Author for correspondence:
Harvey L. Devereux
e-mail: h.devereux@warwick.ac.uk

THE ROYAL SOCIETY
PUBLISHING

Whirligig beetles as corralled active Brownian particles

Harvey L. Devereux[1,4], Colin R. Twomey[5], Matthew S. Turner[2,3,6] and Shashi Thutupalli[4,7]

[1]Department of Mathematics, [2]Department of Physics and [3]Centre for Complexity Science, University of Warwick, Coventry CV4 7AL, UK
[4]Simons Center for the Study of Living Machines, National Centre for Biological Sciences, Tata Institute for Fundamental Research, Bangalore 560065, India
[5]Department of Biology, and Mind Center for Outreach, Research and Education, University of Pennsylvania, Philadelphia, PA, USA
[6]Department of Chemical Engineering, Kyoto University, Kyoto 615-8510, Japan
[7]International Centre for Theoretical Sciences, Tata Institute for Fundamental Research, Bangalore 560089, India

HLD, 0000-0001-9449-7350; ST, 0000-0002-4728-937X

We study the collective dynamics of groups of whirligig beetles *Dineutus discolor* (Coleoptera: Gyrinidae) swimming freely on the surface of water. We extract individual trajectories for each beetle, including positions and orientations, and use this to discover (i) a density-dependent speed scaling like $v \sim \rho^{-\nu}$ with $\nu \approx 0.4$ over two orders of magnitude in density (ii) an inertial delay for velocity alignment of approximately 13 ms and (iii) coexisting high and low-density phases, consistent with motility-induced phase separation (MIPS). We modify a standard active Brownian particle (ABP) model to a corralled ABP (CABP) model that functions in open space by incorporating a density-dependent reorientation of the beetles, towards the cluster. We use our new model to test our hypothesis that an motility-induced phase separation (MIPS) (or a MIPS like effect) can explain the co-occurrence of high- and low-density phases we see in our data. The fitted model then successfully recovers a MIPS-like condensed phase for $N = 200$ and the absence of such a phase for smaller group sizes $N = 50, 100$.

1. Introduction

There is now an extensive body of computer simulations and theoretical work to suggest that aggregation can emerge, even when inter-particle interactions are purely repulsive, e.g. steric contact forces [1–5]. The aggregation can be thought of as arising from the competition between the accretion of free motile particles on contact with a dense cluster and their departure from the surface of that cluster following a re-orientational time lag. A form of non-equilibrium phase separation can arise in which densely and sparsely populated regions coexist. This is now known as motility-induced phase separation (MIPS). This phase separation depends on the relationship between self-propulsion speed and local density. If the speed falls off sharply enough with increasing density then a feedback loop can emerge in which slowing down (due to higher density) promotes further aggregation. High-density clusters then grow and the density in the dilute phase drops. As it does so the rate of accretion onto the clusters drops until it again comes into balance with the evaporation of particles from the cluster surface into the dilute phase. Even steric repulsion is therefore enough to cause active, self-propelled particles to accumulate in regions where they move slowly [6]. MIPS has been shown to arise in active particle systems with a more general density-dependent propulsion speed [7]. Our study provides experimental evidence justifying the use of a power law velocity-density dependence in such models.

This aggregation of motile particles in the presence of purely repulsive interactions has attracted much attention from theorists working on non-equilibrium

statistical mechanics attracted by possible insights to time-reversal symmetries and entropy production. Numerical investigations of MIPS have been carried out to examine the affects of incorporating velocity alignment terms, varying dimensionality and the effect of mixtures of both active and passive Brownian particles. These models focus on particles in finite space, e.g. with periodic boundary conditions [8–13]. However, experimental analogues are rare with the few examples including self-propelled robots, colloid systems and vibrated granular systems [10,14–16]. To our knowledge, few corresponding examples exist in living systems, although active phase separation has been seen in bacteria [17] and mussels [18].

Another emerging strand of literature has begun to focus on the role of inertia in self-propelled particle systems, leading to an equation of motion that is second order in time. The presence of inertia has lead to observations of inertial delay between particle velocity and body axis [19]. If the inertial effect is strong enough it appears that the onset of MIPS occurs at higher Péclet numbers (a dimensionless ratio of self-propulsion speed to the rate of diffusion [20]) and can vanish for large enough particle mass. Furthermore, before the onset of MIPS, a novel phase coexistence between 'hot' and 'cold' regions develops where the kinetic energy per particle (kinetic temperature) is low in the dense phase and high in the dilute phase (a difference of a factor of 100 has been predicted) [21,22]. Similarly, it has been found that the presence of inertia drastically changes the dynamics of a rotating micro swimmer [23]. Experimentally, the realization of inertia in self-propelled particle systems is seen in active granule systems such as macroscopic 'bristle bots' or 'vibro-bots' which use either a small vibrating motor or are placed on a vibrating plate with angled feet to provide self-propulsion [16,24]. Beyond this, experimental realizations of inertial delay are rare, particularly in living systems.

We study experimental footage of whirligig beetles *D. discolor* containing between $N = 50$ and 200 individuals that are moving freely on a water surface within a large circular arena. Figure 1a depicts the experimental set-up with an overlay representing the topological method we employ for local density estimation. These water beetles are ellipsoidal in nature with an aspect ratio (from a top-down perspective) measured as approximately 2:1 and a body length of approximately 12 ± 1 mm [25]. Whirligig beetles are a particularly useful study organism due to their lack of group hierarchy and strong similarity between individuals (both particularly important traits in the context of MIPS). It is also relatively simple to collect top-down two-dimensional video footage of their movement.

Previous studies have focused on their natural behaviour and include observations of large-scale clusters (rafts), which form during the day and can number from 100s to 1000s of individuals [26,27]. These structures are noted for their rapid dispersal (flash expansion events) and reformation when threatened by predators such as fish. The rapid break-up is thought to be caused by a cascading signalling process in which beetles sense (via vision or sensation of water disturbance) the movement of neighbours and react accordingly by moving rapidly and often randomly, with the onset dependent on the number of visibly startled beetles [28]. In particular, the movement is directed away from the group's geometric centre and not the point of highest density or from the location of the original beetle to startle [29]. Here, we neglect any possible role of capillary interactions between individual beetles [30], noting that these are probably less

significant for strongly self-propelled particles. Other studies have focused on the individual movement capabilities of whirligig beetles, with applications to the design of efficient 'fast' bioinspired artificial swimming robots [31–33]. Individual beetles have been previously observed moving with maximum speeds of 160 body lengths per second (in bursting events), reaching accelerations of 2.86 g, and maximum turning rates of 77.3 rad s^{-1} [34–36].

We report evidence for a density-dependent swimming speed which decays as a power law of density for different populations sizes studied. This is indicative of a marked propulsion speed difference between the dense clustered phase and dilute phases that are observed to coexist. We present a simple model, based on the motion of active Brownian particles (ABPs), that is able to capture the empirical density probability density function (PDF) observed in the data. This model, which we call corralled active Brownian particles, generates the turning of particles back towards the geometric centre of the cluster. The turning is proportional to a strength coefficient τ and a power $\alpha \geq 0$ of density. Finally, we demonstrate the presence of inertial effects in the form of a short inertial delay between the beetle's body axis and its velocity vector.

2. Results

2.1. Speed and density

Using individual beetle trajectory data we extract the speed v and density ρ averaged over particles and time. The speed here is defined in the short time limit as the particle displacement between individual frames and is calculated using a central difference method for higher order accuracy. Note that the crossover to diffusive behaviour is on much longer timescales of 10–30 frames (see electronic supplementary material, figure S6, consistent with electronic supplementary material, S5). The density is a local (topological) measure of individual beetle density, see methods. Figure 1b shows the averaged speed for particles with density ρ, written $v(\rho)$. This exhibits an empirical power law scaling across a broad regime of densities and appears to be quite consistent across different population sizes. The slope associated with an exponent of -0.4 is shown in figure 1b as a guide to the eye. At the very highest local densities, we observe a marked break from the power law to movement speeds increasing with density. We speculate that this may be associated with the coordinated motions of high-density domains in the cluster, moving as a rigid body.

2.2. Orientation–velocity correlation

As shown in figures 1c,d, we find the body axis leads the velocity by a positive lag time. The correlation function is given by

$$C(\Delta t) = \langle \hat{V}_i(t + \Delta t) \cdot \hat{n}(t) \rangle_{t,i} \qquad (2.1)$$

and measures the average scalar product between the orientation at time t and the instantaneous velocity direction at time $t + \Delta t$, where Δt is the time lag. All times are discrete in units of the video frame interval ($1/30$ s), hence the discrete points on figure 1d. The average $\langle \ldots \rangle_{t,i}$ is over time t and over the beetle trajectory index i. We only analyse trajectories that we consider to be reliably collision free, that is trajectories in which the minimum inter-particle distance over the entire trajectory is greater than a threshold of one body length, (see electronic supplementary material
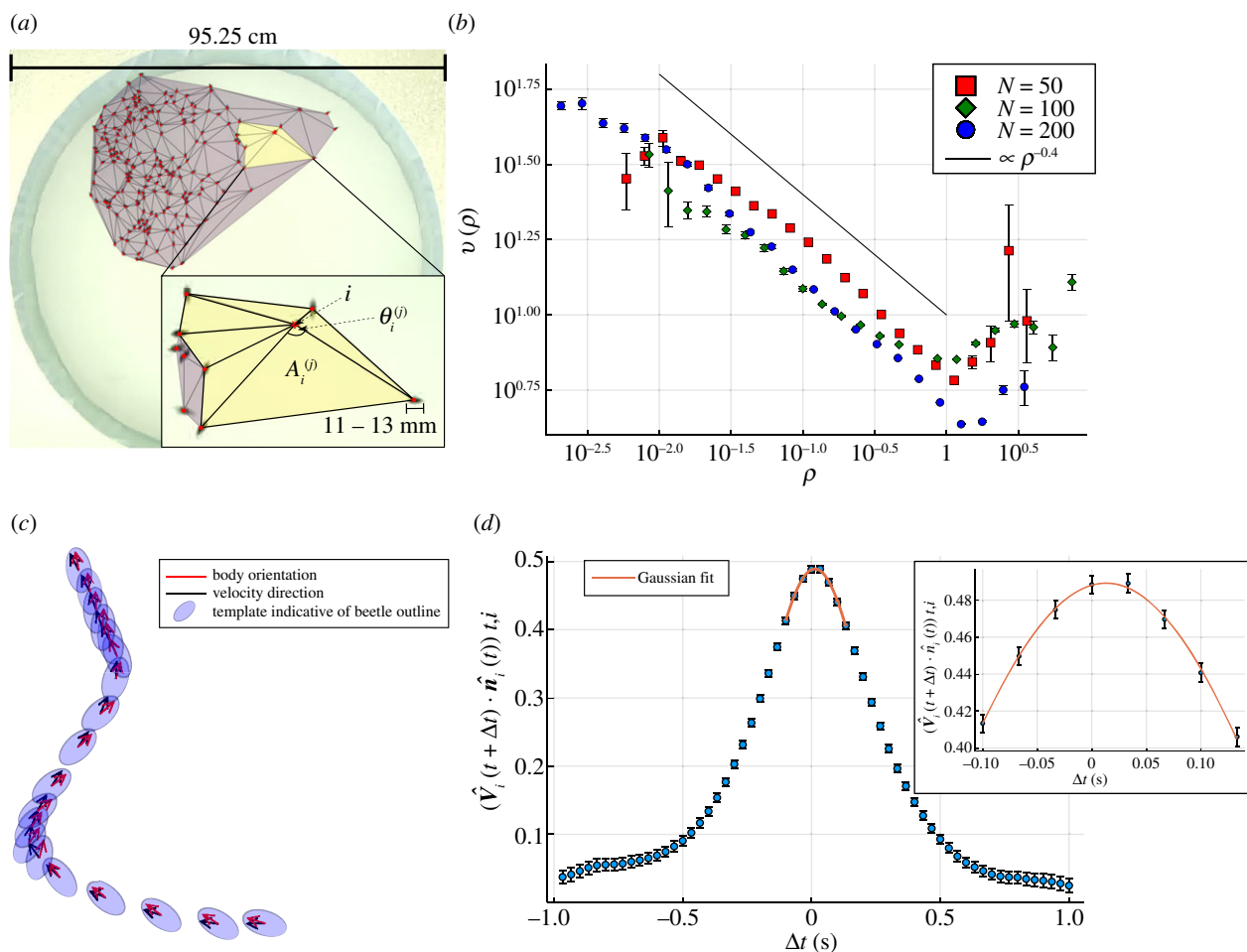
Figure 1. Beetle velocities scale with a power law of local density and exhibit a relatively short inertial delay of 13 ms. (*a*) A snapshot of the experimental set-up and an overlay detailing the method for local density calculation using the Delaunay tessellation-based method (see methods §4.3). In the overlay, the red points are particle (beetle) positions and lines indicate the Delaunay tessellation. The inset labels the angles and areas used to calculate the local density of the central *i*th beetle, shown as a red star. The yellow polygon outlined in bold indicates the union of Delaunay triangles having this beetle as a common vertex. We label this set of Delaunay triangles with the index *j* referring to each triangle as $T_i^{(j)}$, it is area is $A_i^{(j)}$, and the angle subtended at *i* as $\theta_i^{(j)}$. The local density is calculated as the *inverse* of the average weighted areas $A_i^{(j)}$, with the angles $\theta_i^{(j)}$ as weights, further normalized by a factor of $1/2\pi$ to account for the fact internal points satisfy $\sum_j \theta_i^{(j)} = 2\pi$ while points on the boundary satisfy $\sum_j \theta_i^{(j)} < 2\pi$. See methods section for further details. (*b*) The relationship between beetle speed *v* (body lengths per second) and local density ρ (per body length squared) on a log-log scale. Each data point represents a bin-average with error bars showing 1 s.d. The solid black line indicates a power law ($\rho^{-0.4}$) as a guide to the eye. (*c*) Shows the inertial delay between a beetle's velocity vector and its body axis orientation: the orientation leads the velocity. The shaded ellipses represent the moving outline of the beetle over time. This is quantified in (*d*), showing the orientation–velocity time correlation function with a Gaussian fit superimposed near the peak, located at $\Delta t = 13$ ms (see inset). Only in (*d*), we use the set of all $N = 200$ beetle trajectories pre-filtered to remove (near) collisions (see electronic supplementary material for details).

for details). The positive peak location corresponds to an inertial delay of approximately 13 ms.

2.3. Model: corralled active Brownian particles

To model the behaviour of the swarm, we develop a minimal particle-based simulation and fit this to the data. This is based on a standard ABP model (neglecting hydrodynamic interactions and inertia) within the same framework as [1] but with an additional reorientation term included in the orientational dynamics to account for the fact that our system is not contained by periodic boundaries and therefore needs a mechanism to corral the particles into the same region in space—behaviour that is clearly exhibited by the beetles themselves. To this end, we introduce a torque that tends to re-orientate particles back towards the centre of mass of the cluster (figure 2*a*).

This re-orientation, acts separately on all particles and is assigned a strength that depends on the local density.

Models that incorporate similar torques, but designed to promote co-alignment, have been used to study the affect of particle alignment on the onset of MIPS [8]. Torque terms can also arise naturally as a consequence of particle geometry [37] although we neglect this in our model. For simplicity, our model employs self-propelling polar disks of uniform radius *a* (with the diameter 2*a* considered equal to one beetle length), self-propelled along the body axis (polarity) $\hat{\boldsymbol{n}}_i = [\cos\theta_i, \sin\theta_i]^T$, of the *i*th particle. Propulsion along the $\hat{\boldsymbol{n}}_i$ direction involves a constant magnitude speed v_0. The collisions are soft-body interactions with a harmonic force on particle *i* due to *j* of $\boldsymbol{F}_{ij} = -k(2a - r_{ij})\hat{\boldsymbol{r}}_{ij}$ for $r_{ij} < 2a$, $\boldsymbol{F}_{ij} = 0$ otherwise. This involves the interparticle separation vector $\boldsymbol{r}_{ij} = \boldsymbol{r}_j - \boldsymbol{r}_i$, with $\boldsymbol{r}_i$ the position of particle *i*, its magnitude r_{ij} and where a hat (ˆ) denotes a unit vector throughout. The particles follow over-damped Langevin equations of motion given by

$$\partial_t \boldsymbol{r}_i = v_0 \hat{\boldsymbol{n}}_i + \mu \sum_{j \neq i} \boldsymbol{F}_{ij} \qquad (2.2)$$
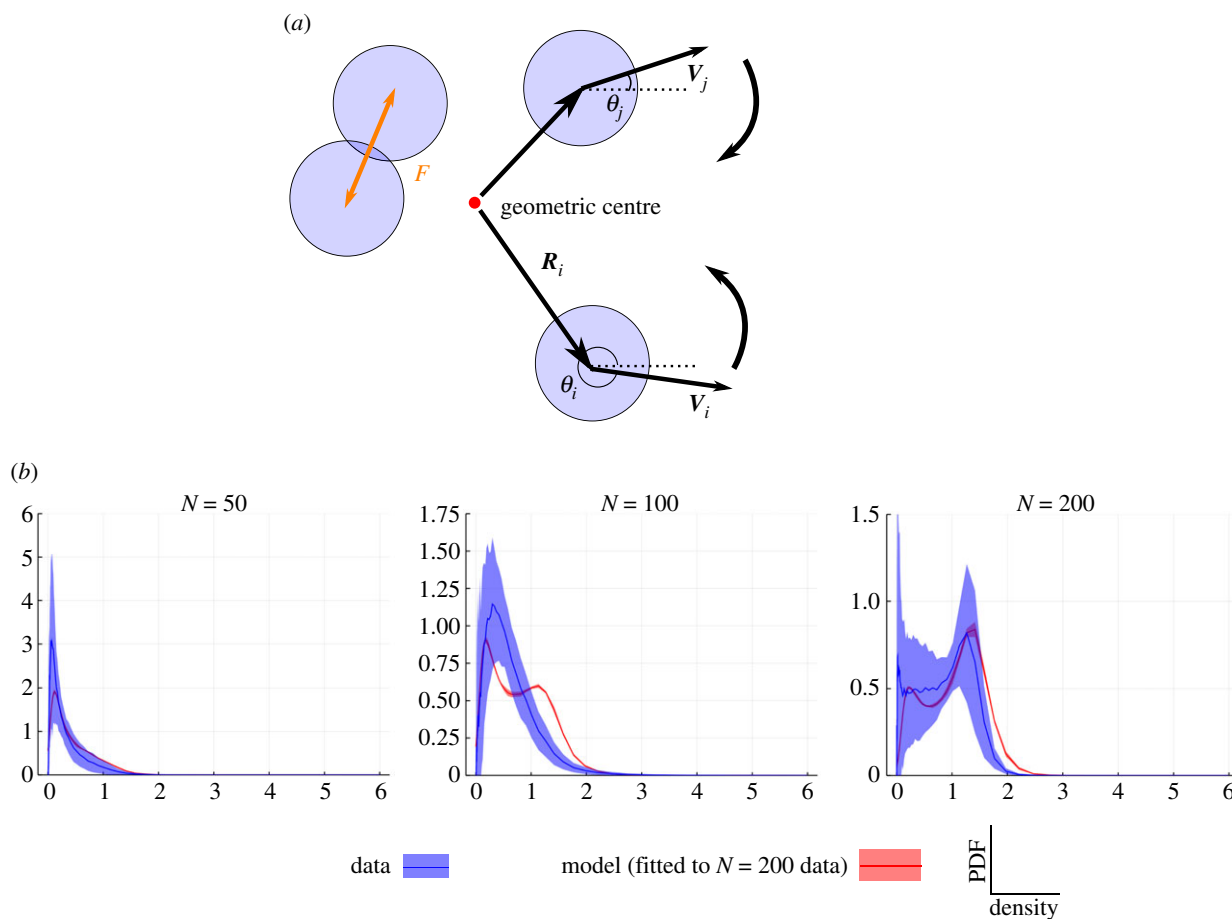
Figure 2. (a) Cartoon of the model. The particles are treated as soft, self-propelled disks that repel with a force F when overlapped, as in a standard ABP model. However, they also experience a density-dependent re-orientation towards the centre of mass of the swarm (annotated geometric centre), represented by the curved arrows. The reorientation depends on the local density and is assumed strongest at low densities. (b) The density PDFs for the experimental data (blue) and the fitted model (red), evaluated for different numbers of particles N, as shown. The model dynamics correctly recover coexisting dilute and dense phases for $N = 200$ and a dilute phase alone for $N = 50$, the phase boundary is around $N = 100$. The model is parameterized once only, by fitting to the PDF for the $N = 200$ human tracked dataset; the red curves for $N = 50, 100$ are the results of this model evaluated with different particle counts. The solid lines are the mean density distributions (kernel density estimates), the error bands indicate 1 s.d. Other simulation parameters in this and subsequent figures are $\mu k = 316.2$, $D_r = 2.34$ rad^2 s^{-1} and $v_0 = 13.19$ body lengths per second.

and

$$\partial_t \theta_i = \eta_i(t) + \tau \rho_i(t)^{-\alpha}(\hat{\boldsymbol{R}}_i \times \hat{\boldsymbol{V}}_i) \cdot \hat{\mathbf{z}}. \tag{2.3}$$

Here, μ is a mobility parameter. However, both this and the elastic constant k only appear in the product μk and so this does not introduce an additional control parameter. The angular noise $\eta_i(t)$ has zero mean and is Gaussian distributed according to $\langle \eta_i(t)\eta_j(t') \rangle = 2D_r \delta_{ij}\delta(t - t')$, where D_r is a rotational diffusion coefficient; a positional noise term can be neglected here. The re-orientation term in the angular dynamics generates reorientation towards the centre of mass with $\boldsymbol{R}_i = \langle r_j \rangle - r_i$ the vector pointing from a particle to the centre of mass, and $\boldsymbol{V}_i$ the instantaneous velocity of particle i. The dot product with $\hat{\mathbf{z}}$ (out of the plane of motion) converts this to a signed scalar, positive for anticlockwise turns and negative for clockwise. Finally, the factor $\tau \rho_i^{-\alpha}$ (with $\alpha > 0$) represents a simple choice for the density dependence of the reorientation. All simulations are carried out in unbounded 2D space, with the density that emerges within the cluster controlled by the values of τ and α. All quantities are reported as dimensionless throughout with times scaled in seconds and lengths scaled with the beetle particle length (disc diameter).

To reduce the dimensionality of the fitting process, we determine the value of the rotational diffusion coefficient by directly fitting to the mean square angular displacement of the $N = 200$ beetle data, yielding $D_r = 2.34$ rad s^{-1}. We also fix the self-propulsion speed v_0 as the average speed of beetles that are freely moving (see electronic supplementary material for a discussion of collision free trajectories), yielding $v_0 = 13.19$ body lengths per second. We fit the free parameters $(\mu k, \alpha, \tau)$ by minimizing the error, using a Bayesian optimization technique (see Methods), between the empirical density distribution (PDF) for the $N = 200$ human tracked beetle data (only) and the density distribution obtained from a simulation with $N = 200$ particles (see the methods section for details). The resultant density distributions, for *all* datasets are shown in figure 2b, together with the results from simulations of the model (fitted to $N = 200$ only, not re-parameterized for each dataset) containing the corresponding number of particles $N = 50, 100, 200$. Best-fit parameter values are shown in table 1. We include an example simulation for the fitted model evaluated with $N = 200$ particles as electronic supplementary material, movie S8. Also shown in figure 2b are the results of the best-fit model, simulated at different N values, as shown. As the number of particles increases, we observe a clear transition from a uni-modal density distribution, with a single peak at low density
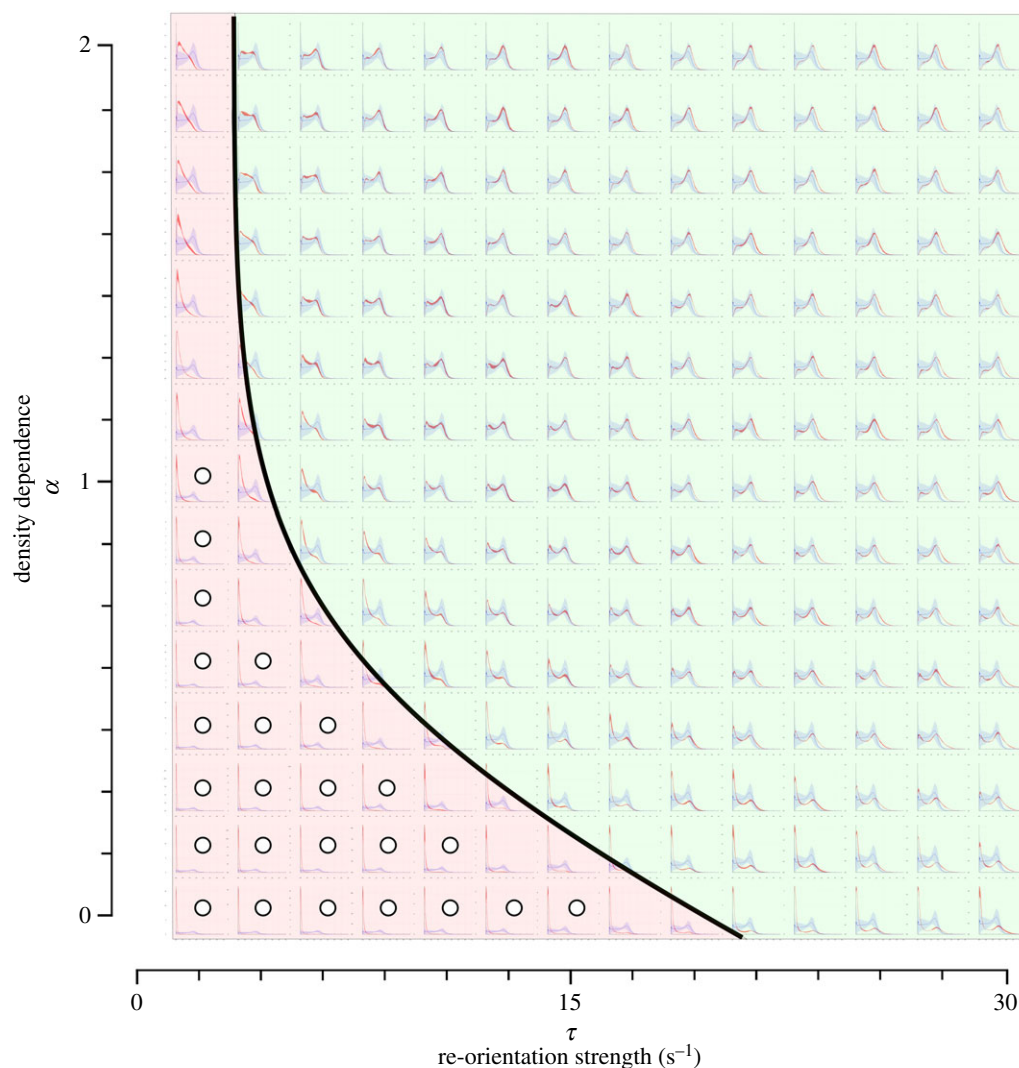
Figure 3. Phase diagram for our CABP model in $\alpha - \tau$ space, with the one phase (dilute gas) highlighted in red and the two phase (gas and dense liquid) in green. Each of the 14×15 sub-panels (the $\tau = 0$ column is excluded) shows the density PDF obtained from a dynamical simulation for $N = 200$ particles at the corresponding τ, α values (red curve, with s.d. error band). Also shown on each panel is the density PDF from the experimental data at the same N value (blue curve, with s.d. error band, the same on each panel). The best-fit parameters from table 1 correspond roughly to the system shown in row 7, column 9, deep in the two-phase region. The approximate location of the phase boundary denoting the appearance of a high-density phase is identified by eye. Circles overlaid on the plots indicate that we are not confident that the density has reached an equilibrium state for these systems within 30 s real time. This means that the gas phase density may be even smaller than shown; the phase boundary is unaffected. The equilibration time used was between 10 s and 30 s, longer times being required as we move towards the bottom left corner, where re-orientation is weak.

(similar to a dilute gas phase) to a bi-modal distribution (similar to a dilute gas coexisting with a dense liquid). A similar trend has been observed in simulations [1], where N directly controls the density and phase separation (bi-modality) appears only above a critical threshold.

2.4. Phase diagram in $\alpha - \tau$ space

Using the fitted value of μk, and the values of v_0 and D_r previously extracted directly from the data, we compute the density PDF generated by dynamical simulation of the model over a large space of reorientation functions, parameterized by different values of α and τ that control the functional form of the reorientation, as defined in equation (2.3). The results are displayed in figure 3 where each of the 210 sub-panels represents a density PDF like those shown in figure 2b. The density PDF for the $N = 200$ beetle data is shown, identically, on every sub-panel (blue). The PDF obtained by simulating each model, parameterized with different α and τ values, differ between sub-panels (red).

Table 1. Best-fit values of the control parameters. These are inferred by fitting the results of simulations performed on $N = 200$ particles, using equations of motion (2.2) and (2.3), to the $N = 200$ beetle dataset. The fit metric is a least-squares measure of the density PDF. We include v_0 and D_r in this table but note that they are fitted to data *before* using Bayesian optimization to find τ, α and μk to reduce dimensionality. We measure all lengths in units of the mean beetle body length.

parameter	α	τ (s^{-1})	μk (s^{-1})	v_0 (s^{-1})	D_r (rad^{2}s^{-1})
best-fit value	1.1	19.6	316	13.19	2.34

The prefactor to the reorientation strength τ increases across the rows, from left to right. This means that the systems represented in panels on the left of a row have an unambiguously weaker turning reorientation than those on the right. Weaker reorientation leads to a more diffuse cloud of particles and a lower mean density. This is consistent

with the one-phase region of dilute gas being located on the left-hand side of (rows of) this diagram. Moving down the columns corresponds to reorientation strength having a stronger density dependence (exponent) α. Since the dense liquid phase has a dimensionless density $1 \lesssim \rho \lesssim 2$ there is relatively little variation of the reorientation strength in the liquid phase moving down the columns but the gas phase typically has a density $\rho \ll 1$ and there is a correspondingly stronger influence on reorientation in this phase.

In figure 3, we also identify the approximate location of the phase boundary between a one-phase gas (low density, uni-modal PDF) and the two-phase gas–liquid coexistence (bi-modal PDF), see electronic supplementary material for similar phase diagrams for $N = 50$ and $N = 100$. The slow equilibration of systems with parameter values in the bottom left corner of figure 3, corresponding to weak reorientation, does not affect our conclusions: the density PDFs in these systems may still be slowly shifting to even *smaller* densities, meaning that they are even deeper in the gas phase than they appear. It is significant that, in all cases, MIPS-like phase coexistence arises for appropriate parameters τ and α, corresponding to sufficiently strong turning, or 'corralling', effect. This is sufficient to maintain high enough overall system densities in much the same way that ABP models need to sit above some critical density for MIPS to arise [1].

3. Discussion

While there is enormous contemporary interest in the physics of active particle systems, it remains unclear how broadly these ideas can be tested experimentally, especially in living analogues. We focus on three separate concepts.

Phenomenological velocity–density relations play a central role in foundational studies of MIPS [7] and can be viewed as a fundamental feature in our current understanding of the phenomenon. However, we are unaware of studies that directly analyse this relationship, with Liu *et al.* [18] a notable exception. In the present work, we report a particularly simple power-law scaling form for this relationship in whirligig beetles that seems to hold over two orders of magnitude in density. This may motivate the development of active field theories that directly encode such a power law.

Secondly, there has been significant recent interest in the role of inertial effects in collections of ABPs [19,21]. We present evidence for the existence of such inertial delay in macroscopic living systems, with inertial delays in the 10 ms regime. This inertial delay can be described as 'short' given that the distance moved in body lengths and the root mean squared angular diffusion are both only $O(10^{-1})$ on this timescale. This gives us some reassurance that non-inertial models, such as employed in §2.3, may be adequate at the semi-quantitative level.

Thirdly, we show that a condensed 'liquid-like' phase of whirligigs can coexist with a dilute 'gas-like' phase. This is highly reminiscent of MIPS [1], a phenomenon that has been widely studied but lacks experimental realizations, particularly in living analogues. In order to study this phenomenon, we developed a model of corralled ABPs that turn inwards, providing a mechanism to control the density of the swarm in open space that seems broadly plausible in its mechanism. We fit this model to experimental data and obtain results that reproduce the bimodal density PDF and MIPS-like coexistence, provided the systems are large enough. Finally, we speculate

that our model could be probed experimentally by studying a group of whirligig beetles 'doped' with robotic beetles programmed to either turn towards the geometric centre, as here, or responding to other interactions, such as nearest neighbour alignment or attraction/repulsion.

Our work provides a new way of understanding the behaviour of this insect, and the mechanism for the formation of dense clusters of individuals, in terms of the MIPS paradigm. MIPS is a phenomenon that has recently been identified in the context of non-equilibrium physics. In this literature, self-propelled particles (ABPs) with purely repulsive contact forces have been shown to phase separate, forming similar clusters. The fact that clustering occurs in the absence of any attraction is a signature of the out-of-equilibrium nature of the particles, i.e. that they are motile. By analogy, we suggest that the phenomenon of beetle clustering need not involve any direct attractive interactions. This was far from obvious at the outset. We have shown that the beetle's behaviour is consistent with a corralled ABP (CABP) model that we develop. This reproduces MIPS-like clustering. We do not view the corralling (turning) in our CABP model as a form of attraction but believe that it is better viewed as providing a global constraint on the density, necessary in unbounded space. This is because it involves no pairwise attractive interactions. Also, in classical ABP systems, density is regulated by the walls of a box (or its periodicity), limiting its volume, but one would not say this confinement provides an 'attraction', rather that it serves to fix the density. The identification of a model insect system that exhibits MIPS-like clustering is also likely to be of keen interest to physicists, both as a rare example in multicellular organisms but also for what it tells us about empirical velocity–density relationships in such systems.

4. Methods

4.1. Experimental data

Our raw data consist of footage from experiments on varying population sizes (50, 100, 200) of whirligig beetles (*D. discolor* collected from the Raquette River in Potsdam, NY, USA) in a cylindrical tank of water with a diameter of 1 m and individual beetles measuring 12 ± 1 mm in length along the major axis. Each video was filmed for a period of around 5 min at a resolution of 1920×1080 pixels and a frame rate of 30 frames per second (see electronic supplementary material, movie S7 for a high contrast example at $N = 200$). The camera was situated 1.96 m above the water level, a fluorescent light illuminated the apparatus at 730 lx. Beetles were not startled or given time-varying external stimulus during filming and were allowed 20 min acclimatization time before filming. Beetles were fed at 07.00 and 19.00 (at the time of filming sunrise and sunset where measured at 05.30–06.00 and 20.30–21.00 each day) and experiments performed between these times. For the $N = 200$ population, footage was tracked by hand, and for each population footage was tracked using a modified version of the network flow formulation (see an outline in our methods, §4.2) which was validated by reference to the human tracked set of data. These highly accurate tracks for individual beetles represent their coordinates (in the laboratory frame) at each time point, together with their orientations, velocity, and density which we use in our analysis of the dynamics of the collective. In the experimental footage, we observe a pronounced change in behaviour across the three swarm sizes. The footage covering the 50 beetles population is more erratic in character with mostly short-lived cluster formation and a generally elevated propulsion speed (see electronic supplementary material, figure S2

for speed distributions). The $N = 100$ and 200 beetles swarm around relatively stable clusters. In these systems, a fraction of individuals reside in a more dilute 'corona' around the main cluster, corresponding to the dense and dilute phases, respectively. The dense regions are notable for significantly decreased self-propulsion speed.

4.2. Tracking

The individual beetle tracks were generated from centre of mass coordinates and (major axis) orientations for each beetle using a modification of the network flow method for multi-object tracking [38]. Our method [39] takes advantage of the fact we can assume conservation of beetle population, whereas Li Zhang *et al.* track pedestrian data in which many individuals leave and enter the video frame. The coordinates and orientations were gathered from raw video footage by training a neural network on a set of validation footage which was tracked by hand (logging positions orientations and linking beetle tracks). Human marked tracks ($N = 200$ data) were also used to validate the tracking method, and to fit the model.

4.3. Density

To calculate the local density at a beetle's centre of mass, we use a method based upon the Delaunay triangulation to assign an area and therefore a number density to each beetle. Delaunay triangulation-based methods of interpolating density fields from point data have been used successfully in astronomy [40], and the estimation of group density has been computed using alpha shapes to calculate an area fraction [41], which is closely related to the Delaunay triangulation.

Our method identifies the number density $\rho(\mathbf{r}_i)$ of a point $\mathbf{r}_i$ as equation (4.1). For notation, we write the Delaunay triangles with common vertex, $\mathbf{r}_i$, as $T_i^{(j)}$ so that j is an index over this set of triangles. Note that points on the edge of the convex hull, will generally be associated with fewer Delaunay triangles since the Delaunay triangulation only tessellates the convex hull. The area of $T_i^{(j)}$ is written as $A_i^{(j)}$, and similarly the angle made at point i by the two edges of the triangle $T_i^{(j)}$ ending at i is $\theta_i^{(j)}$. Refer to figure 1a for a visual representation of these quantities on an example Delaunay tessellation computed from data. Using this notation, the density at point i is equation (4.1)

$$\rho(\mathbf{r}_i) = \frac{1}{2} \frac{\sum_j \theta_i^{(j)}}{\sum_j \theta_i^{(j)} A_i^{(j)}}. \tag{4.1}$$

To derive equation (4.1), we consider particle i as contributing $\theta_i^{(j)}$ of it is 'mass' to the area $A_i^{(j)}$, this gives a normalized 'area per particle' of $\frac{1}{\pi} A_i^{(j)}$ for triangle $T_i^{(j)}$. Then to calculate a normalized density, we take the average (weighted by the angles $\theta_i^{(j)}$) of this normalized area per particle and invert it yielding $\pi \sum_j \theta_i^{(j)} / (\sum_j \theta_i^{(j)} A_i^{(j)})$. Since a point in the interior will satisfy $\sum_j \theta_i^{(j)} = 2\pi$ and the boundary points will satisfy $\sum_j \theta_i^{(j)} < 2\pi$ we divide this expression by 2π to arrive at equation (4.1).

The advantage of this approach, and our key reason for taking it, is that it avoids dividing by areas individually. This can lead to arbitrarily large local densities in the case of one triangle having an extremely small area (sliver triangles). These

small triangles can form when three points in a Delaunay triangulation are arbitrarily close to being collinear.

4.4. Model fitting

Our model parameters were tuned using a Bayesian optimization framework [42]. The quality of fit is reported by the mean square error between the simulation density distribution and the empirical distribution obtained from the data. The fitting itself employs a Gaussian kernel density estimator on the experimental and simulated density distributions with bandwidth parameter chosen according to Silverman's rule [43]. To calculate the fitting error, the mean square error between the two kernel density estimates was computed over a the range of the experimental data. In all our simulations, we take a time step of $1/900$ s. We discard the initial 4500 time steps to eliminate short-lived initial transients in the global density. Initial conditions for the simulations are randomly drawn from uniform distributions, both for position and orientation, with positions restricted to an square region of size corresponding to a density $\rho_{\text{init}} = 0.5$. Results are reported as averages over three simulations with different random initial conditions, each with the same set of fitted parameters.

4.4.1. Optimizing runtime

Running numerous simulations of the CABP model at $N = 200$ as is required for the fitting process is computationally expensive. In order to most efficiently deploy computation resources, we take advantage of large scale parallelization on GPUs. We implement the solver for the ABP model as a GPU (CUDA [44]) algorithm, parallelizing across particle index.

References

1. Fily Y, Marchetti MC. 2012 Athermal phase separation of self-propelled particles with no alignment. *Phys. Rev. Lett.* **108**, 235702. (doi:10.1103/PhysRevLett.108.235702)

2. Palacci J, Sacanna S, Steinberg AP, Pine DJ, Chaikin PM. 2013 Living crystals of light-activated colloidal surfers. *Science* **339**, 936–940. (doi:10.1126/science.1230020)

3. Buttinoni I, Bialké J, Kümmel F, Lüwen H, Bechinger C, Speck T. 2013 Dynamical clustering and phase separation in suspensions of self-propelled colloidal particles. *Phys. Rev. Lett.* **110**, 238301. (doi:10.1103/PhysRevLett.110.238301)

4. Redner GS, Baskaran A, Hagan MF. 2013 Reentrant phase behavior in active colloids with attraction.

Phys. Rev. E **88**, 012305. (doi:10.1103/PhysRevE.88. 012305)

5. Tailleur J, Cates ME. 2008 Statistical mechanics of interacting run-and-tumble bacteria. *Phys. Rev. Lett.* **100**, 218103. (doi:10.1103/PhysRevLett.100.218103)

6. Schnitzer MJ. 1993 Theory of continuum random walks and application to chemotaxis. *Phys. Rev. E* **48**, 2553–2568. (doi:10.1103/PhysRevE.48.2553)

7. Cates ME, Tailleur J. 2015 Motility-induced phase separation. *Annu. Rev. Condensed Matter Phys.* **6**, 219–244. (doi:10.1146/annurev-conmatphys-031214-014710)

8. Sesé-Sansa E, Pagonabarraga I, Levis D. 2018 Velocity alignment promotes motility-induced phase separation. *EPL (Europhysics Letters)* **124**, 30004. (doi:10.1209/0295-5075/124/30004)

9. Wysocki A, Winkler RG, Gompper G. 2014 Cooperative motion of active Brownian spheres in three-dimensional dense suspensions. *EPL (Europhysics Letters)* **105**, 48004. (doi:10.1209/0295-5075/105/48004)

10. Mognetti BM, Šarić A, Angioletti-Uberti S, Cacciuto A, Valeriani C, Frenkel D. 2013 Living clusters and crystals from low-density suspensions of active colloids. *Phys. Rev. Lett.* **111**, 245702. (doi:10.1103/PhysRevLett.111.245702)

11. Stenhammar J, Marenduzzo D, Allen RJ, Cates ME. 2014 Phase behaviour of active Brownian particles: the role of dimensionality. *Soft Matter* **10**, 1489–1499. (doi:10.1039/C3SM52813H)

12. Stenhammar J, Wittkowski R, Marenduzzo D, Cates ME. 2015 Activity-induced phase separation and self-assembly in mixtures of active and passive particles. *Phys. Rev. Lett.* **114**, 018301. (doi:10. 1103/PhysRevLett.114.018301)

13. Kudrolli A, Lumay G, Volfson D, Tsimring LS. 2008 Swarming and swirling in self-propelled polar granular rods. *Phys. Rev. Lett.* **100**, 058001. (doi:10. 1103/PhysRevLett.100.058001)

14. Narayan V, Ramaswamy S, Menon N. 2007 Long-lived giant number fluctuations in a swarming granular nematic. *Science* **317**, 105–108. (doi:10. 1126/science.1140414)

15. Deseigne J, Dauchot O, Chaté H. 2010 Collective motion of vibrated polar disks. *Phys. Rev. Lett.* **105**, 098001. (doi:10.1103/PhysRevLett.105.098001)

16. Giomi L, Hawley-Weld N, Mahadevan L. 2013 Swarming, swirling and stasis in sequestered bristle-bots. *Proc. R. Soc. A* **469**, 20120637. (doi:10. 1098/rspa.2012.0637)

17. Liu G, Patch A, Bahar F, Yllanes D, Welch RD, Marchetti MC, Thutupalli S, Shaevitz JW. 2019 Self-driven phase transitions drive *Myxococcus xanthus* fruiting body formation. *Phys. Rev. Lett.* **122**, 248102. (doi:10.1103/PhysRevLett.122.248102)

18. Liu Q-X, Doelman A, Rottschäfer V, de Jager M, Herman PMJ, Rietkerk M, van de Koppel J. 2013 Phase separation explains a new class of self-organized spatial patterns in ecological systems. *Proc. Natl Acad. Sci. USA* **110**, 11 905–11 910. (doi:10.1073/pnas.1222339110)

19. Scholz C, Jahanshahi S, Ldov A, Löwen H. 2018 Inertial delay of self-propelled particles. *Nat. Commun.* **9**, 5156. (doi:10.1038/s41467-018-07596-x)

20. Bechinger C, Di Leonardo R, Löwen H, Reichhardt C, Volpe G, Volpe G. 2016 Active particles in complex and crowded environments. *Rev. Mod. Phys.* **88**, 045006. (doi:10.1103/RevModPhys.88.045006)

21. Löwen H. 2020 Inertial effects of self-propelled particles: from active Brownian to active Langevin motion. *J. Chem. Phys.* **152**, 040901. (doi:10.1063/ 1.5134455)

22. Mandal S, Liebchen B, Löwen H. 2019 Motility-induced temperature difference in coexisting phases. *Phys. Rev. Lett.* **123**, 228001. (doi:10.1103/ PhysRevLett.123.228001)

23. Zheng Y, Löwen H. 2020 Rototaxis: localization of active motion under rotation. *Phys. Rev. Res.* **2**, 023079. (doi:10.1103/PhysRevResearch.2.023079)

24. Deblais A, Barois T, Guerin T, Delville PH, Vaudaine R, Lintuvuori JS, Boudet JF, Baret JC, Kellay H. 2018 Boundaries control collective dynamics of inertial self-propelled robots. *Phys. Rev. Lett.* **120**, 188002. (doi:10.1103/PhysRevLett.120.188002)

25. Ferkinhoff WD, Gundersen RW. 1983 *A key to the whirligig beetles of Minnesota and adjacent states and Canadian provinces (Coleoptera: Gyrinidae)*. St Paul, MN: Science Museum of Minnesota.

26. Heinrich B, Daniel Vogt F. 1980 Aggregation and foraging behavior of whirligig beetles (Gyrinidae). *Behav. Ecol. Sociobiol.* **7**, 179–186. (doi:10.1007/ BF00299362)

27. Vulinec K, Miller MC. 1989 Aggregation and predator avoidance in whirligig beetles (Coleoptera: Gyrinidae). *J. New York Entomol. Soc.* **97**, 438–447.

28. Romey WL, Lamb AR. 2015 Flash expansion threshold in whirligig swarms. *PLoS ONE* **10**, 1–12. (doi:10.1371/journal.pone.0136467)

29. Romey WL, Smith AL, Buhl J. 2014 Flash expansion and the repulsive herd. *Anim. Behav.* **110**, 171–178. (doi:10.1016/j.anbehav.2015.09.017)

30. Voise J, Schindler M, Casas J, Raphaël E. 2001 Capillary-based static self-assembly in higher organisms. *J. R. Soc. Interface* **8**, 1357–1366. (doi:10.1098/rsif.2010.0681)

31. Philamore H, Rossiter J, Stinchcombe A, Ieropoulos I. 2015 Row-bot: an energetically autonomous artificial water boatman. In *2015 IEEE/RSJ Int.Conf. on Intelligent Robots and Systems (IROS)*, pp. 3888–3893. IEEE.

32. Jia X, Chen Z, Riedel A, Si T, Hamel WR, Zhang M. 2015 Energy-efficient surface propulsion inspired by whirligig beetles. *IEEE Trans. Rob.* **31**, 1432–1443. (doi:10.1109/TRO.2015.2493501)

33. Kwak B, Bae J. 2017 Design of hair-like appendages and comparative analysis on their coordination toward steady and efficient swimming. *Bioinspir. Biomim.* **12**, 036014. (doi:10.1088/1748-3190/ aa6c7a)

34. Fish FE, Nicastro AJ. 2003 Aquatic turning performance by the whirligig beetle: constraints on maneuverability by a rigid biological system. *J. Exp. Biol.* **206**, 1649–1656. (doi:10.1242/jeb.00305)

35. Xu Z, Lenaghan SC, Reese BE, Jia X, Zhang M. 2012 Experimental studies and dynamics modeling analysis of the swimming and diving of whirligig beetles (Coleoptera: Gyrinidae). *PLoS Comput. Biol.* **8**, e1002792. (doi:10.1371/journal.pcbi.1002792)

36. Voise J, Casas J. 2009 The management of fluid and wave resistances by whirligig beetles. *J. R. Soc. Interface R. Soc.* **7**, 343–52. (doi:10.1098/rsif.2009. 0210)

37. van Damme R, Rodenburg J, van Roij R, Dijkstra M. 2019 Interparticle torques suppress motility-induced phase separation for rodlike particles. *J. Chem. Phys.* **150**, 164501. (doi:10.1063/1.5086733)

38. Zhang L, Li Y, Nevatia R. 2008 Global data association for multi-object tracking using network flows. In *2008 IEEE Conf. on Computer Vision and Pattern Recognition, 24–26 June, Anchorage, Alaska*, pp. 1–8. New York, NY: IEEE.

39. Kumar S, Nandakishore P, Thutupalli S. In preparation. Detecting and tracking high density homogeneous objects in low framerate videos.

40. Schaap WE. 2007 The delaunay tessellation field estimator. PhD thesis, University of Groningen, Groningen, The Netherlands.

41. Sosna MMG, Twomey CR, Bak-Coleman J, Poel W, Daniels BC, Romanczuk P, Couzin ID. 2019 Individual and collective encoding of risk in animal groups. *Proc. Natl Acad. Sci. USA* **116**, 20 556–20 561. (doi:10.1073/pnas.1905585116)

42. Nogueira F. 2014 Bayesian optimization: open source constrained global optimization tool for Python. See https://github.com/fmfn/ BayesianOptimization.

43. Silverman BW. 1986 *Density estimation for statistics and data analysis*, vol 26. London, UK: Chapman and Hall.

44. Nickolls J, Buck I, Garland M, Skadron K. 2008 Scalable parallel programming with CUDA. *Queue* **6**, 4053. (doi:10.1145/1365490.1365500)