Environmental Path-Entropy and Collective Motion

Harvey L. Devereux
Department of Mathematics, University of Warwick, Coventry CV4 7AL, United Kingdom

Matthew S. Turner*
Department of Physics and Centre for Complexity Science, University of Warwick, Coventry CV4 7AL, United Kingdom and Department of Chemical Engineering, Kyoto University, Kyoto 615-8510, Japan

(Received 29 September 2022; revised 18 January 2023; accepted 28 March 2023; published 19 April 2023)

Inspired by the swarming or flocking of animal systems we study groups of agents moving in unbounded 2D space. Individual trajectories derive from a "bottom-up" principle: individuals reorient to maximize their future path entropy over environmental states. This can be seen as a proxy for *keeping options open*, a principle that may confer evolutionary fitness in an uncertain world. We find an ordered (coaligned) state naturally emerges, as well as disordered states or rotating clusters; similar phenotypes are observed in birds, insects, and fish, respectively. The ordered state exhibits an order-disorder transition under two forms of noise: (i) standard additive orientational noise, applied to the postdecision orientations and (ii) "cognitive" noise, overlaid onto each individual's model of the future paths of other agents. Unusually, the order increases at low noise, before later decreasing through the order-disorder transition as the noise increases further.

DOI: 10.1103/PhysRevLett.130.168201

Collective motion occurs in both living and synthetic systems. In living systems this arises in a wide variety of species over different length scales, e.g., micro-organisms, cells, insects, fish, birds [1–6], and even dinosaurs [7]. Interest in the physics community often lies in developing models of collective motion that are analogous to living systems, many of which exhibit ordered (coaligned) motion and support a noise-induced transition to disorder [8–15]. Long-ranged behavioral interactions may arise in nature, and there have been some attempts to analyze such interactions [13,16–20]. These can naturally be traced to the nature of information transfer between agents [21,22], noting that senses like vision are long ranged. Other models of swarming behavior incorporate explicit alignment, cohesion, and/or collision avoidance rules directly into an agent-based model [13,17,23]. However, such models cannot easily explain the underlying reason for the emergence of properties like cohesion and coalignment as these are essentially incorporated into the models at the outset. One recent alternative approach is to utilize machine learning based on using a simple form of perception to maintain cohesion directly [24]. Another involves the study of large deviations of nonaligning active particles that are biased, e.g., by effective alignment of self-propulsion with particle velocity [25–28]. While it is possible neural circuitry of animals encodes an algorithm that *directly* targets coalignment and cohesion in the same mathematical manner as in these models, it seems much more likely that some lower-level principle is involved. This principle, almost certainly associated with evolutionary fitness in some way, might then be the origin of cohesion and coalignment. We argue that more satisfactory explanations for the phenomenon of swarming may be offered by testing candidates for this lower-level principle. In this Letter, we analyze one such model.

There is a small but growing literature focusing on the causal understanding of complex behavior, cast as an entropy or state maximization approach. Here some measure of variation across *future* paths accessible from a particular system configuration is computed, and an action that maximizes this variation is selected, e.g., Refs. [29–33]. It is argued that agents that can retain access to the most varied future environments can better select from these to satisfy any immediate requirements or objectives, e.g., resource acquisition or predator evasion. For these reasons such strategies are expected to generally confer evolutionary fitness in an uncertain world. The present work shares similar motivation to Ref. [32] but provides a rigorous mathematical model based on path entropies and focuses on the emergence of order. We believe that such models offer clear advantages in terms of their conceptual clarity and prospects for future development.

To realize such a model here, agents are treated as oriented unit disks that move in discrete time t, defining our length and time units, respectively. The position of the ith agent in the next time step is

$$\mathbf{x}_i^{t+1} = \mathbf{x}_i^t + \mathbf{v}_i^{t+1}. \tag{1}$$

At each discrete time step t agents choose from $z = 5$ velocities: either to move along their current orientation with one of three speeds, namely, nominal v_0, fast $v_0 + \Delta v$, or slow $v_0 - \Delta v$; or to reorientate by an angle $\pm\Delta\theta$ while moving at the nominal speed v_0. Unless noted otherwise we take $v_0 = 10$, $\Delta v = 2$, and $\Delta\theta = \pi/12 = 15°$. The agent's velocity is updated by an operator $A_{\alpha_i^t}$ acting on its previous velocity $\mathbf{v}_i^t$:

$$\mathbf{v}_i^{t+1} = A_{\alpha_i^t}[\mathbf{v}_i^t]. \tag{2}$$

Actions α change the velocity according to

$$A_\alpha[\mathbf{v}] = v_\alpha R(\theta_\alpha)\hat{\mathbf{v}}. \tag{3}$$

The index $\alpha \in [1, z]$ labels possible actions, here with indices dropped for clarity. Hat accents denote unit vectors according to $\hat{\mathbf{v}} = \mathbf{v}/|\mathbf{v}|$ throughout, with $|\cdot|$ the Euclidean norm. The action chosen at each time step determines the corresponding speed of the agent $v_1 = v_4 = v_5 = v_0$, $v_2 = v_0 + \Delta v$, $v_3 = v_0 - \Delta v$ in that time step, where R generates a rotation

$$R(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \tag{4}$$

with rotation angles $\theta_1 = \theta_2 = \theta_3 = 0$, $\theta_4 = \Delta\theta$, and $\theta_5 = -\Delta\theta$. The sequence of such actions realized by this agent α_i^t over time t completely determines the dynamics. In order to select actions, i.e., compute hypothetical path entropy over future states, this model requires that agents model positions of themselves and other agents into the future. Therefore we adopt the notation $\tilde{\mathbf{x}}_k^{t'}$, $\tilde{\alpha}_k^{t'}$, $v_{\tilde{\alpha}_k^{t'}}$, and $\theta_{\tilde{\alpha}_k^{t'}}$, involving a tilde accent, to indicate virtual positions, actions, speeds, and rotation angles of all agents k at time t'. Hence

$$\tilde{\mathbf{x}}_k^{t+s} = \mathbf{x}_k^t + \sum_{t'=t}^{t-1+s} \tilde{v}_{\tilde{\alpha}_k^{t'}} \prod_{t''=t}^{t'} R(\theta_{\tilde{\alpha}_k^{t''}})\hat{\mathbf{v}}_k^t, \tag{5}$$

with $1 \le s \le \tau$ reflecting the time horizon τ.

Equation (5) generates the hypothetical position of both the $k = i$ (self) and $k = j \neq i$ (other) agents. However, we make a simplifying assumption for the motion of the $j \neq i$ (other) agents. Here our default model corresponds to "ballistic" translation of the $j \neq i$ agents in which $v_{\tilde{\alpha}_j^{t'}} = v_0$ and $\theta_{\tilde{\alpha}_j^{t'}} = 0$, $\forall\ t' \ge t$. The speeds and rotations depend neither on the particle index j nor the future time index, and so they can be stated in more condensed form simply as $\tilde{v} = v_0$ and $\tilde{\theta} = \theta_{\tilde{\alpha}} = 0$. The ballistic assumption can often be rather good, in the sense that the trajectories that are realized can have a very high degree of orientational order, and so the assumption is broadly self-consistent [34]. Later in this article we consider models that generate different virtual actions for the $j \neq i$ agents that incorporate noise. See Fig. 1(b) for a sketch of this dynamical scheme.

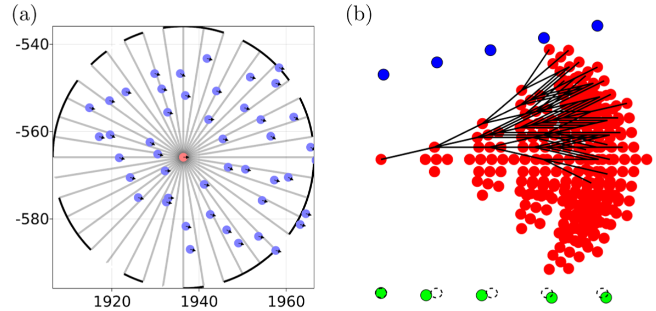

FIG. 1. Snapshot of a system configuration and sketch explaining our model. (a) $N = 50$ agents that take actions to maximize a future path entropy over environmental states (see text); axes show x-y coordinates in units of the agent radius. Overlaid (broken circle) is a representation of the visual state perceived by the red individual. Obtained from a simulation with parameters $\tau = 6$, $\Delta\theta = \pi/12 = 15°$, $v_0 = 10$, and $\Delta v = 2$ (see text for details). (b) In red the tree of hypothetical future actions the agent examines, starting from the present at the root on the far left. Shown in blue and green (with dashed circles) are ballistic and noisy motion assumptions of other agents.

The environmental state of an agent is assumed to be perceived using only vision; see Fig. 1(a). This state encodes information on the relative positions of the other agents in a manner that is broadly consistent with animal vision, abstracted to $d = 2$ dimensions: visual sensing involves a radial projection of all other agents onto a circular sensor array at each agent. Loosely speaking, the radial projection registers 0 "white" along lines of sight not intersecting agents, and 1 "black" along those that do. We discretize this into an n_s-dimensional visual state vector $\boldsymbol{\psi}_i$, for angular subregions of size $2\pi/n_s$. This then resembles a spin state, e.g., $(0, 1, 0, 0, 1\cdots)$.

Mathematically we use two indicator functions, first the distance of shortest approach along a line of sight $\hat{\mathbf{n}}_i = R(\chi)\hat{\mathbf{v}}_i$ originating from the ith agent,

$$I_{ij} = \Theta[1 - |\tilde{\mathbf{x}}_{ij} \times \hat{\mathbf{n}}_i(\chi)|], \tag{6}$$

where the Heaviside function $\Theta[x] = 1$ for $x \ge 0$ and 0 otherwise and $\tilde{\mathbf{x}}_{ij}$ is the separation vector $\tilde{\mathbf{x}}_j - \tilde{\mathbf{x}}_i$, with $|\cdot|$ the Euclidean norm. Equation (6) indicates an agent is visible along this line of sight in *either* direction from the

ith agent, i.e., along χ or $\chi + \pi$. We restrict to χ using the second indicator

$$I'_{ij} = \Theta[\tilde{\mathbf{x}}_{ij} \cdot \hat{\mathbf{n}}_i(\chi)]. \tag{7}$$

The nth component of the visual state vector $\boldsymbol{\psi}_i$ is then

$$\psi_i^n = \Theta\left[\int_{\sigma_n} \Theta\left[\sum_j I_{ij}(\chi) I'_{ij}(\chi)\right] d\chi - \frac{\pi}{n_s}\right], \tag{8}$$

where the nth sensor covers the angular domain $\sigma_n = [2\pi(n-1)/n_s, 2\pi n/n_s]$. The inner Heaviside function registers 1 ("black") if at least one agent intersects the line of sight χ; the integral then measures the coverage of σ_n by "black" regions. The outermost Heaviside function is a further threshold that at least half the sensor must be "black" to activate the nth visual state component.

For a virtual action $\tilde{\alpha}_i^t$ the entropy of the state distribution over (all nodes on) all virtual paths for the ith agent following action $\tilde{\alpha}_i^t$ is

$$S(\tilde{\alpha}_i^t) = -\sum_{\boldsymbol{\psi}} p_i(\tilde{\alpha}_i^t, \boldsymbol{\psi}) \log p_i(\tilde{\alpha}_i^t, \boldsymbol{\psi}), \tag{9}$$

where $p_i(\tilde{\alpha}_i^t, \boldsymbol{\psi})$ is the count of occurrences of a state $\boldsymbol{\psi}$ on these virtual paths, normalized by the count of states on all branches. In this way each action branch $\tilde{\alpha}_i^t$ is associated with an environmental path entropy $S(\tilde{\alpha}_i^t)$. The key step in the decision making process is that each agent then executes the action

$$\alpha_i^t = \arg\max_{\tilde{\alpha}_i^t} S(\tilde{\alpha}_i^t), \tag{10}$$

thereby choosing the branch that maximizes the entropy of future visual states. This process is carried out simultaneously for each agent and repeated, from scratch, at each time step. Degenerate options are selected at random, the only randomness in the baseline algorithm that is otherwise deterministic.

Our model supports various phenotypes. In Fig. 2 we report on the effect of varying the turning rate $\Delta\theta$, the nominal speed v_0, and its variation Δv. We find that a highly ordered and cohesive phenotype is commonly achieved when the agents move relatively fast with moderate turning, resembling those seen in flocks of social birds [37,38], noting that these birds also have relatively fast speed, do not slow significantly, and have limited turning ability relative to an insect. We also find cohesive disordered groups, some showing circulation. The most important conditions for the emergence of cohesive swarms are (i) $\tau \gtrsim 3$, (ii) $10 \lesssim n_s \lesssim 100$ to avoid the visual states becoming largely degenerate (see the Supplemental Material [39] for details).

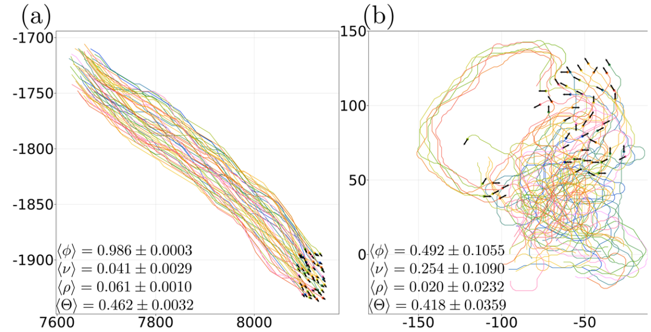

FIG. 2. Agents that maximize environmental path entropy naturally adopt different dynamical modes, or "phenotypes." Each panel shows the agent's trajectories, together with the time-averaged mean order ϕ, root-mean-squared vorticity ν, density ρ, and opacity Θ (see text): (a) the ordered, dense ("bird") phenotype and (b) translation combined with significant rotation (similar to "fish" or "insects"). Averages computed over ten replicates.

We report the visual opacity as the average sensor state $\Theta = \langle 1/n_s \sum_{n=1}^{n_s} \psi_i^n \rangle$, density $\rho = \langle N\pi r^2/\mathcal{A}^t \rangle$ with the convex hull area $\mathcal{A}^t$, and global order $\phi = \langle |1/N \sum_{i=1}^{N} \hat{\mathbf{v}}_i^t| \rangle$, and quantify rotation using a normalized mean-squared vorticity $\nu^2 = \langle (1/N \sum_i \hat{\boldsymbol{r}}_i^t \times \hat{\mathbf{v}}_i^t)^2 \rangle$, with $\boldsymbol{r}_i^t = \boldsymbol{x}_i^t - \langle \boldsymbol{x}_k^t \rangle_k$ and $\boldsymbol{v}_i^t$ the ith agent's position relative to the geometric center and velocity respectively. In each case we average over agents i and times t. We also use a measure of spatial clustering using DBScan [40] (see the Supplemental Material [39], Sec. S2, for details) to both detect fragmentations and measure the quantities above on clusters. We denote the average fraction of agents in the largest cluster as $\mathcal{C}$, and by $\phi_{\mathcal{C}}$ denote the order of the largest cluster.

We have established the emergence of coaligned, cohesive states under environmental path entropy maximizing trajectories; see Fig. 2(a). It is therefore natural to ask about the effect of noise on these dynamics. In this way we will investigate to what extent this model supports an order-disorder transition similar to those extensively studied in other models of collective motion [8–15].

By "cognitive noise" we mean some imprecision in an agent's model of the others. We therefore define a stochastic process for the virtual speeds $\tilde{v}^{t'} = v_0 + \mu_v^{t'}$ and rotations $\tilde{\theta}^{t'} = \mu_\theta^{t'}$ of all $j \neq i$ agents. Here both μ variables (subscript v, θ omitted for clarity) are drawn from zero mean $\langle \mu_j^{t'} \rangle = 0$ Gaussian distributions that are uncorrelated according to $\langle \mu_j^{t'} \mu_{j'}^{t''} \rangle = \eta^2 \delta_{jj'} \delta_{t't''}$ with $j, j' \neq i$ here the particle index of the (other) agents and $t', t'' \geq t$. The root-mean-squared noise amplitude of the speed and orientation are written, with subscripts restored, as η_v and η_θ respectively. An example is shown in the sequence of positions shown in green in Fig. 1(b). All else proceeds as before, without any additional noise applied to realized agent actions.

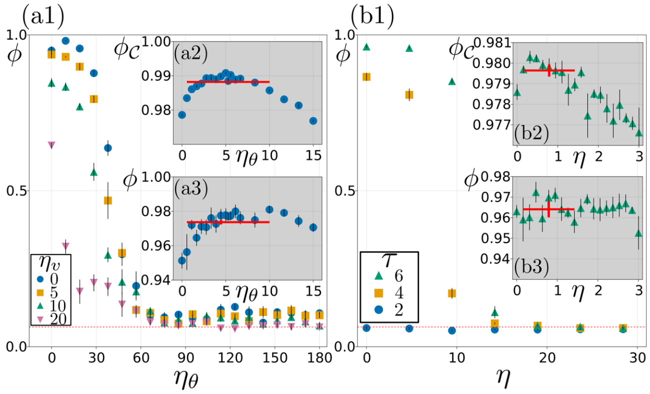

FIG. 3. Ordering transitions in the presence of (a) cognitive noise and (b) post-decision orientational noise. (a1)–(a3) Each agent approximates the future trajectories of others in the presence of *cognitive* noise as a sequence of random rotations and speed changes from their current heading and speed v_0. The noise strengths η_θ and η_v characterize the magnitude of the rotations (degrees) and speed changes (body radii per time step), respectively. (a1) A transition from high order to a disordered phase occurs with increasing cognitive rotational noise η_θ. Insets (a2) and (a3) focus on small η_θ with $\eta_v = 0$. They show the order of the largest cluster ϕ_C and the overall system order ϕ, respectively; note the maximum in order appears at nonzero noise. The red horizontal line shows the order averaged over all runs $0 < \eta \leq 10°$. In (a1)–(a3) the future time horizon is $\tau = 6$. (b1) The effect of postdecision orientational noise on global order ϕ. Here a random rotation with root-mean-squared angle η (degrees) is applied directly to the velocity, before the movement update. (b2) A statistically significant local maximum in ϕ_C for nonzero noise η whereas (b3) now shows no significant maximum in ϕ. The red horizontal line shows the order averaged over all runs $0 < \eta \leq 1.5°$. All systems contain $N = 250$; all error bars are 1 standard error in the mean; the dashed lines represent the mean order $\phi = 1/\sqrt{N}$ of randomly orientated agents. In all statistical tests additional repeats ($n = 16$) were computed for the zero noise case; see text for further details.

The most striking feature from Fig. 3(a) is that the order initially *increases* with the addition of small levels of noise, before later decreasing again. An upper tailed t-test (with unequal variances) on the difference of the mean order in the noise-free case ($\eta_\theta = 0$) and the mean of simulations with nonzero noise $0 < \eta \leq 10°$ rejects the null hypothesis that the mean order ϕ is the same at the level of $p < 10^{-13}$. The same t-test for ϕ_C, the order computed only for agents that are members of the largest cluster, is rejected at $p < 10^{-50}$.

To understand why a small amount of noise might actually increase order we compare the noise level at which the order is maximal, roughly $\eta_\theta = 5$–$7°$, to intrinsic variation in the realized dynamics in a low noise state $\phi \approx 0.98$. There are several ways to achieve this: (i) approximating the order as the mean component of the normalized velocities of the agents along the average direction of motion, $\phi = \langle \cos \delta\theta \rangle \approx 1 - \frac{1}{2}\langle \delta\theta^2 \rangle$, leading to $\delta\theta_{\mathrm{rms}} = 11°$; (ii) crudely assuming moves are uncorrelated extending for τ steps into the future and asking what angular noise amplitude *per time step*, analogous to η_θ would be required to give the realized order ϕ at the *end* of this sequence, leading to $11°/\sqrt{\tau} = 4.7°$; and (iii) using the velocity autocorrelation function $C_{vv}(t) = \langle \hat{\mathbf{v}}_i^{t'} \cdot \hat{\mathbf{v}}_i^{t'+t} \rangle$ (see the Supplemental Material, [39], Fig. 4) and extracting an angular noise per time step by either writing $V_{vv}(1) = 0.987 = \langle \cos \delta\theta \rangle$ or using $V_{vv}(\tau) \approx 0.968$ leading to $\delta\theta \approx 9°$ and $6°$ respectively. All are similar to the observed value of η_θ at which order is maximal. Thus the realized order is maximal at a value of cognitive noise η_θ that is *self-consistent* with the variation in the realized trajectories that arises in the dynamics. We argue that this is the noise level at which the predictive model of the trajectories of other agents will be more accurate, at least in a statistical sense. We propose that this represents the fundamental reason for the increase of order at small noise levels.

To apply postdecision noise, the rotation associated with each action that appears in Eq. (3) is modified to include noise according to $\theta_1 = \theta_2 = \theta_3 = \zeta_i^t$, $\theta_4 = \Delta\theta + \zeta_i^t$, and $\theta_5 = -\Delta\theta + \zeta_i^t$ with the random rotation angle ζ_i^t drawn from a zero mean $\langle \zeta_i^t \rangle = 0$ Gaussian distribution satisfying $\langle \zeta_i^t \zeta_j^{t'} \rangle = \eta^2 \delta_{ij} \delta_{tt'}$. This noise can be interpreted as arising from imperfect implementation of the target velocity, ubiquitous in physical or living systems.

Figure 3(b) shows the effect of this postdecision orientational noise. At large noise amplitude η the order approaches a value $\sim 1/\sqrt{N}$, expected for N randomly orientated agents. This corresponds to a complete loss of orientational order. We find the order-disorder transition occurs around $\eta = 12°$. This is a significantly smaller noise level than for the case of cognitive noise [see Figs. 3(a1)–3(a3)], where the transition occurs around $\eta_\theta = 45°$. This indicates that cognitive noise has a much weaker disordering effect than postdecision noise and could even be seen as providing robustness, by anticipating the possibility of varied trajectories in the future. In contrast, postdecision noise plays no such role.

A one-tailed t-test, to test whether the order at nonzero noise values is significantly different from the zero noise case, was performed for both ϕ and ϕ_C. The result being significant for the mean order for agents in the largest cluster ϕ_C ($p < 10^{-6}$) but insignificant for the global order ϕ ($p = 0.086$). The difference between the two is likely due to rare fragmentations, which we see in large groups $\gtrsim 100$, noting also that ϕ is systematically lower than ϕ_C. The fact that there is a significant increase in ϕ_C is perhaps even more surprising than the similar effect apparent in Fig. 3(a3). The magnitude of the increase in order ϕ_C from $\eta_\theta = 0$ to $\eta_\theta \sim 5°$ that is apparent in Fig. 3(a2) is about 1% (a relatively large difference: the misordering halves). However, the corresponding increase in Fig. 3(b2) is nearly an order of magnitude weaker and occurs at much smaller noise $\eta \sim 0.5°$. This signifies a different mechanism for the much weaker increase in order that occurs under such postdecision noise. We speculate that this might be due to subtle changes in the swarm structure resulting from the addition of noise, noting that the density is systematically lower in the presence of weak postdecision noise (see the Supplemental Material [39] for details). Such changes could plausibly affect path-entropy maximizing trajectories in such a way that they generate a higher order. Although there is no obvious intuitive explanation for this it could be related to the fact that the agents have more information on the global organization at lower densities, where there are fewer particle overlaps in the visual state.

To conclude, we analyze a simple model that could underly evolutionary fitness and hence intelligent behavior. This model involves agents that seek to maximize the path entropy of their future trajectories, analogous to keeping future options open. The entropy is here computed over visual states, such as would be perceived by animals that rely primarily on vision to sense and navigate the world around them. Such path-entropy maximization strategies could be of broader interest within biology, e.g., in the biochemical state space accessible to micro-organisms or cells. However, we believe that it will be easier to test these ideas in higher animals that exhibit swarming motion where the state space is lower dimensional and the dynamics of inertial flying (or swimming) agents is much more simple and well understood than the nonlinear chemical kinetics of cellular biochemistry.

We find that the "bottom-up" principle of maximization of path entropy is a promising candidate to understand the emergence of properties like coalignment and cohesion observed in typical swarming phenotypes. This principle also leads to flocks with opacity values close to 0.5, in agreement with observations on some bird flocks [19].

Although the algorithm is highly computationally demanding it involves a simple mapping from an observed visual state to an action. Heuristics that mimic this process and that could operate under animal cognition in real time are easy to develop. For example, an artificial neural network could be trained on simulation data to choose actions from sensory input. Similar algorithms could also find use in novel forms of active, information-processing ("intelligent") matter that may soon form part of the experimental landscape.

Funding was provided by UK Engineering and Physical Sciences Research Council though the Mathematics for Real World Systems Centre for Doctoral Training Grant No. EP/L015374/1 (H. L. D.). All numerical work was carried out using the Scientific Computing Research Technology Platform of the University of Warwick. M. S. T. acknowledges the support of a long-term fellowship from the Japan Society for the Promotion of Science, a Leverhulme Trust visiting fellowship, and the peerless hospitality of Prof. Ryoichi Yamamoto (Kyoto).

*Corresponding author.
M.S.Turner@warwick.ac.uk

[1] A. Sokolov, I. S. Aranson, J. O. Kessler, and R. E. Goldstein, Concentration Dependence of the Collective Dynamics of Swimming Bacteria, Phys. Rev. Lett. **98**, 158102 (2007).
[2] T. Vicsek and A. Zafeiris, Collective motion, Phys. Rep. **517**, 71 (2012).
[3] A. Cavagna, A. Cimarelli, I. Giardina, G. Parisi, R. Santagati, F. Stefanini, and M. Viale, Scale-free correlations in starling flocks, Proc. Natl. Acad. Sci. U.S.A. **107**, 11865 (2010).
[4] A. Flack, M. Nagy, W. Fiedler, I. D. Couzin, and M. Wikelski, From local collective behavior to global migratory patterns in white storks, Science **360**, 911 (2018).
[5] D. Zitterbart, B. Wienecke, J. Butler, and B. Fabry, Coordinated movements prevent jamming in an emperor penguin huddle, PLoS One **6**, 1 (2011).
[6] M. Moussaïd, D. Helbing, and G. Theraulaz, How simple rules determine pedestrian behavior and crowd disasters, Proc. Natl. Acad. Sci. U.S.A. **108**, 6884 (2011).
[7] D. Pol, A. C. Mancuso, R. M. H. Smith, C. A. Marsicano, J. Ramezani, I. A. Cerda, A. Otero, and V. Fernandez, Earliest evidence of herd-living and age segregation amongst dinosaurs, Sci. Rep. **11**, 20023 (2021).
[8] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, Novel Type of Phase Transition in a System of Self-Driven Particles, Phys. Rev. Lett. **75**, 1226 (1995).
[9] J. Buhl, D. J. T. Sumpter, I. D. Couzin, J. J. Hale, E. Despland, E. R. Miller, and S. J. Simpson, From disorder to order in marching locusts, Science **312**, 1402 (2006).
[10] B. Szabó, G. J. Szöllösi, B. Gönci, Z. Jurányi, D. Selmeczi, and T. Vicsek, Phase transition in the collective migration of tissue cells: Experiment and model, Phys. Rev. E **74**, 061908 (2006).
[11] G. Grégoire and H. Chaté, Onset of Collective and Cohesive Motion, Phys. Rev. Lett. **92**, 025702 (2004).
[12] H. Chaté, F. Ginelli, G. Grégoire, F. Peruani, and F. Raynaud, Modeling collective motion: Variations on the Vicsek model, Eur. Phys. J. B **64**, 451 (2008).
[13] F. Ginelli and H. Chaté, Relevance of Metric-Free Interactions in Flocking Phenomena, Phys. Rev. Lett. **105**, 168103 (2010).
[14] C. A. Weber, T. Hanke, J. Deseigne, S. Léonard, O. Dauchot, E. Frey, and H. Chaté, Long-Range Ordering of Vibrated Polar Disks, Phys. Rev. Lett. **110**, 208001 (2013).
[15] D. S. Calovi, U. Lopez, S. Ngo, C. Sire, H. Chat'e, and G. Theraulaz, Swarming, schooling, milling: Phase diagram of a data-driven fish school model, New J. Phys. **16**, 015026 (2014).
[16] D. Nishiguchi, K. H. Nagai, H. Chaté, and M. Sano, Long-range nematic order and anomalous fluctuations in suspensions of swimming filamentous bacteria, Phys. Rev. E **95**, 020601(R) (2017).
[17] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic, Interaction ruling animal collective behavior depends on topological rather

than metric distance: Evidence from a field study, Proc. Natl. Acad. Sci. U.S.A. **105**, 1232 (2008).

[18] H. Hildenbrandt, C. Carere, and C. Hemelrijk, Self-organized aerial displays of thousands of starlings: A model, Behav. Ecol. **21**, 1349 (2010).

[19] D. J. G. Pearce, A. M. Miller, G. Rowlands, and M. S. Turner, Role of projection in the control of bird flocks, Proc. Natl. Acad. Sci. U.S.A. **111**, 10422 (2014).

[20] R. Bastien and P. Romanczuk, A model of collective behavior based purely on vision, Sci. Adv. **6**, eaay0792 (2020).

[21] A. C. Gallup, J. J. Hale, D. J. T. Sumpter, S. Garnier, A. Kacelnik, J. R. Krebs, and I. D. Couzin, Visual attention and the acquisition of information in human crowds, Proc. Natl. Acad. Sci. U.S.A. **109**, 7245 (2012).

[22] A. Attanasi, A. Cavagna, L. Del Castello, I. Giardina, T. S. Grigera, A. Jelić, S. Melillo, L. Parisi, O. Pohl, E. Shen *et al.*, Information transfer and behavioural inertia in starling flocks, Nat. Phys. **10**, 691 (2014).

[23] J. Toner and Y. Tu, Long-Range Order in a Two-Dimensional Dynamical XY Model: How Birds Fly Together, Phys. Rev. Lett. **75**, 4326 (1995).

[24] M. Durve, F. Peruani, and A. Celani, Learning to flock through reinforcement, Phys. Rev. E **102**, 012601 (2020).

[25] T. Nemoto, E. Fodor, M. E. Cates, R. L. Jack, and J. Tailleur, Optimizing active work: Dynamical phase transitions, collective motion, and jamming, Phys. Rev. E **99**, 022605 (2019).

[26] tienne Fodor, T. Nemoto, and S. Vaikuntanathan, Dissipation controls transport and phase transitions in active fluids: mobility, diffusion and biased ensembles, New J. Phys. **22**, 013052 (2020).

[27] Y.-E. Keta, E. Fodor, F. van Wijland, M. E. Cates, and R. L. Jack, Collective motion in large deviations of active particles, Phys. Rev. E **103**, 022603 (2021).

[28] E. Fodor, R. L. Jack, and M. E. Cates, Irreversibility and biased ensembles in active matter: Insights from stochastic thermodynamics, Annu. Rev. Condens. Matter Phys. **13**, 215 (2022).

[29] A. D. Wissner-Gross and C. E. Freer, Causal Entropic Forces, Phys. Rev. Lett. **110**, 168702 (2013).

[30] S. Mohamed and D. J. Rezende, Variational information maximisation for intrinsically motivated reinforcement learning, in *Advances in Neural Information Processing Systems* (Montreal, Canada, 2015), pp. 2116–2124.

[31] R. P. Mann and R. Garnett, The entropic basis of collective behaviour, J. R. Soc. Interface **12**, 20150037 (2015).

[32] H. J. Charlesworth and M. S. Turner, Intrinsically motivated collective motion, Proc. Natl. Acad. Sci. U.S.A. **116**, 15362 (2019).

[33] H. Hornischer, S. Herminghaus, and M. G. Mazza, Structural transition in the collective behavior of cognitive agents, Sci. Rep. **9** (2019).

[34] It is not strictly self-consistent in the sense that the model for others (ballistic motion) is different to that of self (path entropy-maximisation). It is unclear whether such models can ever be made rigorously self-consistent or whether they then represent uncomputable functions [35,36].

[35] A. M. Turing, Computability and λ-definability, J. Symb. Log. **2**, 153 (1937).

[36] K. Gödel, Über formal unentscheidbare sätze der principia mathematica und verwandter systeme i, Monatshefte Math. Phys. **38**, 173 (1931).

[37] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic, Empirical investigation of starling flocks: a benchmark study in collective animal behaviour, Animal Behaviour **76**, 201 (2008).

[38] C. K. Hemelrijk and H. Hildenbrandt, Some causes of the variable shape of flocks of birds, PLoS One **6**, 1 (2011).

[39] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.130.168201 for long time behavior, clustering methodology, variations in density with the noise models, the velocity auto correlation function, behavior across model parameters, and example movies.

[40] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, A density-based algorithm for discovering clusters in large spatial databases with noise, in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, KDD'96 (AAAI Press, Portland Oregon, 1996), p. 226231.