion coefficient decreases (as before) with \bar{P} . However, the taxis coefficient for the strong alignment behavior increases dramatically with \bar{P} , over the population density range between zero and $\bar{P} \approx P_{\max}/2$. This rise is consistent with previous observations of social taxis [12], in which strong alignment tended to narrow angular distributions within polarized groups climbing noisy gradients. This narrowing tendency was more pronounced as groups increased from solitary individuals to ≈ 50 members. In the present simulations, shifts in the group size distribution across the corresponding range may account for increases in taxis. At population densities above $\approx P_{\max}/2$, χ decreases with \bar{P} . This may be due to repulsion

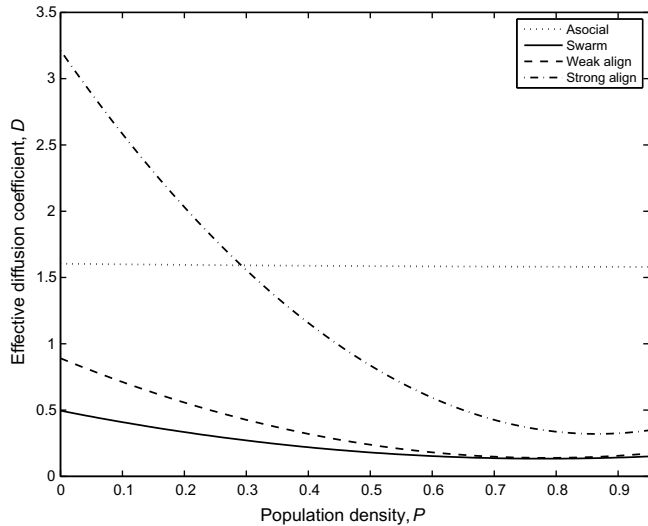


Fig. 2. Graphical comparison of the non-linear diffusion coefficient, $D(P)$, estimated from IBM simulations of asocial, swarm, weak align and strong align behaviors (see Table 1).

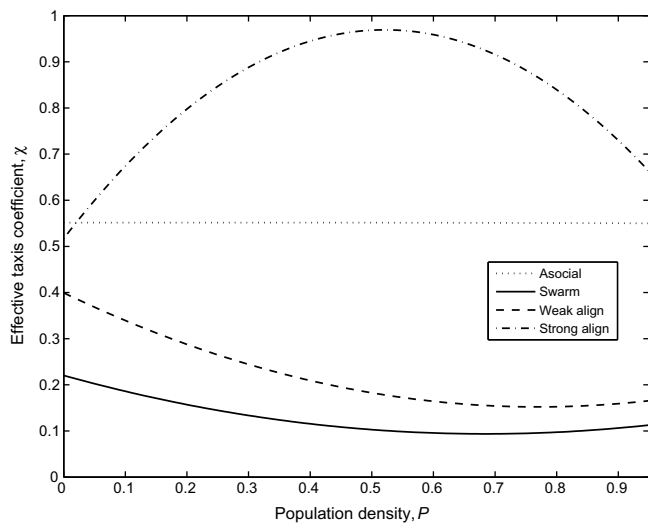


Fig. 3. Graphical comparison of the non-linear taxis coefficient, $\chi(P)$, estimated from IBM simulations of asocial, swarm, weak align and strong align behaviors (see Table 1).

responses during collisions among groups, which become increasingly frequent at high population densities.

4. IBM-PDE comparisons

To assess how well the fitted ADEs approximate the four focal IBM behaviors, matched IBM and ADE simulations were run in a spatially variable environment,

$$g(t, x) = g_1 \cos\left(\frac{2\pi x_i}{L_x}\right). \quad (8)$$

In each case, the population was initially uniformly distributed across the entire y -domain, and within an interval $[L_x/4 \leq x \leq 3L_x/4]$ that did not include the peak of the attractant distribution. This geometry was chosen because it allows comparisons among both initial transients and long-time distributions, and both sharp and smooth population density variations.

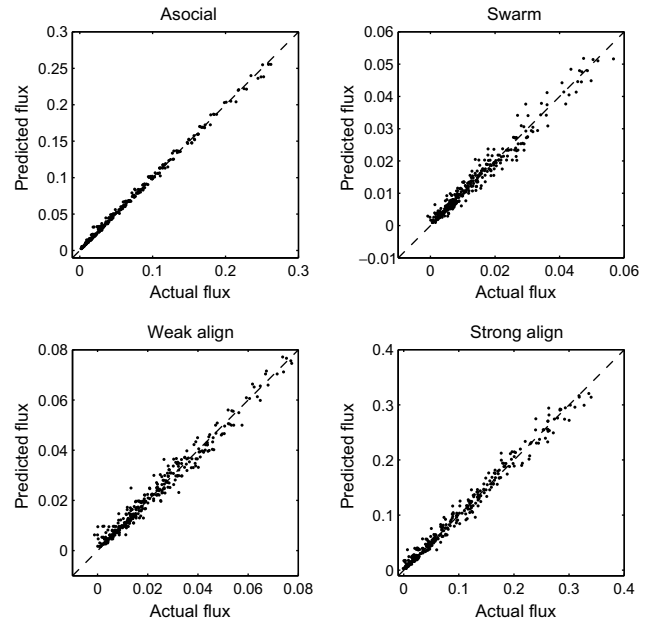


Fig. 4. Comparison of actual fluxes in spatially uniform IBM simulations with fluxes predicted from the non-linear diffusion and taxis parameters in Table 1. Shown are plots for asocial behavior ($R^2 = 0.9975$), swarm behavior ($R^2 = 0.9705$), weak alignment behavior ($R^2 = 0.9762$) and strong alignment behavior ($R^2 = 0.9870$). Note the differences in axes, which reflect overall differences in magnitudes of fluxes. In all plots, the dashed line represents a perfect match between actual and predicted fluxes.

Simulations were run with 2^{17} individuals using time steps $\delta t = 0.025$ for $t_{\max} = 1800$ time units, with spatial domains $L_x = 648, L_y = 486$, and with $g_1 = 0.125$. Spatial distributions from the two modeling approaches at three times are presented for asocial behavior (Fig. 5), swarming (Fig. 6), weak alignment (Fig. 7) and strong alignment (Fig. 8). Also shown in these figures are the instantaneous spatial distributions of D, χ, U and $\frac{\partial S}{\partial x^i}$ and the time series of an R^2 goodness-of-fit metric. These results are typical of many IBM/ADE comparisons run at both smaller and larger spatial scales.

In the first two of these comparisons – for the asocial and swarming behaviors – the ADEs derived from the estimation procedure are very good approximations to the corresponding IBMs. For the weak alignment behavior, the approximation seems to be useful. However, where the ADE model results are nearly indistinguishable from the IBM in the asocial and swarming behavior, in the weak aligning case the ADE gives small but systematic overestimates of high densities, and underestimates of low densities.

These errors become dramatic for the strong alignment behavior. In this case, the population in the ADE model quickly coalesces into a much tighter distribution than in the IBM, greatly overestimating the magnitude of the population density peak and underestimating its breadth. The reasons for the tight population concentration in the ADE model are indicated by the spatial distributions of D and χ : the peak corresponds to a significant local increase in χ , and a very large local decrease in D . From the IBM results, it appears that this peak (and hence presumably the D and/or χ distributions that cause it) are spurious.

This suggestion can be easily confirmed by arbitrarily increasing the diffusion coefficient (see Table 1), whereupon the adjusted ADE provides a very close approximation to the IBM distribution for the strong alignment behavior (Fig. 9). We have confirmed that the adjusted ADE is then a poorer fit to fluxes in the spatially uniform simulations of Section 3. This result suggests that an ADE can accurately approximate the strongly aligning spatially non-uniform

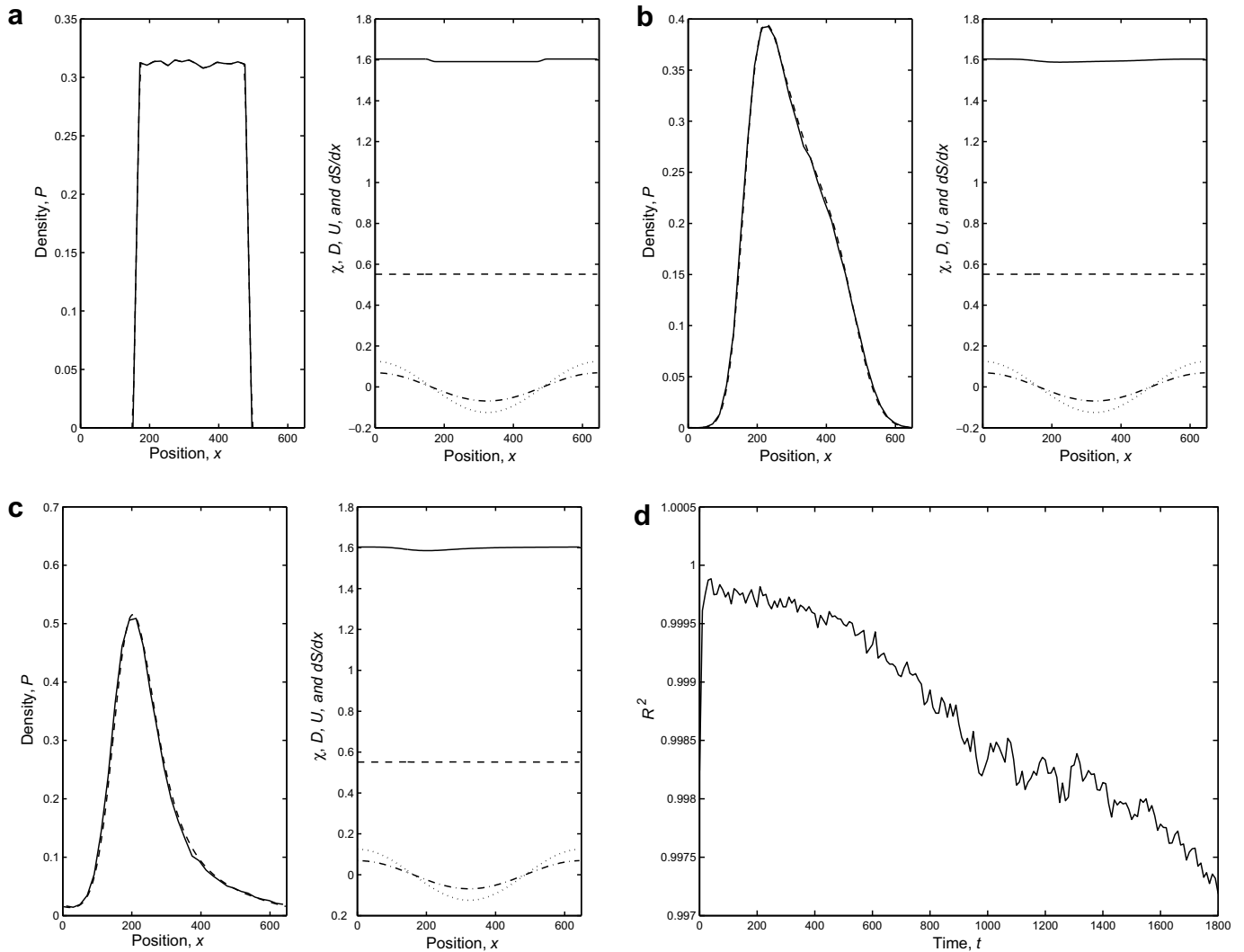


Fig. 5. Comparison of IBM vs. ADE spatial distributions for the asocial behavior, at $t = 0$ (a), $t = 60$ (b) and $t = 180$, (c). In each of these, the left subplot shows the spatial distribution of population density from the two models (IBM: solid; ADE: dashed). In the IBM, population density is estimated by censusing individuals in 32 bins equally spaced in the x -domain. The right subplot shows the spatial distribution of effective diffusion coefficient, $D(P)$ (solid line); effective taxis coefficient, $\chi(P)$ (dashed line); directional signal $g = \partial S / \partial x$ (dotted line); and advective velocity $U = \chi(P)g$ (dot-dashed line). Note changes in vertical scale. Also shown (d) is the time series of an R^2 goodness-of-fit statistic. See text for additional details.

IBM simulation. However, our estimation scheme did not identify that ADE, because an IBM of the same behavior on a different spatial domain is better approximated by a different ADE. That is, in the strongly aligning case there does not appear to be a one-to-one correspondence between individual behavior and approximating population-level ADE.

The suggestion is then that effective diffusion and/or taxis are not functions (only of) of population density for the strongly aligning behavior, though they appeared to be for the asocial, swarming and weakly aligning behaviors. We can assess this suggestion independently of our estimation scheme by calculating effective diffusion directly from individuals' mean squared displacements in both spatially uniform and non-uniform environments (Fig. 10). These results reinforce the idea that the strongly aligning behavior differs qualitatively from the other behaviors: The effective diffusion from the spatially non-uniform IBM falls on the curves from the spatially uniform IBMs for asocial, swarming and weakly aligning behaviors. However, effective diffusion from the spatially non-uniform IBM is much higher than the curve from the spatially uniform IBMs for the strongly aligning behavior.

5. Implications for PDE approximations of social behaviors

In this paper, we implemented a simple estimation procedure to approximate individual-based models (IBMs) of social and taxis behaviors with non-linear advection–diffusion equations (ADEs). We obtained approximating equations for four behaviors: asocial taxis, and taxis with three social behaviors (swarming, schooling with weak alignment, and schooling with strong alignment). The asocial behavior is density-independent. This behavior has linear dynamics, and in theory the constant coefficients for population movements can be derived using existing theoretical approaches. However, in practice, there is a mismatch in complexity and detail between the realistic behaviors typically employed in ecosystem-level IBMs and the more idealized behaviors for which analyses have been developed. For the social behaviors, which have inherently non-linear dynamics, some classes of population-level approximations have been investigated (usually with integral terms), but few theoretical methods exist for linking these explicitly to specific individual-level behaviors. For both asocial and social behaviors, there is an unmet need for population-level

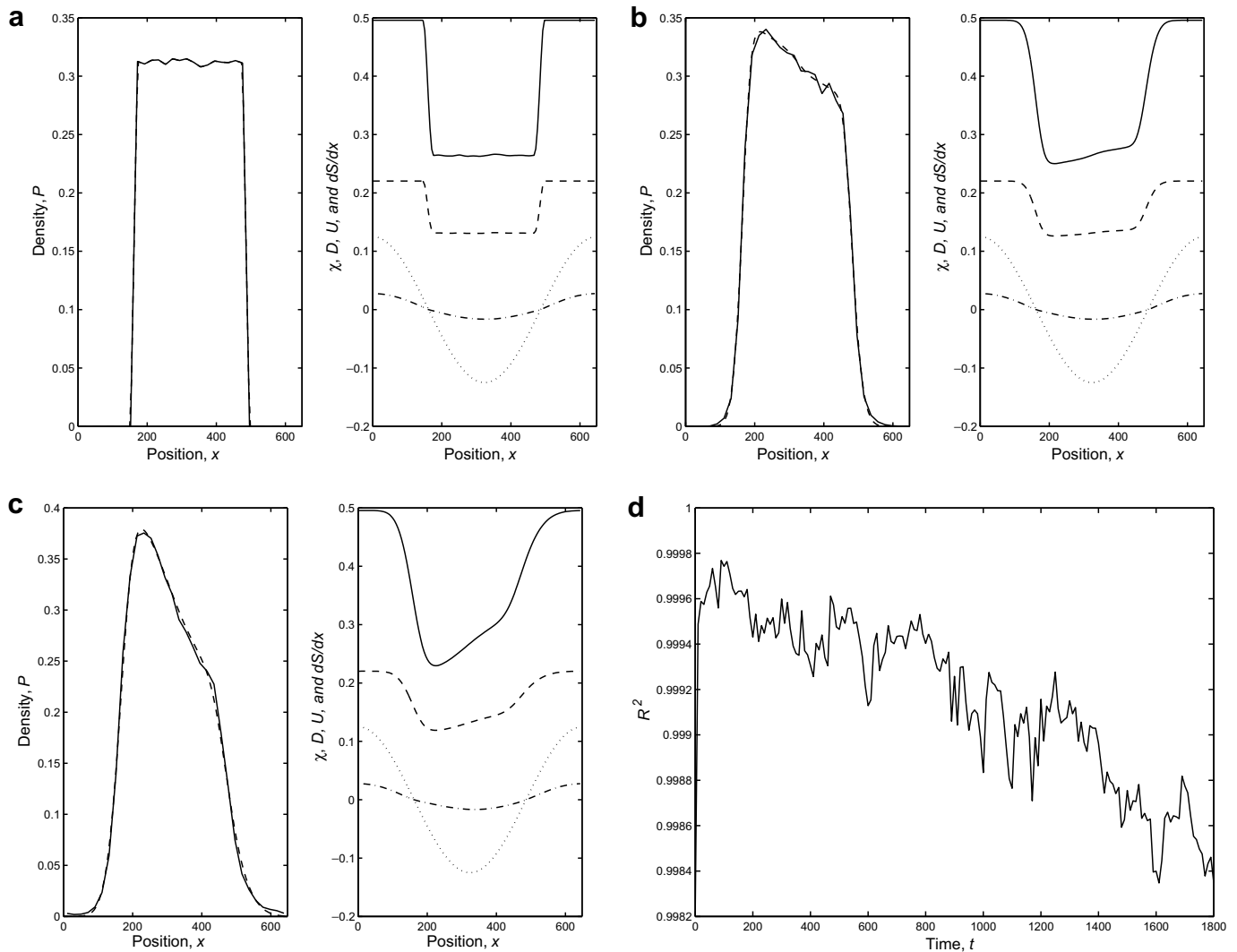


Fig. 6. Comparison of IBM vs. ADE spatial distributions for the swarming behavior, at $t = 0$ (a), $t = 60$ (b) and $t = 180$, (c). In each of these, the left subplot shows the spatial distribution of population density from the two models (IBM: solid; ADE: dashed). In the IBM, population density is estimated by censusing individuals in 32 bins equally spaced in the x -domain. The right subplot shows the spatial distribution of effective diffusion coefficient, $D(P)$ (solid line); effective taxis coefficient, $\chi(P)$ (dashed line); directional signal $g = \partial S / \partial x$ (dotted line); and advective velocity $U = \chi(P)g$ (dot-dashed line). Note changes in vertical scale. Also shown (d) is the time series of an R^2 goodness-of-fit statistic. See text for additional details.

approximations to address complex spatial dynamics relevant to current problems in basic and applied ecology.

By taking a numerical rather than an analytical approach, we sacrificed some measure of generality. However, we gained flexibility in applying our procedure not only to a wide variety of individual-based behavioral models, but also potentially directly to empirical observations of organism movements for which no adequate behavioral models exist.

The results of our parameter estimation are consistent with the general expectations from previous work that social behaviors have dramatic effects on effective population-level advection and diffusion. All the social behaviors resulted in substantial density-dependent non-linearities at the population level. Furthermore, details of social interactions, especially alignment, had profound effects on population-level fluxes. In real organisms, alignment (along with many other behavioral choices of individuals) likely varies facultatively among individuals over short time and space scales [14, e.g., see for an example in fish schools]. It is to be expected then that population fluxes for these organisms also change over commensurate scales.

We have two principal results. The first is that our estimation procedure yielded non-linear advection–diffusion equations that usefully predicted population distributions in IBMs of asocial, swarming and weakly aligning schooling behaviors. For two of those behaviors (asocial and swarming), population distributions were nearly indistinguishable between the IBM and ADE. For the weakly aligning case, systematic differences between the IBM and ADE results were evident. However, even in this case the approximation was good – almost certainly a significant improvement over constant-coefficient ADE descriptions currently used in most large-scale ecosystem models, which entirely neglect density-dependent effects.

For these behaviors, useful non-linear ADE approximations were obtained easily and with little computational effort. The computer time required to estimate parameters from short-term, small-scale IBMs and then to integrate the ADE solutions was lower than the direct simulation of the larger IBM. Furthermore, this is a one-time cost: the non-linear ADE can be applied to many different population distributions and environments, without re-estimating parameters, as long as the underlying movement behaviors are consistent. For

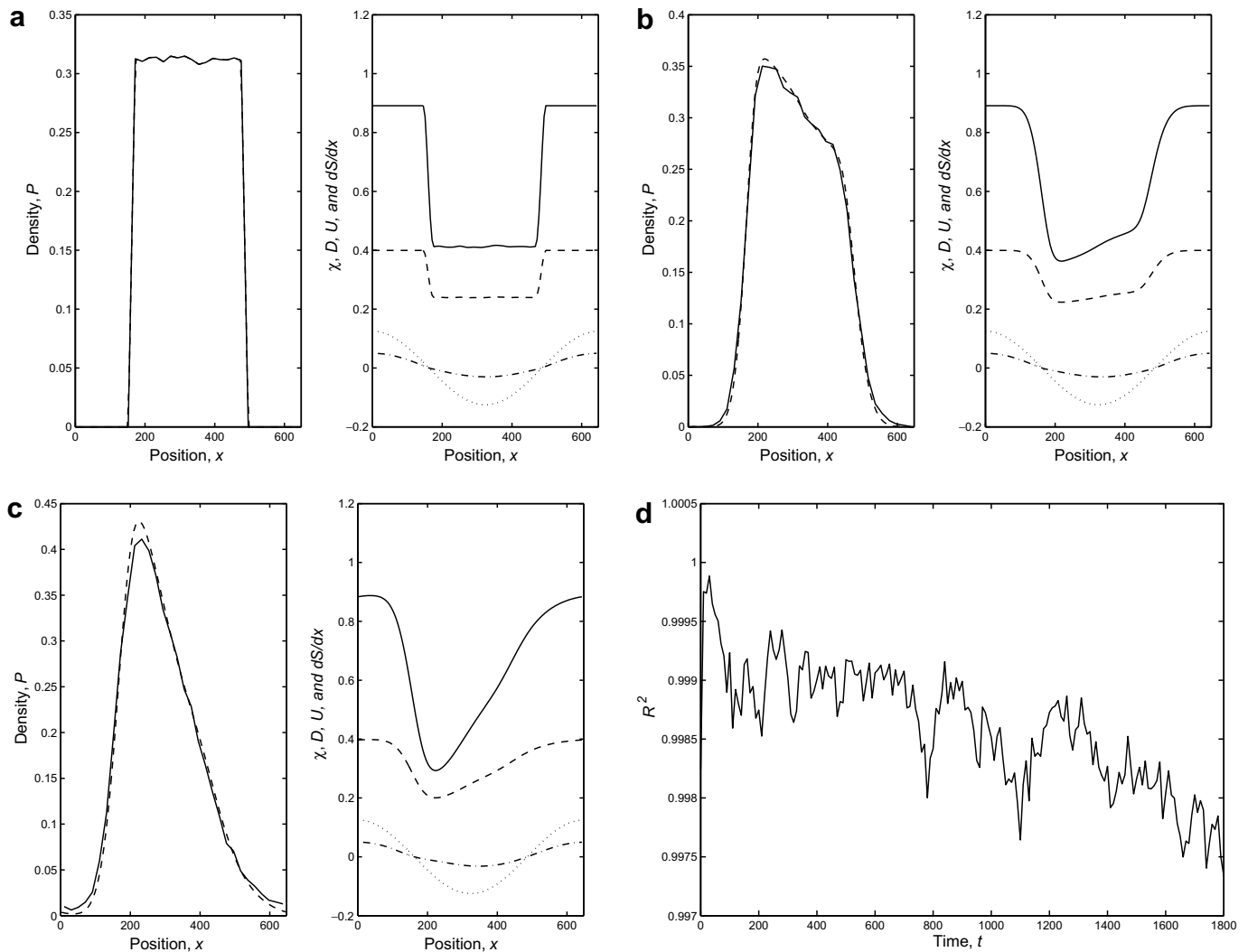


Fig. 7. Comparison of IBM vs. ADE spatial distributions for the weakly aligning schooling behavior, at $t = 0$ (a), $t = 60$ (b) and $t = 180$, (c). In each of these, the left subplot shows the spatial distribution of population density from the two models (IBM: solid; ADE: dashed). In the IBM, population density is estimated by censusing individuals in 32 bins equally spaced in the x -domain. The right subplot shows the spatial distribution of effective diffusion coefficient, $D(P)$ (solid line); effective taxis coefficient, $\chi(P)$ (dashed line); directional signal $g = \partial S / \partial x$ (dotted line); and advective velocity $U = \chi(P)g$ (dot-dashed line). Note changes in vertical scale. Also shown (d) is the time series of an R^2 goodness-of-fit statistic. See text for additional details.

models addressing real-world ecological problems, which are much more complex and costly, we expect the computational gains to be substantial. From the reduced computational requirements, and the analytical advantages of ADEs (however obtained), it appears beneficial to employ simple numerical estimation schemes like ours where analytical derivations of ADE approximations to IBM behaviors are unavailable or impossible. Furthermore, for behaviors similar to those in which the estimation procedure produced an accurate approximation, our results suggest that analytical derivations would be useful if they were made available.

Our second principal result is that our attempts to approximate population distributions that school with strong alignment with non-linear ADEs were not successful. For this behavior, we simulated IBM population distributions in two environments, one spatially uniform and the other spatially non-uniform. Each was well approximated by a non-linear ADE. However, we found no single ADE that adequately described population distributions in both environments.

Whether poor predictive skill of the approximating ADE for the strongly aligning IBM model has broader pessimistic implications for using ADEs to approximate strongly aligning social behaviors

depends on whether the estimation procedure failed, or whether these populations are in general not well described by any PDE that has an ADE functional form.

A limitation of ADEs in general is that they describe population movements accurately only in the so-called diffusion limit. Essentially this means that environmental and population density gradients must be small at the correlation length scales of individual movements. This condition is not satisfied at the leading edge of a population spreading from an initially compact distribution. In this case, the leading edge can advance at no more than the maximum movement speed of individuals, but the diffusion approximation predicts that the tail of the population distribution instantaneously spans the entire spatial domain (that is, the ADE is parabolic). If at the appropriate diffusion rate individuals in the IBM spread more slowly than expected, our estimation procedure may compensate by lowering its estimate of D . Consistent with this, our procedure underestimated D . However, the magnitude of this error is dependent on the time and space intervals ΔL and ΔT . The ADE coefficients estimated by our scheme were consistent across large variations in ΔL and ΔT . Hence, though we cannot exclude the possibility that our estimates had small

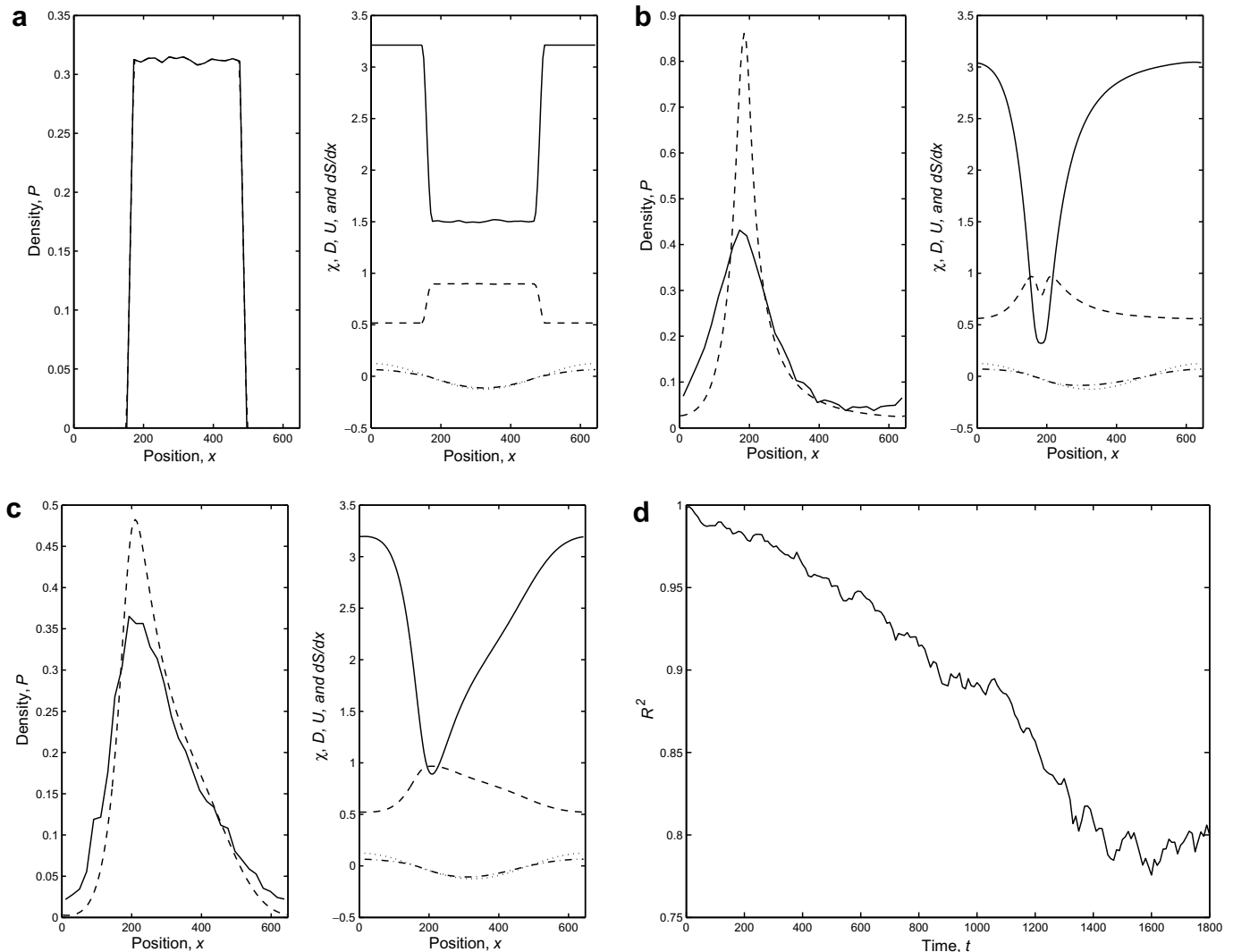


Fig. 8. Comparison of IBM vs. ADE spatial distributions for the strongly aligning schooling behavior, at $t = 0$ (a), $t = 60$ (b) and $t = 180$, (c). In each of these, the left subplot shows the spatial distribution of population density from the two models (IBM: solid; ADE: dashed). In the IBM, population density is estimated by censusing individuals in 32 bins equally spaced in the x -domain. The right subplot shows the spatial distribution of effective diffusion coefficient, $D(P)$ (solid line); effective taxis coefficient, $\chi(P)$ (dashed line); directional signal $g = \partial S / \partial x$ (dotted line); and advective velocity $U = \chi(P)g$ (dot-dashed line). Note changes in vertical scale. Also shown (d) is the time series of an R^2 goodness-of-fit statistic. See text for additional details.

inaccuracies, it is unlikely that this type of error underlies the large discrepancies between the strongly aligning IBM and the best-fit ADE.

We believe the alternative, more pessimistic explanation – that strongly aligning populations do not obey any sort of ADE dynamics of the form (1) – is more central to the poor accuracy of the approximation found by our estimation procedure. According to this explanation, the effective diffusion and taxis coefficients are not functions only of local population density (as they must be in (1)). Instead, strongly social populations with the same density can have different effective diffusion and taxis coefficients depending on non-local variations in population density and environment.

To assess whether different effective diffusion can occur at similar population densities, we estimated diffusion from individual trajectories under comparable conditions in the two environmental configurations we used in our simulations (Fig. 10). Mean squared displacements of these trajectories gave inconsistent estimates of effective diffusion, with a higher effective diffusion in the spatially non-uniform than in the spatially uniform case. This is inconsistent with effective diffusivity being strictly a function of

population density. Thus a key assumption underlying the functional form (1) appears to be violated.

The simplest explanation for the differences we observed in effective diffusion is that, if individuals travel at constant speeds (as they did in our simulations), diffusion and advection velocity are not independent (e.g., [9]). For example, high population-level advection speed can occur only if the heading angle distribution is very tight. This necessarily implies a low diffusion. Hence diffusion and advection co-vary as a function of the heading angle distribution. However, this covariation is quantitatively small until advection velocities are quite close to individuals' travel speed. We attempted to account for this co-variation by modifying (1). Specifically, we assumed that the heading angles of individuals at each x -position followed Wrapped Cauchy distributions. We used local advection velocities to estimate the local-scale parameters for these distributions, from which we calculated the corresponding decrement in effective diffusion. Neither the best-fit coefficients nor the accuracy of the approximating ADE were significantly changed by this modification. This is probably because, in our simulations, advection velocity rarely approaches individuals'

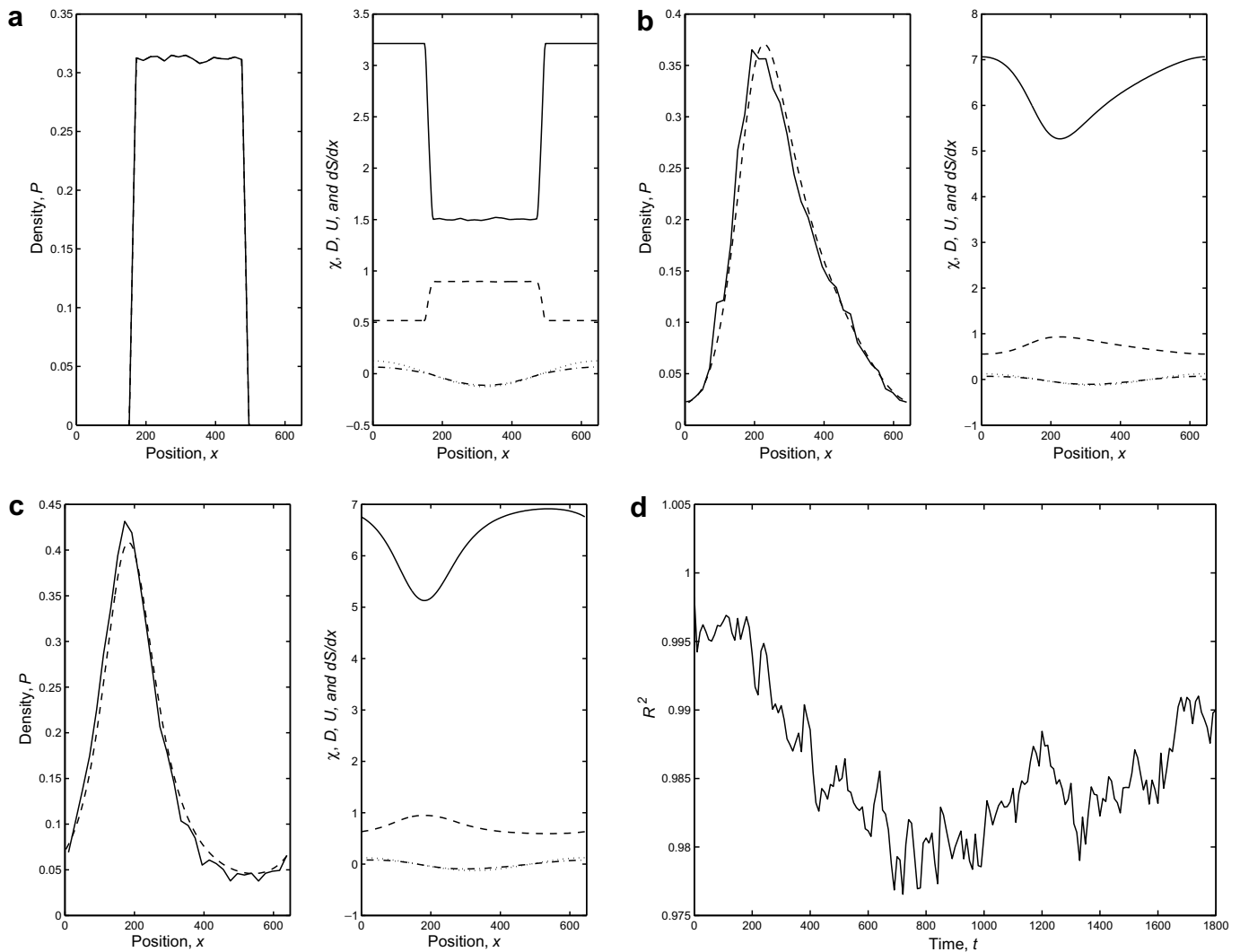


Fig. 9. This IBM vs. ADE comparison is the same as that in Fig. 8, except that the coefficients have been ‘adjusted’ by arbitrarily increasing the parameter D_0 (see Table 1). This adjustment is not in any way optimized. Nonetheless, this comparison shows that an ADE approximation is possible that predicts spatially non-uniform population distributions in the strongly aligning schooling IBM much better than the parameters estimated from the spatially uniform simulations.

travel speed (e.g., Figs. 5–9). Thus, we believe this explanation is not sufficient to explain the non-ADE dynamics we observed in strongly aligning populations.

A fuller explanation probably centers around more complex differences in the underlying distributions of group size, polarity and heading angles in the two spatial configurations. Strong alignment has effects not only within groups; it also tends to produce alignment among groups at larger spatial scales. This is because groups that differ greatly in heading are likely to encounter each other frequently, and subsequently both assume intermediate heading angles. Thus groups with ‘minority’ heading angles tend to reorient more than groups with ‘majority’ heading angles, reinforcing the consensus direction. In our spatially uniform simulations, there was no extrinsic mechanism to limit the length scale of this larger-scale alignment. The result was a high degree of polarity across the entire population, reflected in small diffusion coefficients, and high fluxes, reflected in large taxis coefficients.

In contrast, fluxes in the spatially variable case (Fig. 8) approached zero as the population distribution approached equilibrium. Zero flux is incompatible with population-wide polarity. If

strong alignment promotes polarity at greater than small length scales, and zero flux precludes polarity at large length scales, the likely result would seem to be polarity at intermediate scales. The proposed explanation is then that, in the presence of strong alignment, coherent meso-scale polarized structures arise in the zero-flux case but not the constant flux cases. These meso-scale structures significantly increase effective diffusion (as in Fig. 10), in ways that cannot be predicted from local population density and environment alone.

While our initial, simple ADE form appears inadequate to approximate population-level fluxes arising from social behaviors with strong alignment, it seems likely that a similar numerical approach using more suitable types of PDEs might prove successful. A possible experimental and theoretical analog for this has already been described in recent literature as ‘low Reynolds number turbulence’ arising in some dense populations of bacteria as a result of neighbor–neighbor contact and hydrodynamic forcing [3,7,16,29]. PDE forms abstracted from this and other fluid mechanics-inspired treatments of multi-scale spatial processes seem to be a good starting point from which to begin future approximation efforts.

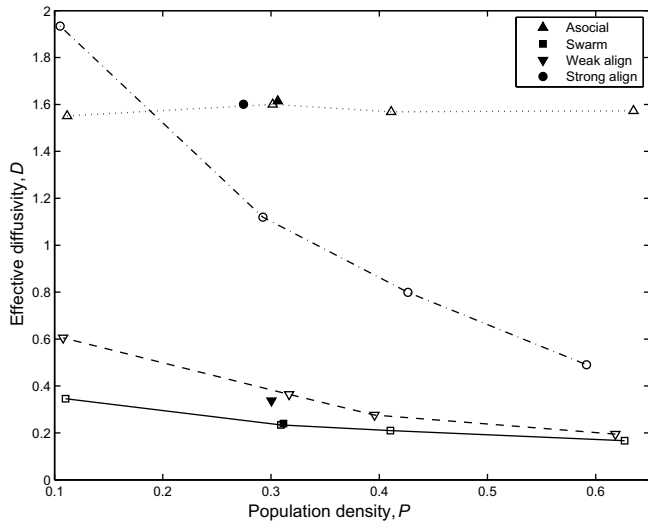


Fig. 10. Comparisons of the non-linear diffusion coefficient, $D(P)$, estimated directly from mean squared deviations of movements in IBM simulations of asocial, swarm, weak align and strong align behaviors. Solid symbols represent estimates from the spatially non-uniform simulations (Figs. 5–8) from Section 4. Open symbols represent estimates from the spatially uniform simulations in Section 3. Effective diffusion is consistent between the uniform and non-uniform simulations for asocial, swarming and weakly aligning social behaviors. However, for the strongly aligning social behavior, the effective diffusion is much higher in the spatially non-uniform simulation than in the spatially uniform simulation. All estimates were for individual movements over the time interval $40 \leq t \leq 60$, for individuals within the spatial interval $310 \leq x \leq 330$ at $t = 40$. In both spatial and non-spatial simulations, this time and space interval has comparable, approximately constant directional signal, $|g| \approx 0.125$.

Acknowledgments

The authors gratefully acknowledge support from the National Science Foundation (award EF-0434340 to D.G.). The manuscript was greatly improved by discussions with S.A. Levin, J.M. Moehlis, and Y.G. Kevrekidis; by insightful suggestions from an anonymous reviewer; and by stimulating lectures and discussions at the 2007 BIOCAMP conference. KC gratefully acknowledges support from the Sir Edward Youde Memorial Fund. E.T. and M.T.N. also acknowledge support from Washington State Sea grant (NA040AR170032 to D.G.).

References

- [1] M. Adiou, J.P. Treuil, O. Arino, Alignment in a fish school: a mixed Lagrangian–Eulerian approach, *Ecol. Modell.* 167 (1–2) (2003) 19.
- [2] W. Alt, Biased random walk models for chemotaxis and related diffusion approximations, *J. Math. Biol.* 9 (1980) 147.
- [3] I.S. Aranson, A. Sokolov, J.O. Kessler, R.E. Goldstein, Model for dynamical coherence in thin films of self-propelled microorganisms, *Phys. Rev. E* 75 (4) (2007) 040901.
- [4] P. Babak, K.G. Magnusson, S. Sigurdsson, Dynamics of group formation in collective motion of organisms, *Math. Med. Biol. – J. IMA* 21 (4) (2004) 269.

- [5] E. Bonabeau, L. Dagorn, P. Freon, Scaling in animal group-size distributions, *Proc. Natl. Acad. Sci. USA* 96 (8) (1999) 4472.
- [6] Y.L. Chuang, M.R. D’Orsogna, D. Marthaler, A.L. Bertozzi, L.S. Chayes, State transitions and the continuum limit for a 2d interacting, self-propelled particle system, *Phys. D – Nonlinear Phenomena* 232 (1) (2007) 33.
- [7] L.H. Cisneros, R. Cortez, C. Dombrowski, R.E. Goldstein, J.O. Kessler, Fluid dynamics of self-propelled microorganisms, from individuals to concentrated populations, *Exp. Fluids* 43 (5) (2007) 737.
- [8] R.B. Dickinson, R.T. Tranquillo, Transport equations and indices for random and biased cell migration based on single cell properties, *SIAM J. Appl. Math.* 55 (5) (1995) 1419.
- [9] B. Faugeras, O. Maury, Modeling fish population movements: from an individual-based representation to an advection–diffusion equation, *J. Theor. Biol.* 247 (4) (2007) 837.
- [10] G. Flierl, D. Grünbaum, S. Levin, D. Olson, From individuals to aggregations: the interplay between behavior and physics, *J. Theor. Biol.* 196 (4) (1999) 397.
- [11] D. Grünbaum, Translating stochastic density-dependent individual behavior with sensory constraints to an Eulerian model of animal swarming, *J. Math. Biol.* 33 (2) (1994) 139.
- [12] D. Grünbaum, Schooling as a strategy for taxis in a noisy environment, *Evol. Ecol.* 12 (5) (1998) 503.
- [13] D. Grünbaum, Advection–diffusion equations for internal state-mediated random walks, *SIAM J. Appl. Math.* 61 (1) (2000) 43.
- [14] D. Grünbaum, S. Viscido, J.K. Parrish, Extracting interactive control algorithms from group dynamics of schooling fish, in: *Lecture Notes in Control and Information Sciences*, vol. 309, 2005, pp. 103–117.
- [15] R. Humston, J.S. Ault, M. Lutcavage, D.B. Olson, Schooling and migration of large pelagic fishes relative to environmental cues, *Fisheries Oceanogr.* 9 (2) (2000) 136.
- [16] J. Lega, T. Passot, Hydrodynamics of bacterial colonies, *Nonlinearity* 20 (1) (2007) C1.
- [17] K.G. Magnusson, S.T. Sigurdsson, E.H. Deredsdottir, A simulation model for capelin migrations in the North Atlantic, *Nonlinear Anal.–Real World Appl.* 6 (4) (2005) 747.
- [18] W.Q. Meeker, L.A. Escobar, Teaching about approximate confidence-regions based on maximum-likelihood estimation, *Am. Stat.* 49 (1) (1995) 48.
- [19] V. Mirabet, P. Auger, C. Lett, Spatial structures in simulations of animal grouping, *Ecol. Modell.* 201 (3–4) (2007) 468.
- [20] A. Mogilner, L. Edelstein-Keshet, A non-local model for a swarm, *J. Math. Biol.* 38 (6) (1999) 534.
- [21] A. Mogilner, L. Edelstein-Keshet, L. Bent, A. Spiros, Mutual interactions, potentials, and individual distance in a social aggregation, *J. Math. Biol.* 47 (4) (2003) 353.
- [22] D. Morale, V. Capasso, K. Oelschläger, An interacting particle system modelling aggregation behavior: from individuals to populations, *J. Math. Biol.* 50 (1) (2005) 49.
- [23] H.S. Niwa, Power-law versus exponential distributions of animal group sizes, *J. Theor. Biol.* 224 (4) (2003) 451.
- [24] A. Okubo, Dynamical aspects of animal grouping: swarms, schools, flocks and herd, *Adv. Biophys.* 22 (1986) 1.
- [25] A. Okubo, D. Grünbaum, L. Edelstein-Keshet, The dynamics of animal grouping, in: A. Okubo, S.A. Levin (Eds.), *Diffusion and Ecological Problems: Mathematical Models*, Springer-Verlag, New York, 2001 (chapter 7).
- [26] A. Okubo, S.A. Levin, *Diffusion and Ecological Problems: Mathematical Models*, Lecture Notes in Biomathematics, Springer-Verlag, New York, 2001.
- [27] H.G. Othmer, S.R. Dunbar, W. Alt, Models of dispersal in biological systems, *J. Math. Biol.* 26 (1988) 263.
- [28] O. Pays, S. Benhamou, R. Helder, J.-F. Gerard, The dynamics of group formation in large mammalian herbivores: an analysis in the European roe deer, *Animal Behav.* 74 (2007) 1429.
- [29] A. Sokolov, I.S. Aranson, J.O. Kessler, R.E. Goldstein, Concentration dependence of the collective dynamics of swimming bacteria, *Phys. Rev. Lett.* 98 (15) (2007) 158102.
- [30] C.M. Topaz, A.L. Bertozzi, Swarming patterns in a two-dimensional kinematic model for biological groups, *SIAM J. Appl. Math.* 65 (1) (2004) 152.
- [31] C.M. Topaz, A.L. Bertozzi, M.A. Lewis, A nonlocal continuum model for biological aggregation, *Bull. Math. Biol.* 68 (7) (2006) 1601.
- [32] E. Weinan, B. Engquist, X. Li, W. Ren, E. Vanden-Eijnden, Heterogeneous multiscale methods: a review, *Commun. Comput. Phys.* 2 (3) (2007) 367.