Coarse-grained analysis of stochasticity-induced switching between collective motion states

Allison Kolpas*, Jeff Moehlis[†‡], and Ioannis G. Kevrekidis[§]

Departments of *Mathematics and [†]Mechanical Engineering, University of California, Santa Barbara, CA 93106; and [§]Department of Chemical Engineering and Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544

Edited by Simon A. Levin, Princeton University, Princeton, NJ, and approved February 15, 2007 (received for review September 20, 2006)

A single animal group can display different types of collective motion at different times. For a one-dimensional individual-based model of self-organizing group formation, we show that repeated switching between distinct ordered collective states can occur entirely because of stochastic effects. We introduce a framework for the coarse-grained, computer-assisted analysis of such stochasticity-induced switching in animal groups. This involves the characterization of the behavior of the system with a single dynamically meaningful "coarse observable" whose dynamics are described by an effective Fokker–Planck equation. A "lifting" procedure is presented, which enables efficient estimation of the necessary macroscopic quantities for this description through short bursts of appropriately initialized computations. This leads to the construction of an effective potential, which is used to locate metastable collective states, and their parametric dependence, as well as estimate mean switching times.

coarse-graining | equation-free | individual-based model | self-organization | schooling

Fish, birds, and honey bees, as well as many other animal groups, display collective types of motion such as schooling, flocking, and swarming (1, 2). A single animal group can display different types of collective motion at different times, with <1 day of residence time in each state (3). Although such transitions could be due to changing behavioral rules or environmental factors, they also can occur entirely due to stochastic effects, as will be demonstrated for the model considered in this paper.

One class of biologically motivated, individual-based models for group formation, frequently used for schooling fish, abstracts animal behavior by placing zones around individuals in which they respond to others through repulsion, alignment, and/or attraction (4–11). In the three-dimensional model of Couzin *et al.* (10), long-time steady-state computations revealed four different types of stable collective motion in different parameter regions: swarm, torus, dynamic parallel, and highly parallel. It was also shown that by changing the quantitative features of the behavioral rules (increasing or decreasing the radius of alignment), the collective state of the school could be changed.

In ref. 10, stochasticity is incorporated by adding a small deviation to the heading of each individual obtained from the deterministic evolution algorithm. Our simulations show that if one instead considers relatively rare but substantial variations, namely that there is a small probability of each individual changing its direction substantially from that obtained from the deterministic algorithm, then for certain parameter regions multiple successive transitions between the torus and the dynamic parallel state can occur. See supporting information (SI) Fig. 7.

In this paper, we study a one-dimensional individual-based model for group formation with stochasticity included along the lines of the variation described above. This system exhibits repeated stochasticity-induced switching between distinct ordered collective motion states. This switching appears similar in nature but is, as we will show, different in detail, from results obtained by using other models. In those cases, collective motion transitions between "symmetry-related" states [e.g., between clockwise and counterclockwise motions for marching locusts constrained to a ring (12), or the "alternating flock" in refs. 13 and 14], stochastically driven transitions between ordered and disordered states mediated by clustering (15), mixed-phase states at phase transition boundaries (16), or transitions that do not occur repeatedly (17) were observed.

Individual-based models are often used in the study of animal groups because they can incorporate biologically realistic behavioral responses and social interactions that might be discontinuous (e.g., characterized by thresholds or if/then rules) or stochastic in nature; they also can support complex network topologies, allow for individual variability, and enable the study of the relationship between adaptive individual behavior and emergent properties (18, 19). Most analysis of individual-based models, however, relies on long-time simulations, which can be extremely costly and difficult to interpret and analyze (2, 19). In this paper, we introduce a framework for the coarse analysis of stochasticity-induced switching between collective motion states for individual-based models. We characterize the behavior of the model with a single "coarse observable," $A(t)$, a scalar variable that quantifies the global structure of the school. We show computational evidence to support that $A(t)$ parameterizes a one-dimensional, attracting, invariant "slow manifold," which characterizes the long-term dynamics of the system. This suggests that we can use an effective Fokker–Planck (FP) equation to describe the dynamics of the probability distribution $P(A)$, whose drift and diffusion coefficients are determined by the short-time evolution of the first two moments of A. We locally estimate these coefficients by developing a "lifting" procedure, which enables the initialization of brief bursts of simulations of the individual-based model at a given value of A (20). This framework allows us to construct an effective potential, thereby enabling coarse bifurcation analysis and estimation of the mean residence times in each state, without having to rely on computationally expensive "long-time" equilibrium simulations of the individual-based model.

Results

Model. N agents with positions $c_i(t)$ and unit directions $v_i(t) = \pm 1$, $i = 1, \ldots, N$, move on the line with constant speed s. Time is partitioned into steps of size τ units, corresponding to finite response time of the agents. At the beginning of every time step, each agent instantly updates its direction by considering the positions and directions of agents surrounding it in three nonoverlapping zones (see Fig. 1): the zone of repulsion $Z_{r_i}(t) =$

Author contributions: A.K., J.M., and I.G.K. performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Abbreviation: FP, Fokker–Planck.

[‡]To whom correspondence should be addressed: E-mail: moehlis@engineering.ucsb.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0608270104/DC1.

Fig. 1. Nonoverlapping behavioral zones for the model: zone of repulsion Z_r, zone of orientation Z_o, and zone of attraction Z_a.

$(c_i(t) - r_r, c_i(t) + r_r)$, the zone of orientation $Z_{o_i}(t) = (c_i(t) - r_o, c_i(t) + r_o) \backslash Z_{r_i}(t)$, and the zone of attraction $Z_{a_i}(t) = (c_i(t) - r_a, c_i(t) + r_a) \backslash (Z_{r_i}(t) \cup Z_{o_i}(t))$, where r_r is the radius of repulsion, r_o is the radius of orientation, and r_a is the radius of attraction. We define $\Delta r_o = r_o - r_r$ and $\Delta r_a = r_a - r_o$. These zones are used to define rules that are abstractions of the behavioral tendencies seen in animal groups in nature, the first being that animals tend to repel away from those that are too close, and the second that if they are not so repelled, they tend to align with and feel an attraction toward their neighbors (21, 22). Specifically, if agents can be found within an individual i's zone of repulsion, then it orients its direction away from the average relative directions of those within its zone of repulsion. The desired direction of agent i is then given by

$$v_i(t + \tau) = -\sum_{\substack{c_j(t) \in Z_{r_i}(t) \\ i \neq j}} \frac{c_j(t) - c_i(t)}{|c_j(t) - c_i(t)|}. \quad [1]$$

If individual i does not find agents within its zone of repulsion, it orients its direction toward an equally weighted combination of the average orientations of itself and those within its zone of orientation, $o_i(t)$, and the average relative directions of those within its zone of attraction, $a_i(t)$. The desired direction of agent i is then given by

$$v_i(t + \tau) = o_i(t) + a_i(t), \quad [2]$$

$$o_i(t) = \left(v_i(t) + \sum_{c_j(t) \in Z_{o_i}(t)} v_j(t) \right) \Bigg/ \left| v_i(t) + \sum_{c_j(t) \in Z_{o_i}(t)} v_j(t) \right|,$$

$$a_i(t) = \sum_{c_j(t) \in Z_{a_i}(t)} \frac{c_j(t) - c_i(t)}{|c_j(t) - c_i(t)|} \Bigg/ \left| \sum_{c_j(t) \in Z_{a_i}(t)} \frac{c_j(t) - c_i(t)}{|c_j(t) - c_i(t)|} \right|.$$

If $v_i(t + \tau) = 0$, the agent maintains its direction from the previous time step, so that $v_i(t + \tau) = v_i(t)$. Otherwise, the desired direction $v_i(t + \tau)$ of agent i is normalized as

$$v_i(t + \tau) \rightarrow \frac{v_i(t + \tau)}{|v_i(t + \tau)|}, \quad v_i(t + \tau) \neq 0. \quad [3]$$

Stochasticity is added so that each agent changes the sign of its desired direction with probability p. Each agent's position is updated according to

$$c_i(t + \tau) = c_i(t) + s v_i(t + \tau)\tau. \quad [4]$$

To begin a simulation, N individuals are placed randomly on the interval $[-N/4, N/4]$ with random directions, chosen so that each agent initially interacts with at least one other agent.

We observe that the model can display two metastable cohesive collective states, which we call "stationary" and "mobile." In the stationary state, the individual dynamics are driven by repulsion. The school remains approximately stationary in time, with neighboring agents typically having opposite directions, and each agent typically changing its direction at each time step to avoid neighbors to its right or left. We interpret this ordered stationary state as a one-dimensional analog of circular milling behavior. In the mobile state, the individual dynamics are driven by orientation and attraction, and the school coherently travels in the positive or negative direction. This is (one-dimensional) parallel motion. For certain values of the parameters we find "stick/slip behavior," in which the school alternates at apparently random times between the stationary and mobile states; such transitions arise from random fluctuations in the directions of individuals because of the stochasticity of the model (see Fig. 2). In what follows, we focus primarily on parameters for which both the stationary and mobile states are metastable.

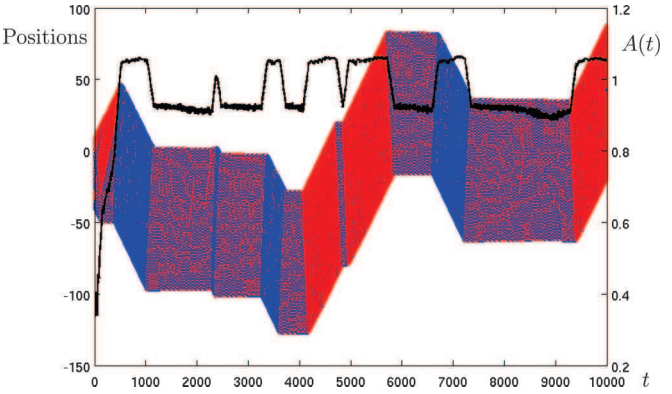

Fig. 2. Positions of $N = 100$ agents for a 10^4 step run, with parameters $s = 0.75$, $\tau = 0.1$, $r_r = 1$, $\Delta r_o = 0.6$, $\Delta r_a = 1$, $p = 0.001$, blue (resp., red) indicates motion of an agent in the positive (resp., negative) direction. The black line shows the corresponding time series plot of the coarse observable $A(t)$.

Coarse Observable. Through simulations we are led to hypothesize that the dynamics of this model can be suitably characterized by a single coarse observable

$$A(t) = \frac{1}{N} \sum_{i=1}^{N} \min_{\substack{j \\ j \neq i}} |c_j(t) - c_i(t)|, \quad [5]$$

the average nearest neighbor distance. This variable has been used in fish schooling models as a measure of the global structure of the school (6). $A(t)$ can distinguish between the stationary and mobile states as long as the school is not fragmented into subgroups displaying different collective dynamics. When the system is in the stationary state, typically $A(t) < r_r$ (repulsion driven), and when the system is in the mobile state, typically $A(t) > r_r$ (orientation and attraction driven; see Fig. 2).

Computational Observations. For our simulations, we fix $N = 100$, $s = 0.75$, $\tau = 0.1$, $r_r = 1$, $\Delta r_a = 1$, and $p = 0.001$, and let Δr_o vary. (For these parameters, when fragmentation into smaller, noninteracting subschools occurs, typically one subschool is composed of only a few agents, so that $A(t)$ remains a good measure of the collective state.) For each value of Δr_o studied, data were taken from 100 runs lasting 10^4 steps. For Δr_o sufficiently small, the school remains in the stationary state for the duration of our simulation. As Δr_o is increased to ≈ 0.14, the school exhibits stick/slip behavior in which it transitions at apparently random times between the stationary and mobile states. Transitions between these states typically begin with a stochastic change in direction of an agent at the edge of the school, which then "propagates" through the rest of the school (cf. refs. 13 and 14). As an example, for $\Delta r_o = 0.6$, the mean residence time for the stationary state is 994 steps, and for the mobile state, 509 steps. For $\Delta r_o > 1.08$, the school remains in the mobile state for the entire duration of our simulations.

Observing the steady-state probability distributions for various values of the parameter Δr_o, shown in Fig. 3, one can see the signature of the transitions between the stationary and mobile

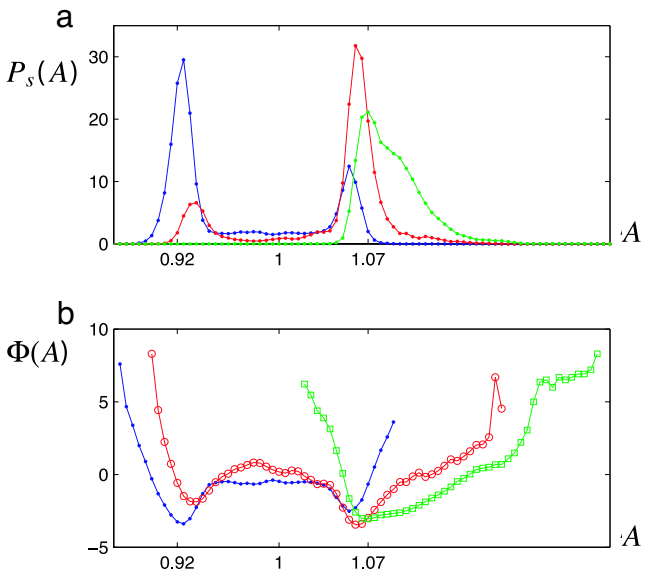

Fig. 3. Probability distribution functions (*a*) and effective potentials (*b*) for 100 trials with 10^4 steps per trial: $N = 100$, $r_r = 1$, $\Delta r_a = 1$, $s = 0.75$, $\tau = 0.1$, $p = 0.001$, $\Delta r_o = 0.6$ (blue), $\Delta r_o = 1$ (red), and $\Delta r_o = 1.1$ (green).

states. The probability distribution peaks at approximately $A = w_1 \equiv r_r - s\tau$, corresponding to the stationary state, and at approximately $A = w_2 \equiv r_r + s\tau$, corresponding to the mobile state, a distance $\approx 2s\tau$ from the stationary state. This may be rationalized by considering the dynamics of the stationary state, which is characterized by each agent typically changing its direction at every single time step to avoid its neighbor to the right or left, with A remaining nearly constant. Thus, one expects the individuals to be spaced approximately at alternating distances of d_1 and d_2 where $d_1 < r_r$ and $d_2 \approx d_1 + 2s\tau > r_r$. When the group exhibits a transition from the stationary to the mobile state, the distance d_2 "propagates" throughout the school so that $A \approx d_2$ in the mobile state.

In our simulations we found that the locations of the peaks of the stationary probability distribution depend somewhat on the details of the initial positions of the agents. This is a straightforward consequence of an important property of the model: in determining the desired direction of a given agent at the next time step, the only positional information used is which (if any) zone the other agents are in. Thus, agents can be moved slightly without changing their zones and hence with no change to the dynamics, but with a change to the value of A. These "neutrally stable" states have consequences for our lifting procedure described in *Methods*.

Probabilistic Description. We assume that the system dynamics at the macroscopic level may be suitably characterized by our single coarse variable $A(t)$. We therefore consider describing the evolution of the probability distribution function $P(A)$ with an effective FP equation (cf. refs. 23–25):

$$\frac{\partial P(A,t)}{\partial t} = -\frac{\partial}{\partial A}[V(A)P(A,t)] + \frac{\partial^2}{\partial A^2}[D(A)P(A,t)], \quad [6]$$

where $V(A)$ is the drift coefficient, and $D(A) > 0$ is the diffusion coefficient (26–28). These terms are related to the short-time evolution of the first two moments of A as

$$V(A_0) = \left.\frac{\partial\langle A(t;A_0)\rangle}{\partial t}\right|_{t=0}, D(A_0) = \frac{1}{2}\left.\frac{\partial\sigma^2(t;A_0)}{\partial t}\right|_{t=0}, \quad [7]$$

where $A(t; A_0)$ denotes a trajectory initialized at A_0 at $t = 0$, angular brackets denote ensemble averaging over different realizations of the trajectory, and σ^2 denotes the variance of A for such an ensemble. See *SI Appendix* for a detailed derivation. The FP equation is equivalent to the Itô stochastic differential equation

$$dA = V(A)dt + \sqrt{2D(A)}dW, \quad [8]$$

with $W(t)$ a Wiener process (26). In the limit $D(A) = 0$, Eq. **8** describes the deterministic motion of A subject to the "effective potential" $U(A) = -\int_{-\infty}^{A} V(A')dA' + const.$ In general, an effective potential $\Phi(A)$ can be obtained from the stationary probability distribution function $P_s(A)$, which satisfies the steady state ($\partial/\partial t = 0$) FP equation. Defining

$$P_s(A) \sim \exp(-\Phi(A)), \quad [9]$$

it follows that $\Phi(A)$ satisfies (23)

$$\Phi(A) = \log(D(A)) - \int_{-\infty}^{A} \frac{V(A')}{D(A')}dA' + const. \quad [10]$$

When $D(A) = const.$, this corresponds to Brownian motion of A subject to an effective potential proportional to $U(A)$.

Effective Potential from Long-Time Simulations. We first construct the effective potential by obtaining a stationary probability distribution function directly from an ensemble of long-time simulations and by using Eq. **9** to compute the effective potential as $\Phi(A) = -\log(P_s(A)) + const.$ An ensemble of 100 simulations of 10^4 steps each was performed for a wide range of values of the parameter Δr_o. The probability distributions and corresponding effective potentials for three characteristic values of Δr_o are shown in Fig. 3. To study the dependence of the behavior of the system on parameters, the critical points of the effective potential were followed as Δr_o was varied. This is a useful practical analog of deterministic bifurcation diagrams for this stochastic case. The minima of the effective potential correspond to points on the stable branch of the bifurcation diagram and the maxima correspond to points on the unstable branch. (In determining the maxima, we perform a quadratic fit of the effective potential between the two prominent wells. This filters out spurious small minima and maxima, which can arise because of fragmentation of the school.) The bifurcation diagram is shown in Fig. 4. Two saddle node bifurcations are found at approximately $\Delta r_o = 0.14$ and $\Delta r_o = 1.09$, and the system appears bistable for values of Δr_o between these values.

We also construct the effective potential by using ensembles of long-time simulations as a database. Here A is discretized over a grid of values that appear in the database: $A_0 = 0.88 + mk, m = 0.005$ (mesh size), $k = 0,1, \ldots, 42$ for $\Delta r_o = 0.6$. Then, for each A_0 over the grid, $V(A_0)$ and $D(A_0)$ are approximated by using Eq. **7** as follows. Every appearance of A_0 (within a certain error tolerance) as well as its subsequent values over a short fixed time interval of length $t = 10$ steps are saved. The ensemble mean $\langle A(t;A_0)\rangle$ and variance $\sigma^2(t;A_0)$ are then computed by averaging over these short trajectories. $V(A_0)$ (resp., $D(A_0)$) are then estimated by taking the slope of the linear regression of $\langle A(t;A_0)\rangle$ (resp., $\sigma^2(t;A_0)$). Finally, $\Phi(A)$ is estimated by numerically approximating the integral in Eq. **10**.

Results are shown for both methods for $\Delta r_o = 0.6$ in Fig. 5. The effective potential obtained by using the second approach agrees quite well with that obtained by using the first approach, which confirms that A is a good dynamic observable. These methods, however, do not offer any computational savings because we had to compile data from sufficiently extensive temporal simulations

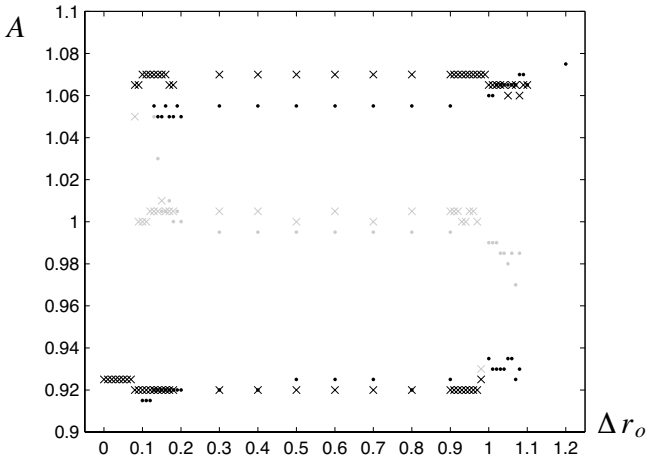

Fig. 4. Coarse bifurcation diagram showing the critical points of the effective potential as Δr_o is varied. The minima of the effective potential correspond to the stable branch (black) and the maxima correspond to the unstable branch (gray). Dots (resp., ×) show critical points of the long-time (resp., short-time) simulation estimate of the effective potential. In both cases, the unstable solutions were located by performing a quadratic fit of the data between the two wells and then locating the maxima of the fit.

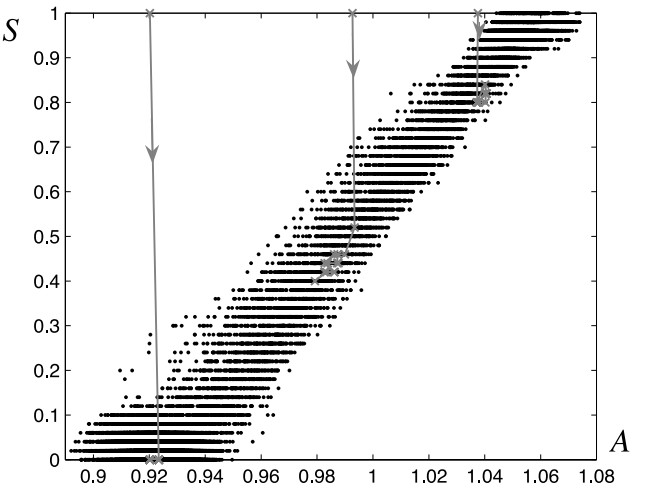

Fig. 6. Sample trajectories initialized using the lifting procedure approaching an apparent slow manifold parameterized by A. Trajectories (gray) evolve for 15 steps with arrows showing the direction of increasing time. Black dots represent a numerical approximation of the slow manifold in the (A, S) plane. Data were taken from an ensemble of five 10^4-step simulations after steady state was reached ($\approx 10^3$ steps) for $\Delta_{r_o} = 0.6$.

(an "equilibrium run"), which take ≈10 h for 100 trials of 10^4 steps each running with Matlab (MathWorks, Natick, MA) on a standard workstation.

Effective Potential from Short-Time Simulations. Recently, Kevrekidis *et al.* (20) developed an "equation-free" computational framework for extracting population-level information from individual-based models; the term "equation-free" arises because the population-level equations are not explicitly known (20). The approach relies on the assumption that the system state variables can be separated into a subset of fast variables and a low-dimensional subset of slow or coarse variables (in our case A), which parameterize an attracting invariant slow manifold. If a simulation is appropriately initialized at a prescribed value A_0, then after a short time, once all of the fast variables have equilibriated, one will in effect sample the slow manifold at A_0. This framework can be used to estimate on demand (without long-time simulation) the drift and diffusion terms in the effective FP equation (24, 25).

The drift and diffusion coefficients were estimated by performing an ensemble of 10^3 simulations for each A_0 initialized by using the lifting procedure described in *Methods*. A linear fit of the data was performed after waiting a short "healing" time of 15 steps. The local drift and diffusion terms and the resulting effective potential, obtained by using the designed initializations, are in good agreement with those obtained by averaging multiple long-time simulations (see Figs. 4 and 5) and require approximately a factor of 5 less computation.

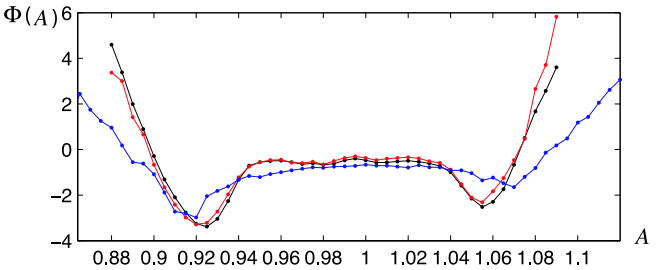

Fig. 5. Comparison of effective potentials for $N = 100$, $r_r = 1$, $\Delta r_o = 0.6$, $\Delta r_a = 1$, $s = 0.75$, $\tau = 0.1$, and $p = 0.001$. Black line, $\Phi(A) = -\log(P_s(A))$ where $P_s(A)$ was obtained from 100 trials with 10^4 steps per trial. Red line, Eq. **10** with drift and diffusion terms estimated from the same database. Blue line, Eq. **10** with drift and diffusion terms estimated by using the lifting procedure to initialize ensembles of short trajectories with the same A_0.

The construction of an effective potential from short-time simulations using the lifting procedure also allows one to predict mean residence times using the Kramers formula (26) without having to observe the average time spent in each collective state from ensembles of long-time simulations. We find that such predictions are reasonably accurate: for example, for $\Delta r_o = 0.6$ the mean residence time predictions are 1,537 steps for the stationary state and 424 steps for the mobile state.

Discussion

The individual-based model of self-organizing group formation analyzed in this paper shows that animal groups can repeatedly switch between qualitatively different ordered collective motion states entirely due to stochastic effects. In particular, changes to behavioral rules or the environment are not necessary for such transitions to occur. Instead, such switching relies on the presence of at least two metastable collective motion states, and stochasticity of appropriate type and strength to allow transitions to occur. The use of effective potentials constructed from long- as well as short-time simulations allows a powerful characterization of such switching.

Because the stochasticity that leads to switching is imposed at the level of individuals, this analysis suggests that random decisions by a small number of individuals can change an entire population's collective behavior, in particular when these individuals are near the edge of the school. The stochasticity-induced switching discussed in this paper complements recent simulations for a related model, which indicate that a small number of informed individuals can influence group dynamics (29). One can imagine that a combination of these effects might also be important: for example, a small number of individuals might spot a predator and quickly, randomly change their directions, an "informed stochasticity," which leads to a change in the entire group's motion, which could allow all individuals to escape (cf. ref. 30).

The framework developed in this paper provides a useful, "equation-free" computer-assisted approach to the analysis of emergent phenomena in individual-based aggregation models. Most analysis of individual-based models in the field of group formation has relied on costly long-time simulations, which has limited the number of individuals that can be simulated as well as the types of analysis that can be realistically performed (11). Our approach allows one to achieve a new level of understanding and quantification of biological self-organization by bridging individual-based modeling with coarse population-level analysis.

Methods

The most challenging step of equation-free computation is "lifting": the construction of one or more states "consistent with" the prescribed value of the coarse observable A_0. Because of the neutrally stable states mentioned in *Computational Observations*, to equilibrate to the desired slow manifold we must take care when placing agents at a given value of A_0. From these computations, we found that the distribution of distances between individuals in the stick/slip state is bimodal, with peaks at w_1 and w_2. We use this information in our lifting algorithm to place individuals at a given A_0 as follows.

1. Calculate proportions p_1 and p_2 of distances w_1 and w_2 so that $p_1w_1 + p_2w_2 = A_0$. If $A_0 < w_1$, set $w_1 = A_0$ and if $A_0 > w_2$, set $w_2 = A_0$.
2. Draw the appropriate proportion p_1 and p_2 of distances to first nearest neighbors from tight Gaussian distributions centered at w_1 and w_2. Place them randomly in the vector d_1.
3. Draw distances to second nearest neighbors from a tight Gaussian distribution centered at w_2. Place them randomly in the vector d_2.
4. Start the first agent at some position c_1. Place the second agent at position $c_2 = c_1 + d_1(1)$. Place the third agent at position $c_3 = c_2 + d_2(1)$. Place the fourth agent at position $c_4 = c_3 + d_1(2)$. Continue this process until N agents have been positioned.
5. Let $v_i = 1$, $\forall$ i.

To validate this lifting procedure, we consider another coarse variable

$$S(t) = \frac{1}{N}\left|\sum_{i=1}^{N} v_i(t)\right|, \qquad [11]$$

the group polarization, which has also been used in many fish schooling models as a measure of school structure (6, 10). When $S \approx 1$ (resp., $S \approx 0$) the school is in the mobile (resp., stationary) state. Our lifting procedure initializes the population with $S = 1$, and the time scale of approach to the slow manifold is comparable whether $A_0 > r_r$ or $A_0 < r_r$ (see Fig. 6). The quick relaxation to the slow manifold for $A_0 < r_r$ occurs because rule **1** causes the agents to try to immediately move away from each other.

We thank Danny Grünbaum, Simon Levin, Iain Couzin, and Bjorn Birnir for helpful discussions related to this work. This work was supported by National Science Foundation Grant NSF-0434328. J.M. also was supported by an Alfred P. Sloan Research Fellowship in Mathematics, and I.G.K. was supported by a Guggenheim Fellowship.

1. Camazine S, Deneubourg J-L, Franks NR, Sneyd J, Theraula G, Bonabeau E (2001) *Self-Organization in Biological Systems* (Princeton Univ Press, Princeton).
2. Parrish JK, Edelstein-Keshet L (1999) *Science* 284:99–101.
3. Koltes KH (1984) *Mar Biol* 78:113–122.
4. Aoki I (1982) *Bull Jap Soc Fish* 48:1081–1088.
5. Warburton K, Lazarus J (1991) *J Theor Biol* 150:473–488.
6. Huth A, Wissel C (1994) *Ecol Model* 75:135–146.
7. Reuter H, Breckling B (1994) *Ecol Model* 75:147–159.
8. Romey WL (1996) *Ecol Model* 92:65–77.
9. Czirók A, Vicsek T (2001) in *Fluctuations and Scaling in Biology*, ed Vicsek T (Oxford Univ Press, Oxford), pp 177–242.
10. Couzin ID, Krause J, James R, Ruxton GD, Franks NR (2002) *J Theor Biol* 218:1–11.
11. Parrish JK, Viscido S, Grünbaum D (2002) *Biol Bull* 202:296–305.
12. Buhl J, Sumpter DJT, Couzin ID, Hale JJ, Despland E, Miller ER, Simpson SJ (2006) *Science* 312:1402–1406.
13. O'Loan OJ, Evans MR (1999) *J Phys A* 32:L99–L105.
14. Raymond JR, Evans MR (2006) *Phys Rev E* 72:036112.
15. Huepe C, Aldana M (2004) *Phys Rev Lett* 92:168701.
16. Grégoire G, Chaté H, Tu Y (2003) *Physica D* 181:157–170.
17. Erdmann U, Ebeling W, Mikhailov AS (2005) *Phys Rev E* 71:051904.
18. Grimm V, Railsback SF (2005) *Individual-Based Modeling and Ecology* (Princeton Univ Press, Princeton).
19. Bonabeau E (2002) *Proc Natl Acad Sci USA* 99:7280–7287.
20. Kevrekidis IG, Gear CW, Hyman JM, Kevrekidis PG, Runborg O, Theodoropoulos T (2003) *Comm Math Sci* 1:715–762.
21. Krause J, Ruxton GD (2002) *Living in Groups* (Oxford Univ Press, Oxford).
22. Partridge BL (1982) *Sci Am* 246:90–99.
23. Haataja M, Srolovitz DJ, Kevrekidis IG (2004) *Phys Rev Lett* 92:160603.
24. Hummer G, Kevrekidis IG (2003) *J Chem Phys* 118:10762–10773.
25. Kopelevich DI, Panagiotopoulos AZ, Kevrekidis IG (2005) *J Chem Phys* 122:044908.
26. Gardiner CW (2004) *Handbook of Stochastic Methods* (Springer, Berlin).
27. Risken H (1996) *The Fokker–Planck Equation: Methods of Solution and Applications* (Springer, Berlin).
28. Coffey WT, Kalmykov YP, Waldron JT (2004) *The Langevin Equation: With Applications to Stochastic Problems in Physics, Chemistry, and Electrical Engineering* (World Scientific, Singapore).
29. Couzin ID, Krause J, Franks NR, Levin SA (2005) *Nature* 433:513–516.
30. Inada Y, Kawachi K (2002) *J Theor Biol* 214:371–387.

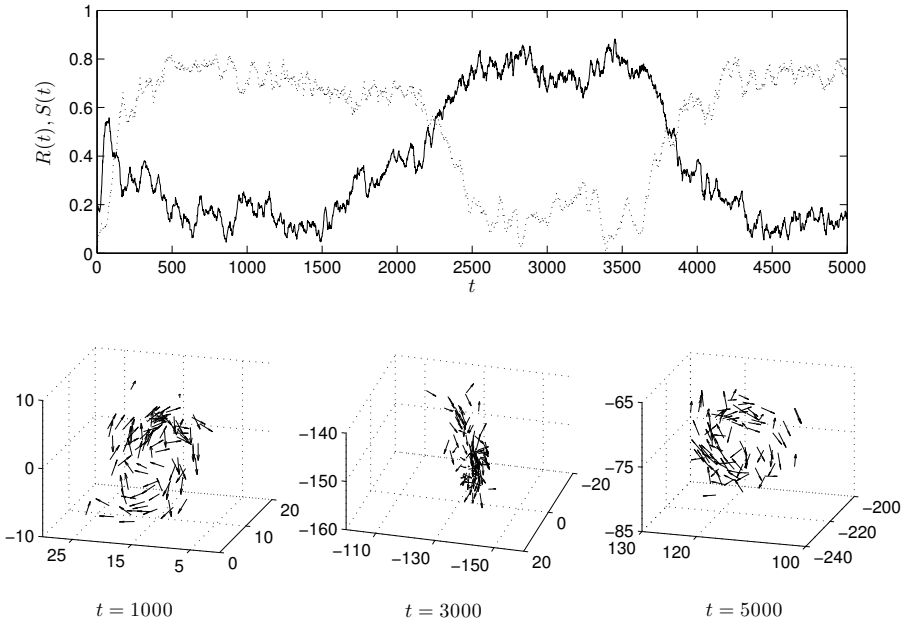

Figure 7: Repeated transitions between the torus state and the dynamic parallel state for a modified version of the three-dimensional schooling model of Couzin *et al.* (12) with $r_r = 1$, $\Delta r_o = 2$, and $\Delta r_a = 9$. We modify (12) by considering rare but substantial variations in the headings of individuals obtained from the deterministic algorithm. Specifically, each agent changes the sign of its desired direction with probability $p = 0.005$, with the new direction chosen according to a spherically wrapped Gaussian distribution with standard deviation $\sigma = 5$. (This differs from (12) for which $p = 1$ and $\sigma = .05$, so that all agents modify their direction at each timestep with small variation.) Two coarse variables measure the collective behavior (12): (dashed) $R(t)$ = angular momentum, (solid) $S(t)$ = group polarization; both are normalized so that they reach their maximum value at 1 and minimum at 0.

Appendix: Derivation of the Fokker-Planck Equation

Let $\{X(t) : t \geq 0\}$ be a one-dimensional stochastic process with $t_1 > t_2 > t_3$. We use $P(X_1, t_1; X_2, t_2)$ to denote the joint probability distribution, i.e., the probability that $X(t_1) = X_1$ and $X(t_2) = X_2$, and $P(X_1, t_1 \mid X_2, t_2)$ to denote the conditional (or transition) probability distribution, i.e., the probability that $X(t_1) = X_1$ given that $X(t_2) = X_2$, defined as $P(X_1, t_1; X_2, t_2) = P(X_1, t_1 \mid X_2, t_2)P(X_2, t_2)$. We will assume $X(t)$ is a Markov process, namely,

$$P(X_1, t_1 \mid X_2, t_2; X_3, t_3) = P(X_1, t_1 \mid X_2, t_2). \quad [1]$$

For any continuous state Markov process, the following Chapman-Kolmogorov equation is satisfied (1, 2):

$$P(X_1, t_1 \mid X_3, t_3) = \int P(X_1, t_1 \mid X_2, t_2)P(X_2, t_2 \mid X_3, t_3)dX_2. \quad [2]$$

In the following, we will also assume $X(t)$ is time homogeneous:

$$P(X_1, t_1 + s; X_2, t_2 + s) = P(X_1, t_1, X_2, t_2), \quad [3]$$

so that X is invariant with respect to a shift in time. For simplicity of notation, we use $P(X_1, t_1 - t_2 \mid X_2) \equiv P(X_1, t_1 \mid X_2, t_2)$.

We will now outline the derivation of the Fokker-Planck equation, a partial differential equation for the time evolution of the transition probability density function. This closely follows the derivation in ref. 3. Consider

$$\int_{-\infty}^{\infty} h(Y)\frac{\partial P(Y, t \mid X)}{\partial t}dY, \quad [4]$$

where $h(Y)$ is any smooth function with compact support. Writing

$$\frac{\partial P(Y, t \mid X)}{\partial t} = \lim_{\Delta t \to 0} \frac{P(Y, t + \Delta t \mid X) - P(Y, t \mid X)}{\Delta t}, \quad [5]$$

and interchanging the limit with the integral, it follows that

$$\int_{-\infty}^{\infty} h(Y)\frac{\partial P(Y, t \mid X)}{\partial t}dY = \lim_{\Delta t \to 0} \int_{-\infty}^{\infty} h(Y) \left[\frac{P(Y, t + \Delta t \mid X) - P(Y, t \mid X)}{\Delta t}\right] dY. \quad [6]$$

Applying the Chapman-Kolmogorov identity (Eq. **2**), the right hand side of Eq. **6** can be written as

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \left[\int_{-\infty}^{\infty} h(Y) \int_{-\infty}^{\infty} P(Y, \Delta t \mid Z)P(Z, t \mid X)dZdY - \int_{-\infty}^{\infty} h(Y)P(Y, t \mid X)dY\right]. \quad [7]$$

Interchanging the limits of integration in the first term of Eq. **7**, letting $Y \to Z$ in the second term, and using the identity $\int_{-\infty}^{\infty} P(Y, \Delta t \mid Z)dY = 1$, we have

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \left[\int_{-\infty}^{\infty} P(Z, t \mid X) \int_{-\infty}^{\infty} P(Y, \Delta t \mid Z) \left(h(Y) - h(Z)\right) dY dZ\right]. \quad [8]$$

Taylor expanding $h(Y)$ about Z gives

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \left[\int_{-\infty}^{\infty} P(Z, t \mid X) \int_{-\infty}^{\infty} P(Y, \Delta t \mid Z) \sum_{n=1}^{\infty} h^{(n)}(Z)\frac{(Y - Z)^n}{n!}dY dZ\right]. \quad [9]$$

Defining the jump moments as

$$D^{(n)}(Z) = \frac{1}{n!} \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{-\infty}^{\infty} (Y - Z)^n P(Y, \Delta t \mid Z)dY, \quad [10]$$

it follows that

$$\int_{-\infty}^{\infty} h(Y)\frac{\partial P(Y,t \mid X)}{\partial t}dY = \int_{-\infty}^{\infty} P(Z,t \mid X)\sum_{n=1}^{\infty} D^{(n)}(Z)h^{(n)}(Z)dZ. \quad [11]$$

Integrating each term on the right side of Eq. **11** by parts n times and using the assumptions on h, after moving terms to the left hand side, it follows that

$$\int_{-\infty}^{\infty} h(Z)\left(\frac{\partial P(Z,t \mid X)}{\partial t} - \sum_{n=1}^{\infty}\left(-\frac{\partial}{\partial Z}\right)^n \left[D^{(n)}(Z)P(Z,t \mid X)\right]\right)dZ = 0. \quad [12]$$

Now, because h is an arbitrary function, it is necessary that

$$\frac{\partial P(Z,t \mid X)}{\partial t} = \sum_{n=1}^{\infty}\left(-\frac{\partial}{\partial Z}\right)^n \left[D^{(n)}(Z)P(Z,t \mid X)\right]. \quad [13]$$

We define the probability distribution function $P(X,t)$ of $X(t)$ as the solution of Eq. **13** with initial condition given by a δ-distribution at X_0 at $t = 0$. In this case, $P(X,t) \equiv P(X,t \mid X_0,0)$ and we may write Eq. **13** as

$$\frac{\partial P(X,t)}{\partial t} = \sum_{n=1}^{\infty}\left(-\frac{\partial}{\partial X}\right)^n \left[D^{(n)}(X)P(X,t)\right], \quad [14]$$

with

$$D^{(n)}(X_0) = \frac{1}{n!}\lim_{\Delta t\to 0}\frac{1}{\Delta t}\left\langle [X(t+\Delta t) - X(t)]^n\right\rangle|_{t=0}, \quad [15]$$

which is commonly called the Kramers-Moyal expansion. Now, if we assume $D^{(n)}(X) = 0$ for $n > 2$, then we have the Fokker-Planck equation:

$$\frac{\partial P(X,t)}{\partial t} = -\frac{\partial}{\partial X}\left[V(X)P(X,t)\right] + \frac{\partial^2}{\partial X^2}\left[D(X)P(X,t)\right], \quad [16]$$

where, $V(X) \equiv D^{(1)}(X)$ is the drift coefficient and $D(X) \equiv D^{(2)}(X) > 0$ is the diffusion coefficient, which can be written as

$$V(X_0) = \left.\frac{\partial\langle X(t;X_0)\rangle}{\partial t}\right|_{t=0}, \qquad D(X_0) = \left.\frac{1}{2}\frac{\partial\sigma^2(t;X_0)}{\partial t}\right|_{t=0}, \quad [17]$$

where angular brackets denote ensemble averaging, σ^2 denotes the variance of X, and $X(t;X_0)$ denotes a realization with $X(0) = X_0$. Any stochastic process $X(t)$ whose probability distribution function satisfies the Fokker-Planck equation is known mathematically as a diffusion process (1).

References

1. Gardiner, C. W. (2004) *Handbook of Stochastic Methods.* (Springer, Berlin).

2. Risken, H. (1996) *The Fokker-Planck Equation: Methods of Solution and Applications.* (Springer, Berlin).

3. Coffey, W. T, Kalmykov, Y. P, & Waldron, J. T. (2004) *The Langevin Equation: With Applications to Stochastic Problems in Physics, Chemistry, and Electrical Engineering.* (World Scientific, Singapore).