The statistical mechanics of schooling fish captures their interactions

Maksym Romenskyy,1 James E. Herbert-Read,1 Ashley J. W. Ward,2 and David J. T. Sumpter1

1Department of Mathematics, Uppsala University, Box 480, Uppsala 75106, Sweden
2School of Biological Sciences, University of Sydney, Sydney, New South Wales, Australia
(Dated: September 1, 2015)

While a rich variety of self-propelled particle models propose to explain the collective motion of fish and other animals, rigorous statistical comparison between models and data remains a challenge. Plausible models should be flexible enough to capture changes in the collective behaviour of animal groups at their different developmental stages and group sizes. Here we analyse the statistical properties of schooling fish (Pseudomugil signifer) through a combination of experiments and simulations. We make novel use of a Boltzmann inversion method, usually applied in molecular dynamics, to identify the potential of the mean force of fish interactions. Specifically, we show that larger fish have a larger repulsion zone, but stronger attraction, resulting in greater alignment in their collective motion. We model the collective dynamics of schools using a self-propelled particle model, modified to include varying particle speed and a local repulsion rule. We demonstrate that the statistical properties of the fish schools are reproduced by our model, thereby capturing a number of features of the behaviour and development of schooling fish.

PACS numbers: 87.18.-h, 05.65.+b, 05.10.-a, 05.40.-a

In sufficiently large collective systems, the behaviour of an individual can be dominated by the generic statistical effects of many individuals interacting, rather than its own behaviour [1]. Much of the progress in understanding collective motion of animal groups has involved applying ideas borrowed from the statistical physics of materials like magnets or fluids [2–6]. For example, changes in group densities produce phase transitions at critical group sizes [7, 8]. More complex collective states, such as swarm, mills and polarised groups depend on the density of a group and the noise within the system [9]. Recent studies of starlings and midges have looked at spatial velocity fluctuations [10, 11], long range correlations [12], and diffusive [13] and entropic characteristics of flocks [14]. Other experiments with artificial particles have looked for similarities and differences between self-organised living matter and thermal equilibrium systems [15]. These latter approaches gather statistical information about self-organising structures in order to parameterise models (see for example the maximum entropy approach [11, 16]). However, none of these have explicitly solved the inverse problem of using the macro-level properties of animal groups to find out how the individuals within them interact.

This inference problem is essentially a statistical physics problem. The last few decades have seen a major increase in research at the interface of molecular dynamics and biophysics. In soft matter systems, estimating the potential energy of an interaction and the corresponding potentials is of particular importance, as the strength of intermolecular interactions determines the state of matter and many of its properties [17]. At the same time, molecular interaction potentials are difficult to measure experimentally and hard to compute from first principles. An alternative approach, therefore, is to estimate them from experimentally determined structures of molecules. The interactions in these structures are usually strongly coupled and assemblies are typically driven by weak forces (e.g., hydrophobicity or entropy) [18]. Therefore, estimation of these potentials requires application of sophisticated coarse-grained techniques such as reverse Monte-Carlo [19], inverse Monte Carlo [20] or iterative Boltzmann inversion [21]. These methods adjust the force field iteratively, until the distribution functions of the reference system are reproduced as accurately as possible. In other cases, when the potentials are uncoupled or weakly coupled, a more straightforward direct Boltzmann inversion approach can be applied, which approximates the potential by the negative logarithm of the radial distribution function [22]. In collectively moving animal groups the interactions between members are usually assumed to be of hierarchical structure, with repulsion having highest priority at small distances [23]. Thus one can expect that the latter method can be applied to some animal self-organised systems to infer the interactions within these groups from experimental data.

Here we investigate the schooling behaviour of fish using Boltzmann inversion and related methods. Unlike molecules and physical particles, fish change their behaviour as they go through various developmental stages. For example, onset of schooling is only possible when the central nervous system of fish is sufficiently developed to support a high level of coordination of visual and mechanosensory information [24]. Therefore, fish of different sizes can not be considered simply as particles of different physical size, since their behaviour changes with their size. We thus expect the statistical properties of the group, and of individuals, to change both with the density of fish and their developmental stage.

We used groups of 10 to 60 Pacific blue-eyes (Pseudo-
mugil signifier) with approximately three different body lengths (from here and thereafter referred as fish sizes): \( \sim 7.5 \text{ mm} \) (small), \( \sim 13 \text{ mm} \) (medium) and \( \sim 23 \text{ mm} \) (large) (see Supplemental Material [25]). The fish were confined into a large shallow circular arena (760 mm diameter) and filmed from above at high spatial and temporal resolution. The positions of fish were subsequently tracked using DIDSON tracking software [26]. On average on 80% of time we recorded 93% of fish indicating relatively high levels of tracking accuracy as compared to other studies [9].

We first investigated the spatial distribution of fish in the arena. Small fish were distributed over the arena (Fig. 1(a)) with limited collective motion. For example, 10 small fish tended to form a dispersed group, where most of the fish moved very little (Fig. 1(b)). Larger groups of 60 small fish showed slightly more collective motion, but not all fish moved in the same direction at the same time (Fig. 1(c)). In contrast, even small groups of large fish showed highly aligned collective motion (Fig. 1(d)). We calculated the average area, \( A \), covered by the group using a convex hull algorithm (see Supplemental Material [25]). The average value of \( A \) is plotted for different group and fish sizes (Fig. 2(a)). For all three fish sizes, larger groups occupied a larger area. The density, \( \rho = N/A \), also increased with the number of fish in the group (Fig. 2(b)), suggesting that the fish pack closer together in larger groups. Figure 2(a) indicates that groups of small fish occupied a larger area than the groups of medium-size or big fish. This finding supports the results of the spatial analysis (Fig. 1); small fish were more dispersed over the arena. As a result, groups consisting of small fish were less dense (Fig. 2(b)) than groups of larger fish.

To better quantify the spatial arrangement of groups, we measured their packing fraction \( a \) (Fig. 2(c)). This is the ratio between the total body area of all fish in a group (\( A_f = \sum_{i=1}^{N} A_i \)) and the global area of a group (\( A \)): \( a = A_f/A \), where \( A_i \) is the body area of individual fish.

For all body sizes of fish, packing fraction increased with group size. Groups of smaller fish had the lowest packing fraction ranging between 0.001 and 0.004 for groups of 10 and 60 individuals respectively. In contrast, groups of medium-size and large fish had higher packing fractions of \( a > 0.043 \) and \( a > 0.054 \) respectively. The lowest packing fractions in groups of small fish are comparable to those observed in bird flocks [27, 28], whilst the larger packing fractions approach those of some bacteria [29].

In physical systems, small values of \( a \) typically correspond to gases, while larger values (\( a > 0.4 \)) to liquids or crystals [30]. All packing fractions observed in our experiments, therefore, are compatible with an atomistic system in its gaseous state. At the same time, the observed range of packing fraction is wide (\( a \approx 0.001 - 0.054 \)), allowing us to presume a statistically different behaviour between different group and fish sizes in this system.

We next characterised ordering in our system using the polar order parameter [31] \( \varphi = \langle \sum_{i=1}^{N} \exp(i \theta_i) \rangle \), where \( i \) is the imaginary unit, \( \theta_i \) is the direction of motion of individual fish, and \( \langle \cdot \rangle \) denotes the time average. Values of \( \varphi \) are plotted in Fig. 2(d) as a function of group and fish size. Small fish (7.5mm) did not display much...
ordering for any group size with $\varphi \approx 0.4 - 0.55$, confirming previous observations (Figs. 1(b),(c)). Groups of medium-size fish were the most ordered, with values of the order parameter ranging between 0.73 and 0.84 depending on group size. Large fish displayed slightly lower order than medium-size fish.

We then investigated the nature of the interactions between the fish using statistical mechanics. We started by looking at the pair distribution function (PDF) \cite{32} which allowed us to study how the local density around each fish varied with respect to the average density in the system. It is defined by

$$g(r) = \frac{1}{S(r)} \frac{1}{N(r)} \left( \sum_{i=1}^{N} \sum_{j \neq i}^{N} \delta(r - |r_{ij}|) \right),$$

where $\delta$ is the Dirac delta function, $S$ is the surface area of a shell, $N$ is the number of fish inside a shell and $\langle \cdot \rangle$ stands for the time average (for details on the calculation of PDF in confinement see Supplemental Material \cite{25}).

A set of PDF-curves $g(r)$ for various fish sizes is presented in Fig. 3(a). Small fish tended to form aggregations with densities up to 4 times above the average density of the system and with a maximal half-radial width of more than 25 fish body lengths. For medium and large fish, the maximum density in a cluster was 8 times larger than the average in the system. The size of the aggregation of medium and big fish was as large as 17 and 10 body lengths, respectively. Another notable difference between the three curves is the location of the local density peak.

For small fish, the peak is at 24.5 mm, whereas for larger fish it is significantly shifted towards 30 - 40 mm. Figure 3(d) shows PDF plots for medium-sized fish at varying group size. For the smallest groups of 10 fish, the maximum density observed is 25 times above the system’s average density. The peak value of $g(r)$ decreases with increasing group size (from 25 for groups of 10 fish down to 6.5 for the largest groups of 60 fish) whilst the position of the maximum remains unchanged at approximately 30 mm for all group sizes. The maximal half-radial width of the aggregation increases with group size from 145 mm for groups of 10 fish to 240 mm for groups of 60 fish.

The differences in the pair distribution function suggest there is large variation in the underlying pair potential. In our system, in the absence of any external field, this potential is the potential of the mean force of the interactions between fish. Studies of active matter show that the resulting steady states of such systems often satisfy Boltzmann distributions \cite{33}, even given that these active systems are essentially out of equilibrium. Here we apply the opposite route: we start with the assumption that the steady-state configurations observed in the experiment are drawn from the Boltzmann distribution

$$P(r) = Z^{-1} \exp[-\beta U(r)],$$

where $\beta = 1/k_B T = 1$ with a Boltzmann constant $k_B$ and the system’s temperature denoted by $T$; $Z = \int \exp[-\beta U(r)] dr$ is the partition function and $U(r)$ is the potential of the interaction. The fact that we observe large differences in the stationary distributions of different sized fish and different group sizes is related to the changes in motion of individuals when the configuration of neighbours changes. This results in the entropy increasing, producing a number of unique spatial states. Nevertheless, the average number of individuals within each shell of the pair distribution function remains constant over the whole duration of an experiment, defining a steady-state in our system.

To derive the potential of the mean force of the interactions we use the direct Boltzmann inversion \cite{22}

$$U(r) = -k_B T \ln g(r).$$

Figures 3(b),(e) present the potential energy profiles for different sized fish and different group sizes, respectively. Note that all the curves on both figures have practically the same slope for the decreasing part of the potential down to $U(r) = -0.55$. The minimum of these curves occurs at a greater distance for larger fish (Fig. 3(b)), indicating that larger fish have a larger repulsion zone. The increasing portions of the curves have similar slopes only for medium-sized and large fish (Fig. 3(b)). This suggests that fish of different body sizes and in groups of various size have similar repulsion strength (or collision avoidance potential) at short distances, but different attraction strengths towards neighbours. To get a conclusive picture of the variation of the interaction strength over the separation distance between the individuals, we calculate the mean force of the interaction $F(r)$:

$$F(r) = \frac{d}{dr} U(r).$$

In differently-sized fish (Fig. 3(c)), the conservative force (positive and short portion of the $F(r)$ curves) is stronger than the dissipative one (negative and long portion of the $F(r)$ curves). In other words, repulsion is independent of body size and spans a much shorter distance than attraction. Constant attraction force at large distances in our system arises because fish can perceive conspecifics over the entire arena. In other systems, a change of the attraction force over distance could be used as an alternative to some approaches \cite{34} for identifying topological interactions between individuals. For example, an abrupt change in the strength of dissipative force would correspond to metric-type interactions.

To validate the fish interactions established from the experiments, we simulated the collective motion of the fish using a two-dimensional metric self-propelled particle model that accounts for variable fish speed and geometrical confinement (see Supplemental Material \cite{25}). The fish are represented by point particles interacting with each other through two mechanisms: short range repulsion (collision avoidance in \cite{23}) or velocity alignment (as
in [31]) with repulsion rule having an absolute priority in the model [32]. At every time step, the motion of the particles is subjected to a Gaussian noise and speed adjustment in accordance with a local orientational order inside the zone of alignment (see Supplemental Material [25]) so that the speed takes a maximum (preset) value when the particles’ velocities are perfectly aligned. Wall avoidance is modelled through rotation of a particle velocity with a turning rate [35] dependent on the distance to the wall and current heading of an agent (see Supplemental Material [25]). Fish of different size are modelled by scaling the size of the alignment zone with a factor $k$ proportionally to experimentally measured differences in body lengths so that $k = 1$, $k = 1.73$ and $k = 3.07$ correspond to small, medium and large fish, respectively.

Figures 4(a)-(c) present plots of PDF, interaction potential and mean force of the interaction for the three different sizes of simulated fish. Overall these qualitatively match the experimental data. Medium ($k = 1.73$) and large ($k = 3.07$) simulated fish form more dense groups than the small simulated fish ($k = 1$) with a density of 12.5, 14 and 10 times above the average, respectively (Fig. 4(a)). The density peak is also observed at larger distances for bigger simulated fish. The repulsive force is practically the same for all three cases at any given distance (Fig. 4(c)) as the radii of the repulsion zone are constant for all fish sizes. The attractive portion of the $F(r)$-curves has a complex shape as in the experiment indicating strongest attraction at $r \approx 28$, $r \approx 90$ and $r \approx 200$ mm for small, medium and large simulated fish, respectively. For all five group sizes of the medium sized simulated fish ($k = 1.73$) the maximum of the local density is well above the system’s average (Fig. 4(d)) and is largest for groups of 10 fish ($g(r) = 52$). All the curves
cross the $g(r) = 1$ line in the same order as in the experiment: The average half-radial width of groups of 10 and 60 simulated fish are $\sim 100$ mm and $\sim 210$ mm, respectively. As in the experiment, the repulsive force (Fig. 4(f)) decays with a distance and takes very similar values for all five cases. The attraction in groups of 10 and 20 individuals is maximal at $r = 90$ and $r = 115$, respectively, then decreases steeply with a distance and reaches zero at $r \approx 190$. For the other three curves (N=30, 40 and 60), the attractive force has a maximum at $r \approx 100$. At all intermediate distances the dissipation is stronger in small groups of simulated fish which is in agreement with experiments.

Using a combination of Boltzmann inversion and traditional statistical methods, we have inferred how fish interact in schools. Previous studies have applied a force-matching approach to infer the interactions of schooling fish from their movements [36]. Our method can infer these interactions directly from the static spatial distribution of individuals in groups. Whilst repulsion forces had the same strength for different sized fish, attraction strength increased in larger fish, consistent with how a fish’s movement develops with age [24]. The interactions between fish also changed as a function of group size, as suggested by other studies [37]. Our model, refined on the basis of these observations, could capture the dynamics of schooling fish. Application of the approaches used in statistical physics, coupled with informed models of collective motion, now allows us to uncover the intricacies of how individuals in groups interact.

This work is supported by a Knut and Alice Wallenberg foundation grant No. 2013.0072 to DJTS.
Supplemental Material for: The statistical mechanics of schooling fish captures their interactions

Maksym Romenskyy,1 James E. Herbert-Read,1 Ashley J. W. Ward,2 David J. T. Sumpter1

1Department of Mathematics, Uppsala University, Box 480, Uppsala 75106, Sweden
2School of Biological Sciences, University of Sydney, Sydney, New South Wales, Australia

(Dated: September 1, 2015)

EXPERIMENTAL DETAILS

Materials

Pacific blue-eye fish (Pseudomugil signifer) were caught in hand nets from Narrabeen Lagoon, New South Wales, Australia (33°43′03 S, 151°16′17 E). Fish were kept in filtered freshwater in 150 l glass tanks at 22 − 25° and fed crushed flake food ad libitum. All fish were housed for at least three weeks prior to experimentation. The experimental arena was circular with a diameter of 760 mm. It was filled to a depth of 70 mm with aged and conditioned tap water. The arena was lit by fluorescent lamps and was visually isolated. For each trial, we randomly selected N fish (N = 10, 20, 30, 40, 50 or 60 for small and medium fish and N = 10, 20, 30, 40 for large fish) of similar size (see Fig. S1) from the housing tanks and placed them in the experimental arena. Fish were left to acclimate to the arena for at least five minutes, after which they were filmed for 15-20 minutes at 15 frames per second using a camera (Logitech Pro 9000) placed directly above the centre of the arena. The number of trials for each group size ranged between 3-10 (see Table I) due to limitations in the number of fish we could obtain for large or small body sizes. Because of the large numbers of fish we used for the experiment, we reused fish between trials. Fish were never used more than once per day and fish were used a maximum of 5 times.

Data collection and acquisition

Films were recorded in .mov format using original camera manufacturer software and subsequently converted to .avi using DirectShowSource and VirtualDub (v 1.9.2). The tracking was performed using DIDSON tracking program. The raw data consisted of x and y coordinates, fish identity and a time stamp. The accuracy of the tracking process was checked by projecting the raw tracking data onto experimental videos.

MOTION STATISTICS ANALYSIS

Calculation of the shell area for the pair distribution function in confinement

When calculating the pair distribution function for confined systems particular care needs to be taken when the particles are located close to the wall, at a distance smaller than the radius of the largest shell. In such situations only area of the shell lying inside the boundaries of the confining geometry should be considered. For all cases the shell area can be calculated as a difference between the areas of two neighbouring circles. The two latter areas for convenience can be computed separately as intersections excluding the areas outside the constraint. If confinement is circular (see Fig. S2), as is in our case, to find the intersection area we can use a formula for the circular segment of triangular height d′ and radius R′ [S1]:

\[ A(R', d') = R'^2 \cos^{-1}(d'/R') - d' \sqrt{R'^2 - d'^2}, \]  

with radii \( R_1 = R \) and \( R_2 = r \) and heights \( d_1 \) and \( d_2 \) calculated as

\[ d_1 = \frac{d^2 - r^2 + R^2}{2d}, \]  

\[ d_2 = \frac{d^2 + r^2 - R^2}{2d}. \]
FIG. S1. Distribution of standard lengths of fish for all group sizes.

| Group size | Average body length |
|------------|--------------------|
|            | 7.5 mm  | 13 mm  | 23 mm |
| 10         | 4       | 10     | 3     |
| 20         | 4       | 8      | 3     |
| 30         | 4       | 8      | 3     |
| 40         | 3       | 9      | 3     |
| 50         | 3       | 6      | –     |
| 60         | 3       | 6      | –     |

TABLE I. Number of videos recorded for each group size and body size. Each film is 15-20 min long.

FIG. S2. Illustration of the intersection of the pair distribution function circle $r$ with confining circle $R$. Only one circle for the pair distribution function is drawn for simplicity.
FIG. S3. Pair distribution function $g(r)$ for a homogeneous test system constituting of 1000 particles (average over $5 \times 10^6$ positional configurations), $R = 380$.

To calculate the total area of the intersection Eq. (S1) needs to be solved two times, once for each segment. Thus, combining Eqs. (S1-S3) we get

\[
A(r') = A(R, d_1) + A(r, d_2) = \int r^2 \cos^{-1} \left( \frac{d^2 + r^2 - R^2}{2dr} \right) + R^2 \cos^{-1} \left( \frac{d^2 + R^2 - r^2}{2dR} \right) - \frac{1}{2} \sqrt{(-d + r + R)(d + r - R)(d - r + R)(d + r + R)}. \tag{S4}
\]

Figure S3 shows a plot of the pair distribution of particles for highly homogeneous system in circular confinement (distribution of particles is uniform). The black curve displays a clear linear decay of $g(r)$ with increasing inter-particle separation distance $r$. It represents the case where the area outside the constraint has been also included in calculations. The red curve belongs to the case when all the shell areas have been calculated with the method described above. It shows highly regular distribution $g(r)$ of particles for all separation distances $r$ (1 : 1 relation between the local shell density and average density in the system).

**Calculation of the surface area occupied by a fish group**

The surface area of a group was calculated for every frame of a video. It was computed [S2] from a convex hull $C$ of a set of $N$ points representing spatial positions of the geometric centres of fish bodies $r_i$. The convex hull is defined by

\[
C \equiv \sum_{i=1}^{N} \lambda_i r_i : \lambda \geq 0 \text{ for all } i \text{ and } \sum_{i=1}^{N} \lambda_i = 1, \tag{S5}
\]

where $i$ is the point (fish) index and $\lambda$ is the non-negative weight coefficient. The resulting convex hull gives a complex polygon as an output. The area $A$ of this polygon can be computed as [S3]

\[
A = \frac{1}{2} \sum_{i=1}^{N-1} (x_i y_{i+1} - x_{i+1} y_i), \tag{S6}
\]

where $x_i$ and $y_i$ represent coordinates of the vertices of a polygon.

**Additional motion statistics**

Figure S5(a) shows speed distributions for fish of three average body lengths at fixed group size ($N=40$). Small fish ($l=7.5$ mm) most of the time have speeds below 100 mm/s and very often within a range of 0-20 mm/s. Very rarely
FIG. S4. Typical complex polygon returned by a convex hull algorithm. The displayed frame is taken from a video for 40 medium-sized fish.

FIG. S5. Experimental speed distributions for: (a) three sizes of fish in groups of 40 individuals, (b) small (l=7.5 mm), (c) medium (l=13 mm) and (d) large (l=23 mm) fish at variable group size.

small fish have speed above 200 mm/s. For medium-sized individuals the distribution is much wider and has a peak at \( V = 125 \) mm/s. It spans up to \( V \approx 250 \) mm/s and has another minimum at 0-20 mm/s. For large fish the number of events when the individuals are stationary or barely move decreases further and the peak is observed at \( V \approx 200 \). Figures S5(b)-(d) show speed distributions for small, medium and large fish and various group sizes. All histograms for the same body size practically overlap. Therefore, while speed regime of fish depends strongly on body size it is not effected by the number of individuals in a group.
FIG. S6. The illustration of the interaction parameters in the SPP model. The particle shown in red turns away from the nearest neighbours within the zone of repulsion (zor) to avoid collisions and aligns itself with the neighbours within the zone of alignment (zoa).

NUMERICAL SIMULATION METHODS

Self-propelled particle model

In our model, we consider \( N \) point particles at number density \( \rho \) and variable particle speed \( v_i \). The system undergoes discrete-time dynamics with a time step \( \Delta t \). The direction of motion of each particle (Fig. S6) is affected by repulsive or alignment interactions with other particles located inside the zone of repulsion (zor) or zone of alignment (zoa), respectively. Time evolution therefore consists of two steps: velocity updating and streaming. In the first computational step, position of each particle \((r_i)\) is compared to the location of the nearest neighbours. The repulsion rule has an absolute priority in the model and is modelled as a typical collision avoidance [S23]

\[
\hat{u}(t)_i = \sum_{j\neq i}^{n_r} \frac{r_{ij}(t)}{d_{ij}(t)} \mid r_{ij} = r_j - r_i \text{ and } n_r \text{ being a number of particles inside zor.}
\]

The alignment rule similar to one used in the Vicsek model takes into account velocities of all particles located inside the zone of alignment

\[
\hat{u}(t)_i = \sum_{j=1}^{n_a} \frac{V_j(t)}{n_a} \mid n_a \text{ being a number of particles inside zoa.}
\]

The velocities of particles are updated according to

\[
V_i(t) = v_i(t)\hat{u}(t)_i R_1(\xi_i(t))R_2(\theta_i(t)) \quad \text{(S7)}
\]

with \( v_i(t) = \psi(t) \) defining the particle individual speed \( v_i(t) \) based on the averaged local order \( \psi(t) \) inside both behavioural zones [S5–S8]. \( v_i(t) \) takes its maximal value \( v_i(t) = v_0 \) when velocities of particles inside zor and zoa are perfectly aligned \( \psi(t) = 1 \) while absence of local order \( \psi(t) = 0 \) results in \( v_i(t) = 0 \). The exponent \( \gamma \) controls the sharpness of the speed change. The misaligning noise is introduced through a random rotation \( R_1(\xi_i(t)) \) of the resulting particle velocity according to a Gaussian distribution \( P(\xi_i(t)) = e^{-\xi_i^2(t)/2\eta^2} / \sqrt{2\pi\eta} \), where \( \xi_i(t) \) is a random variable and \( \eta \) is the noise strength.

Wall avoidance is modelled as a particle orientation adjustment through rotation \( R_2(\theta_i(t)) \) of the particle velocity with a time-dependent turning rate \( \theta_i(t) = v_0\phi_i(t)/d_i(t) \). \( \phi_i(t) \) is the angle between the heading of a fish and normal to a time-dependent point of impact on the wall. \( d_i(t) \) denotes a distance from particle \( i \) to the impact point. Such construction of the rotating rate allows to achieve its strong damping at large distances from the wall and for smaller angles of approach of a collision point on the wall. At these conditions its influence on particle’s motion is insignificant.

When the velocity update step is complete, the particle positions are updated by

\[
r_i(t + \Delta t) = r_i(t) + V_i(t)\Delta t \quad \text{(S8)}
\]
The unit of length in our simulations is equivalent to the metric length used in the experiment. To set the unit of time we choose a particle speed \( v_0 = bv_0^e \), where \( b \) is the behavioural reaction time \([S9]\) of fish (\( b = 0.05 \) s) and \( v_0^e \) is the average speed in experiment. The integration time step was set to \( \Delta t = 1 \) for all simulations. The differences in body size of fish were introduced through a scaling of the radius of the attraction zone proportional to the experimentally observed distribution of fish size.

**Statistics collection**

All code was written in FORTRAN. Total number of time steps in each run was \( 1 \times 10^6 \). The statistics was collected in the steady state and each characteristic of motion was calculated by averaging over five independent runs. The radius of the arena was fixed at \( R = 380 \) for all simulations. The initial conditions for fish positions and velocities were chosen at random from the uniform distribution.

---

[S1] E. W. Weisstein, “Circle-circle intersection,” MathWorld – A Wolfram Web Resource. http://mathworld.wolfram.com/Circle-CircleIntersection.html.

[S2] C. B. Barber, D. P. Dobkin, and H. Huhdanpaa, ACM Trans. Math. Softw. 22, 469 (1996).

[S3] P. Bourke, “Calculating the area and centroid of a polygon,” http://local.wasp.uwa.edu.au/leith.

[S4] I. D. Couzin, J. Krause, N. R. Franks, and S. A. Levin, Nature 433, 513 (2005).

[S5] W. Li and X. Wang, Phys. Rev. E 75, 021917 (2007).

[S6] J. Zhang, Y. Zhao, B. Tian, L. Peng, H.-T. Zhang, B.-H. Wang, and T. Zhou, Physica A: Statistical Mechanics and its Applications 388, 1237 (2009).

[S7] S. Mishra, K. Tunsstrøm, I. D. Couzin, and C. Huepe, Phys. Rev. E 86, 011901 (2012).

[S8] S. Lu, W. Bi, F. Liu, X. Wu, B. Xing, and E. K. L. Yeow, Phys. Rev. Lett. 111, 208101 (2013).

[S9] P. Domenici and R. W. Blake, J. Experimental Biol. 200, 1165 (1997).