Chapter 3

A data-driven method to investigate the integration of information in fish schools

Valentin Lecheval, Hanno Hildenbrandt, Clément Sire, Guy Theraulaz and Charlotte K. Hemelrijk

Abstract

Collective motion of animals in groups results from interactions among group-members and interactions of group-members with their physical environment (e.g. obstacles, temperature and light gradients). Studying these systems is challenging, in particular because of the difficulty to access how an individual perceives its environment and how it combines and reacts to multiple sources of information in its reaction. Integration of information has rarely been explored in previous studies. Current models of fish schools usually assume that fish average their reactions with respect to each separate stimulus. We develop a data-driven method based on the reconstruction of behavioural action maps derived from empirical data. These maps describe the reaction of fish for each social or non-social stimulus. The rationale of this method is to test hypotheses regarding the integration of stimuli from multiple sources without making assumptions regarding the interactions of fish. We investigate a simple hypothesis in which fish react only to the strongest stimulus. We assume that the strength of a stimulus is proportional to the angle a fish turns in reaction to it. Namely, the stronger a stimulus, the more a fish would turn. We use this in our simulations to compute the assumed strength of each possible stimulus on the
basis of the turns performed by fish experiencing similar stimulus in empirical data. We tested this method with empirical data in a ring-shaped tank with non-social (the walls of the corridor) and social (in groups of 2 and 5 fish) stimuli. We find that the hypothesis that fish react only to the strongest stimulus is not sufficient to reproduce the collective-level properties observed in experiments. This suggests that fish react to more than one information source at each moment of decision. We discuss further improvements of our method.

Contribution of authors

V.L., C.S., G.T., H.H., and C.K.H. designed research; V.L., H.H. and C.K.H. performed research and developed the model; H.H. implemented model; V.L. analysed data; V.L., G.T. and C.K.H. wrote the paper.

3.1 Introduction

The study of complex phenomena such as collective motion in animal societies is a challenging field of research. Recently, the combination of experimental and theoretical work has contributed to identify the main questions that have to be tackled when investigating such a system. Modelling what individuals in a moving group perceive (Lemasson et al., 2009, 2013; Rosenthal et al., 2015; Collignon et al., 2016), defining a neighbourhood (Ballerini et al., 2008) and measuring the interactions between individuals (Katz et al., 2011; Herbert-Read et al., 2011; Gautrais et al., 2012; Calovi et al., 2017) are crucial steps to fully connect the individual behaviour to the collective patterns that emerge at a group-level. However, only a few studies explicitly address the question of the individual-level integration of information from multiple sources by individuals in the context of collective motion. In most models of collective motion, it is commonly assumed that, given the reactions to all the perceived information (usually pair-wise interactions with the neighbours and the physical environment), an individual will react by taking the average of the reactions, possibly weighted (e.g. by the distance to each neighbour) (Vicsek et al., 1995; Lopez et al., 2012). The validity of this assumption is in general untested although it can be questioned. Averaging leads to the damping of behaviours that might have been elicited by external perturbations (such as the attack of a predator) and results in simulated groups being less reactive to perturbations than groups of real fish (Katz et al., 2011). It has been suggested that fish would react considering the sum of the probability density functions asso-
3.2. MATERIAL AND METHODS

3.2.1 Experimental procedures and data collection

70 rummy-nose tetras (*Hemigrammus rhodostomus*) were used in our experiments. Fish were purchased from Amazonie Labège (http://www.amazonie.com) in Toulouse, France. They were kept in 150L aquariums on a 12:12 hour, dark:light photoperiod, at 27.5°C (±0.8°C) and were fed *ad libitum* with fish flakes. These tropical freshwater fish swim in a highly associated with each perceived stimulus weighted by their visual importance (Collignon et al., 2016). Despite the novelty and elegance of this approach, this assumption regarding the integration of information has been tested in combination with many other assumptions (a model of visual perception, the reaction to the wall, ...) and has thus not been properly tested in itself.

In this chapter, we aim to develop a method to test hypotheses about the integration of information by fish in schools. We use experiments of schools of rummy-nose tetras (*Hemigrammus rhodostomus*) of several sizes (1, 2 and 5) in a ring-shaped tank. We investigate first the burst-and-coast behaviour analysed in the previous chapter and show that the interactions between individuals also depend on the control of their speed, in agreement with previous studies (Katz et al., 2011). The empirical data are used to simulate new trajectories of fish without assuming a mathematical equation for the interactions. First, from empirical data we map the sensory inputs of fish to their motor actions (i.e. turning and speeding behaviours). Second, we make hypotheses about the way the sensory inputs are considered and combined by the individuals and we simulate trajectories accordingly by using the mapped motor actions. Thus, the model concerns only the perceived information, an asynchronous, discrete decision-making process (validated in the previous chapter for 1 and 2 fish) and the integration of information from multiple sources, i.e. social (from several neighbours) and non-social (from the walls of the tank). In particular, we investigate what are the collective patterns of schooling that emerge if, at each time of decision, fish only react to the strongest stimulus they perceive. There are many examples in nature of animals choosing without compromise when facing opposite information (Burgess et al., 2010), that contradicts the assumption of information integration being only about averaging over all possible reactions. We investigate whether such a simple rule of selective attention to information suffices to lead to a schooling state, where individuals remain close to each other and highly aligned and react strongly to internal and external perturbations (see Part II of this manuscript).
Figure 3.1: Experimental set-up. The ring-shaped tank is built from two walls, an outer wall of radius 35 cm and a conic inner wall with an average radius of 21 cm at the water level creating a corridor of 14 cm width. Photograph shot by David Villa, ScienceImage CBI CNRS, Toulouse, France.

Table 3.1: Group size, number of trials, their total duration and average body length of individuals.

<table>
<thead>
<tr>
<th>Group Size</th>
<th>Number of trials</th>
<th>Total duration</th>
<th>Body length (mm, mean ± se)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>260 min</td>
<td>33.1 ± 1.8</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>652 min</td>
<td>33.3 ± 0.8</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>543 min</td>
<td>31.5 ± 0.3</td>
</tr>
</tbody>
</table>

synchronised and polarised manner. Inside an experimental tank, a ring-shaped corridor 10 cm wide with a circular outer wall of radius 35 cm was filled with 7 cm of water of controlled quality (50% of water purified by reverse osmosis and 50% of water treated by activated carbon) heated at 27.6°C (±0.9°C) (Figure 3.1). The shape of the circular inner wall was conic and its radius at the bottom was 25 cm. The conic shape was chosen to avoid the occlusion on videos of fish swimming too close to the inner wall.

For each trial, \( n \) fish \((n \in \{1, 2, 5\})\) were randomly sampled from their breeding tank. Each fish only participated per day in a single experiment. Fish were introduced in and acclimatised to the experimental tank during a period of 10 minutes before the trial started. During each trial of one hour, individuals were swimming freely without external perturbation. For each group size, we performed 10 replications (Table 3.1). Note that six experiments with a single fish have been discarded because of the inactivity of the individuals. Trajectories of the fish were recorded by a Sony HandyCam HD camera filming from above the set-up at 50Hz in HDTV resolution (1920×1080p).
3.2. MATERIAL AND METHODS

A

![Diagram A]

B

![Diagram B]

Figure 3.2: Symbols for analysis of fish motion in experiments with 1 fish (A) and 2 fish (B). Fish $i$ has a position $(x_i, y_i)$ at the onset of the burst. We measure its current heading $\varphi_i$, its angle to the outer wall $\theta_w$, its distance to the outer wall $r_w$, its distance to neighbour $d$, the angular position of the neighbour $\psi$ and the difference of heading between the two fish $\Delta \phi$.

3.2.2 Data extraction and pre-processing

We tracked the positions of each individual using idTracker 2.1 (Pérez-Escudero et al., 2014). Sometimes, the tracking software lost certain individuals, for instance when two fish were swimming too close to each other. All sequences that were missing a maximum of 50 consecutive positions were interpolated.

Time series of positions were converted from pixels to meters with the origin of the coordinate system set to the centre of the ring-shaped tank. Body lengths and headings of fish were measured on each frame using the first axis of a principal component analysis of the fish shape issued by idTracker. Details regarding the detection of the walls, of the fish shape and the conversion to metric frame of reference are presented in the first chapter. Table 3.1 summarises the data collected in our study.

3.2.3 Segmentation

Following our findings of Chapter 1, we assume that fish take decisions at the onset of the burst phase of their burst-and-coast cycle. Thus trajectories are segmented with respect to the swimming speed of the individuals with the same method as presented in the first chapter. In short, the seg-
Our measures are thus taken at the onset of the burst. For instance, the distance between two neighbours \( d \) is considered as the distance between the two neighbours when the focal fish starts its burst. For a fish \( i \), at position \( (x_i, y_i) \) at the onset of the burst, swimming in a tank of outer radius \( R \) with a neighbour \( j \), we measure its current heading \( \varphi_i \), the angle turned \( \delta \phi \), its angle to the outer wall \( \theta_w \), its distance to the outer wall \( r_w \), its distance to the neighbour \( d \), the angular position of the neighbour \( \psi \) and the difference of heading between the two fish \( \Delta \phi \) (Figure 3.2). As for the speed parameters, durations of the burst and coast periods (respectively \( b \) and \( c \)), the acceleration during the burst \( a \) and the top speed reached at burst \( v_{\text{max}} \) are measured. All these quantities constitute one data point for each burst-and-coast segment measured on both experimental and simulated trajectories. Experimental data will be analysed to investigate further the burst-and-coast behaviour of the rummy-nose tetra and to study the influence of social information on the control of speed by the fish. They will also be used to check the consistency of the hypotheses about the way information is integrated by fish.

### 3.2.4 Computational model based on behavioural action maps

We develop a computational model that samples experimental data to model fish motion. We investigate the role of integration of information from multiple sources, which may be social (the neighbours) or non-social (the wall). The model is based on behavioural action maps that connect each sensory input (e.g. distance and relative orientation to obstacles and neighbours) to a pattern of actions (angular and velocity change). These behavioural action maps are based on experimental data. The rationale of this data-driven approach of sampling experimental data is to avoid making any assumptions regarding fish interactions.

**Algorithms and assumptions**

Agents react only at the onset of each burst, meaning that they only use information available at this point of time. We assume that all social information perceived can be decomposed into a sum of pair-wise interactions with each neighbour (Katz et al., 2011; Gautrais et al., 2012; Lopez et al., 2012). We also assume that if a larger angle \( \delta \phi \) is turned in experimental data, the stimulus that elicited the reaction is greater and thus the priority of this stimulus is higher.

When a focal agent enters the burst phase:
3.2. MATERIAL AND METHODS

Figure 3.3: Discretisation of the local environment of a focal fish into 8 sectors. The individuals in the squares are labelled as influential because they are the nearest neighbours of the focal fish in each sector.

1. Its local environment is given by the number and location of its influential neighbours and by the nearest point of all walls (in agreement with findings of Chapter 1);

2. It evaluates the priority of each stimulus (social and non-social);

3. It integrates the information according to the priority of each stimulus;

4. It reacts by controlling its actual velocity and orientation.

The following paragraphs describe each component of this algorithm.

1 – Local environment of a focal agent  Regarding the wall, as the previous chapter showed, taking only the closest point of the wall of a circular arena into account suffices to reproduce with simulations the empirical effect of the wall. Thus, in the model we can treat the point of the walls that is the closest to the focal fish as a potential stimulus. As for social interactions, the environment of the focal agent is discretised into $s = 8$ sectors (Figure 3.3). In each sector, the nearest individual to the focal agent is considered to be possibly influential.

2 – Evaluation of the priority of each stimulus  For each potential stimulus, its priority is given by assuming that it is positively correlated with the turning angle induced in a fish subject to a similar environmental configuration in empirical data. Thus, we need a way to measure situations that are similar in experiments. First, we doubled the size of the data by
Figure 3.4: Discretisation of the space $S_w(|\theta_w|, r_w)$ used to sample reactions of a single fish in response to the influence of the wall. Blue lines stand for the bin boundaries. Data points are grey and transparent – the darker a bin, the more points it contains.

adding the symmetric data, as the trajectories had also been recorded from the underside. Namely, we satisfy the symmetry condition

$$\delta \phi(r_w, -\theta_w) = -\delta \phi(r_w, \theta_w),$$

(3.1)

for experiments with 1 fish and

$$\delta \phi(r_w, -\theta_w, d, -\psi, -\Delta \phi) = -\delta \phi(r_w, \theta_w, d, \psi, \Delta \phi),$$

(3.2)

for experiments with 2 fish. When the potential stimulus is a point of the walls, we use the data from experiments with 1 fish and discretise the space $S_w(\theta_w, r_w)$ in $12 \times 6 = 72$ bins (Figure 3.4). When the potential stimulus is a neighbour, we use the data from experiments presented in the previous chapter with 2 fish swimming in the circular tank and away from the wall (65,156 data points before applying the symmetry condition with $r_w > 2.5 \text{BL}$). This is done to remove the influence of the wall from the social interactions. We discretise the space $S_n(d, \psi, \Delta \phi)$ in $10 \times 20 \times 20 = 4000$ bins (Figure 3.5).

The number and the size of the bins are set manually to obtain a good resolution (so that the simulation can sample many different behaviours) while minimising the number of empty bins. Thus, binning is different for each condition of distance between neighbours (see Figures 3.5B, C and D).

Given the potential stimuli, $p$ turning angles are sampled from the experimental turning angles $\delta \phi$ found in the bin that corresponds to each
Figure 3.5: Discretisation of the space $\mathcal{S}_n(d, \psi, \Delta\phi)$ used to sample reactions of several fish in response to the neighbours, from experiments with 2 fish in the circular tank and away from the walls. Blue lines stand for the bin boundaries. A). Discretisation regarding distance to the closest neighbour. B-D). Discretisation of $(\psi, \Delta\phi)$ for three different bins regarding distance to the closest neighbour, as shown by the letters on (A). Data points are grey and transparent – the darker a bin, the more points it contains.
stimulus. The average of the $p$ turning angles is the estimation of the priority of each stimulus. This average can be weighted by the arbitrary parameters “Weight for turning angle in response to the wall” and “Weight for turning angle in response to neighbours” to change the relative influence of either the wall or neighbours (both are set to the same value by default).

3 – Integration of information Given the measured priority of each potential stimulus, the highest one only is selected as the source that elicits a behavioural reaction in the focal agent.

4 – Behavioural reaction of the agent A behavioural reaction is a set of a turning angle $\delta\phi$ (rad), an acceleration $a$ (BL/s$^2$), a duration of burst $b$ (s) and a duration of coast $c$ (s), randomly sampled from the bin that corresponds to the selected stimulus. A parameter controls whether these 4 parameters of swimming are correlated (i.e. taken from the same data point) or not (i.e. a different data point is sampled from the same bin for each parameter). The position at the onset of the burst $k$ $(x_k, y_k)$ is updated as follows:

$$\begin{align*}
x_{k+1} &= x_k + l_k \cos(\phi_k) \\
y_{k+1} &= y_k + l_k \sin(\phi_k),
\end{align*}$$

(3.3)

with $\phi_k = \phi_{k-1} + \delta\phi$, and $l_k = l_b + l_c$. $l_b$ stands for the distance travelled during burst where acceleration is assumed constant. Thus $l_b = \frac{1}{2}ab^2 + v_kb$, where $v_k$ is the speed of the agent at the onset of the burst. $l_c$ stands for the distance travelled during coast where only drag due to water frictions is applied (see Chapter 1 and Figure 3.6). Thus $l_c = v_m(\exp(-Dc) - 1)/D$, with $v_m$ the top speed of the burst phase $v_m = v_k + ab$ and $D$ the drag coefficient. If there are no experimental data in the bin that corresponds to the configuration of the focal agent and its selected stimulus, default values are chosen (i.e. default value for acceleration strength, burst duration and coast duration and a null angle turned – see parameter values in Table 3.2). The binning of $S_w(\theta_w, r_w)$ and $S_n(d, \psi, \Delta\phi)$ is set to avoid empty bins.

Implementation

The computational model has been implemented in C++ by Hanno Hildenbrandt. All parameters used in simulations are shown in Table 3.2. The geometry of the tank in simulations has been measured from the experiments of a single fish (radius of the inner ring: $R_i = 6.27$ BL and outer ring: $R_o = 10.40$ BL).
3.2. MATERIAL AND METHODS

Figure 3.6: Decay of speed during coasting phases relative to top speed in experimental data (red curve) and as modelled by a drag due to water frictions with exponential decay of parameter $D \approx 0.71$ (blue). The grey shade shows the bootstrapped 95% confidence interval.

Figure 3.7: Parameters of the rejection algorithm (see section 3.2.4). The real radii of the outer wall ($R_o$) and of the inner wall ($R_i$) are not used to delimit the walls. Because of the length $L$ of the front end of the fish to its barycentre used in simulations and of the conic shape of the inner wall that interplays with the water level, an effective outer radius $R_{o\text{eff}}$ and inner radius $R_{i\text{eff}}$ are fitted to define the tank boundaries used by the rejection algorithm.
Table 3.2: Default parameters used in simulations.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body length of individual</td>
<td>33.08</td>
<td>mm</td>
</tr>
<tr>
<td>Body width</td>
<td>6.02</td>
<td>mm</td>
</tr>
<tr>
<td>Radius inner wall $R_i$</td>
<td>6.44</td>
<td>BL</td>
</tr>
<tr>
<td>Radius outer wall $R_o$</td>
<td>10.69</td>
<td>BL</td>
</tr>
<tr>
<td>Effective inner radius $R_{ieff}$</td>
<td>0.7</td>
<td>BL</td>
</tr>
<tr>
<td>Effective outer radius $R_{oeff}$</td>
<td>-0.15</td>
<td>BL</td>
</tr>
<tr>
<td>Drag coefficient $D$</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>Default coasting duration</td>
<td>0.50</td>
<td>s</td>
</tr>
<tr>
<td>Default acceleration at burst $a$</td>
<td>4.00</td>
<td>BL / $s^2$</td>
</tr>
<tr>
<td>Default burst duration</td>
<td>0.20</td>
<td>s</td>
</tr>
<tr>
<td>Default turning angle</td>
<td>0</td>
<td>rad</td>
</tr>
<tr>
<td>Number of sectors $s$</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Correlated data points</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Weight for turning angle in response to the wall</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Weight for turning angle in response to neighbours</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Simulation length</td>
<td>$5.10^5$</td>
<td>s</td>
</tr>
<tr>
<td>Frame rate</td>
<td>50</td>
<td>Hz</td>
</tr>
<tr>
<td>Number of samples for the rejection algorithm</td>
<td>10000</td>
<td></td>
</tr>
<tr>
<td>Number of samples $p$</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

To avoid collisions with the wall, we use the same rejection algorithm as discussed in the first chapter. Namely, a set $(x_{k+1}, y_{k+1})$ will be resampled a certain number of times (see Table 3.2) as long as they fall outside the corridor defined by two parameters $R_{oeff}$ and $R_{ieff}$, respectively standing for “outer effective” and “inner effective” (Figure 3.7).

By default, agents cannot see through the inner wall. Agents can perceive all around themselves, i.e. they have no blind angle.

Parameters $R_{ieff}$ and $R_{oeff}$ are fitted by comparing the simulated distribution of the distance to the wall to the experimental one (see Figure 3.13A).

**Output of simulations**

The output of simulations are the position and velocity of agents at the same frequency as the one in the movies of experiments (50 Hz) so that it can be analysed with the same code as used for experimental data. For each position, we save whether the agent was bursting or coasting at this time step. There is also a second data set with the different quantities asso-
3.2. MATERIAL AND METHODS

Figure 3.8: Frame of the real-time visualisation of a simulation with two fish. For parameters, see Table 3.2.

associated with each burst-and-coast period (namely with the index of the focal burst-and-coast period, the distance of the agent to the wall, its angle to the wall, the distance, angular position and difference of headings between neighbours as well as the chosen parameters of swimming – turning angle, acceleration strength, burst duration and coast duration). Simulations are visualised in real-time as a method to assess their quality (Figure 3.8).

Conditions tested in simulations

All simulations with 1 fish are run with the default parameters shown in Table 3.2. For simulations with 2 and 5 fish, we test the influence of the following parameters (values of parameters in italics):

- Correlated data points: this controls whether the sampled swimming parameters come from the same data point (yes, default) or not (no).

- Weighted wall interaction: by default, the comparison of the priority of a stimulus from a neighbour and the stimulus from the wall are equally weighted (parameters Weight for angular turn in response to the wall and Weight for angular turn in response to neighbours are both set to 1) (value set to no). Otherwise (yes), the parameter Weight for angular turn in response to the wall is set to 0.5 and in response to neighbours is set to 1 (social stimuli are thus favoured).

- Invisible walls: by default, the walls are opaque and fish cannot see through them (value set to no), so they are not influenced by an agent
Table 3.3: Combination of parameters for the five conditions tested for the simulations with 2 fish. nocor (no correlation), reg (regular), wei (weighted wall interaction), inv (invisible walls) and wei-inv (weighted wall interaction and invisible walls) are the labels of each condition used in the text and the figures.

<table>
<thead>
<tr>
<th>Correlated data points</th>
<th>nocor</th>
<th>reg</th>
<th>wei</th>
<th>inv</th>
<th>wei-inv</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weighted wall interaction</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Invisible walls</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

at the opposite side of the tank. Otherwise (yes) the radius of the inner wall \( R_i \) was set to 0 and \( R_{\text{eff}} = 7.14 \) (fish cannot go through the invisible wall).

Five different combinations with different values of these three parameters have been studied (Table 3.3) and labelled as nocor (no correlation), reg (regular), wei (weighted wall interaction), inv (invisible walls) and wei-inv (weighted wall interaction and invisible walls). The regular (reg) condition corresponds to all parameters set to their default values (Table 3.2): swimming parameters are correlated, the respective weights for social and non-social stimuli are identical and agents cannot react to stimuli occluded by walls. The no correlation (nocor) condition tests the effect of the correlation between the sampled swimming parameters (Parameter Correlated data points set to no)). The weighted wall interaction (wei) condition aims to decrease the influence of the wall (with respect to social stimuli) by dividing the strength of wall stimuli by 2. In the invisible walls (inv) condition, the agents are not occluded by the inner wall to their neighbours anymore. The condition aims to test whether real fish remember the presence of occluded fish or not. The last condition, weighted wall interaction and invisible walls (wei-inv), is with diminished influence of the wall and invisible walls.

3.3 Results

3.3.1 Experimental results

The distributions of the duration of the bursting and coasting periods have the same mode (≈ 0.2 s) and both are positively skewed yet the tail of the distribution of coast durations is longer and heavier, meaning that the durations of the coasting are on average longer than of the bursting...
3.3. RESULTS

Figure 3.9: Distributions of the burst duration (A), coast duration (B), acceleration at burst (C) and absolute value of the turning angles (D) of 1 fish/agent in empirical data (bars) and simulations (red curve).
3.3. RESULTS

Figure 3.10: Effect of the individual’s position to the outer wall on the average bursting duration (A and B), on the average coasting duration (C and D) and on the average acceleration of fish (E and F) in experiments with 1 fish (page 72). Left figures show effect of the distance to the outer wall, right figures the effect of the absolute value of the angle to the outer wall $\theta_w$ ($0$: facing the outer wall, $\pi$: facing the inner wall). Red vertical lines stand for ± standard error.

(Figures 3.9A and B). The acceleration of fish during the bursting period has a large variance: all accelerations between 0.5 BL/s$^2$ and 12 BL/s$^2$ are common (Figures 3.9C). As found in the previous chapter, fish tend to turn away from the outer wall (Figures 3.9D).

As to the results regarding control of the speed by fish depending on their location to the wall, for fish swimming alone, we find no effect of the distance to the wall on the average duration of bursts (Figure 3.10A) while it seems that fish tend to burst for a longer time when facing the walls (Figure 3.10B). On average, the coast duration is longer when the fish are away by approx. more than 1 BL from the walls and swimming aligned with the walls (Figures 3.10C and D). Acceleration seems weaker when fish are close to the walls and facing them (Figures 3.10E and F).

As for the effect of the neighbour, fish adjust the duration of the coast depending on the angle and distance to their neighbour, with a relatively complex interaction between these two variables (Figures 3.11A-B and 3.12A). Fish coast for a longer period of time when their neighbour is at their side, at distances between 1 and 2 BL, than any other distance. Regarding acceleration, fish accelerate stronger when the neighbour is in front of them, at large distances (greater than 2 BL) (Figures 3.11C-D and 3.12B). The combination of these bursting and coasting behaviours result in a clear pattern of top speed where bursting fish are faster when their neighbour is in front or behind them (Figures 3.11E-F and 3.12C).

In the previous chapter, speed was not a behavioural component of the interactions between individuals – only an effect on turning angle has been modelled and the speed parameters were sampled randomly from the experimental distributions. The results we show in the current section suggest that fish may also control their speed in reaction to the spatial location of their neighbour. This supports our choice to develop the computational model presented in the previous section, where the assumptions regarding interactions between individuals are minimal: the model shown in the previous chapter might suffer from the absence of key elements in the in-
CHAPTER 3.

A. Coast duration (s) vs. Nearest neighbour distance (BL).

B. Coast duration (s) vs. Angular position of the neighbour (rad).

C. Acceleration at burst (BL / s^2) vs. Nearest neighbour distance (BL).

D. Acceleration at burst (BL / s^2) vs. Angular position of the neighbour (rad).

E. Top speed at burst (BL / s) vs. Nearest neighbour distance (BL).

F. Top speed at burst (BL / s) vs. Angular position of the neighbour (rad).
Figure 3.11: Effect of the angular position of the neighbour on the average coasting duration (A and B), the average acceleration of fish (C and D) and the average top speed at burst (E and F) in groups of 2 fish (page 74). Left figures show effect of the distance from the neighbour, right figures the effect of the angular position of the neighbour $\psi$ (0: in front of the focal fish, $\{-\pi; \pi\}$: behind the focal fish). Red segments stand for ± standard error.

Figure 3.12: Effect of the relative position of the neighbour on the coasting duration (A), the acceleration of fish (B) and top speed at burst (C) in groups of 2 fish. White areas stand for absence of data or cells with less than 25 data points.
interactions between individuals (especially regarding control of speed) that would prevent an accurate investigation of the integration of various stimuli in groups of more than 2 fish.

3.3.2 Simulation results vs experimental data

The distributions of the quantities used as input for the model are reproduced by the simulations with 1 fish (Figures 3.9A-D). Thus the model samples the behaviours that are of importance, which would not be the case if it was sampling a small number of different bins of the input maps during simulations. As for the emergent quantities, the model fairly reproduces the main features of the experimental data (Figures 3.13A-C).
Namely, agents are closer to the outer wall than to the inner one, aligned with the walls, and they swim at an average speed of approximately 5 BL. Nevertheless, we note a significant deviation from the experimental distributions, in particular for the distance and angle to the outer wall. Agents in simulations are rarely very slow (speed less than 1 BL) because data points of fish that were not moving have been removed during the segmentation procedure used for the input of the model. We note an important effect of the binning used for the data input of the model on the quantitative adjustment to experimental data. In particular, when there are too many bins, many of them are empty and the default behaviour (see Table 3.2) is over-represented.

For simulations with at least 2 fish, we measure, at each new burst, whether the selected stimulus was social (a neighbour) or not (the wall). The results for the five conditions of simulations detailed in Table 3.3 are shown in Figure 3.14A. The no correlation (wei) and regular (reg) conditions give similar results, with the wall being selected in 20% of all decisions and the neighbour in 80% of all decisions. The other conditions decrease the frequency of the wall being the selected stimulus, with a minimum reached for the condition weighted wall interaction and invisible walls (wei-inv) (4.2%). These results are not surprising since the weighted wall interaction (wei) condition disadvantages the selection of the wall as a stimulus while the invisible walls (inv) condition favours social interactions since the neighbour of the focal agent is never occluded by the inner ring. We find that the non-social stimuli are mostly chosen when the agent is at very short distances (less than 0.33 BL) from the outer wall (Figure 3.15).

The quantitative adjustment between simulation results and experimental data are shown in Figures 3.14B, C and D for the quantities measured at the onset of the burst and in Figures 3.16A-D for the quantities measured on all frames. Overall, the model fits fairly well the monotonicity and the modes of all distributions but with quantitative differences. Agents in simulations stay at short distances from each other, although there are cases where the group splits (for distances to neighbour greater than 6 BL) that do not occur in experimental results (Figure 3.14B). In experiments, the neighbour of the focal fish is mainly found in front or behind it and not at its side (Figure 3.14C). In simulations, agents are too often found at the side of the focal agent ($|\psi| \approx \pi/2$) and not enough in front ($|\psi| \approx 0$). In experiments, fish are mainly found swimming in the same direction, with $\Delta \phi \approx 0$ (Figure 3.14D), leading to a highly polarised group (Figure 3.16D). Strong polarisation is absent in simulations (Figure 3.16D) although there
Figure 3.14: A). Relative frequency of the selected stimulus for each condition of simulation (see Table 3.3). Distributions of the distance to the neighbour (B), angular position of the neighbour $\psi$ (C) and difference in heading between fish $\Delta \phi$ (D) measured at the onset of bursts in experiments (bars) and simulations (curves) with 2 fish.
3.3. RESULTS

Figure 3.15: Model distributions of the distance to the outer wall and of the absolute value of the angle to the outer wall when the wall is selected as the strongest stimulus (respectively (A) and (C)) and when the neighbour is selected as the strongest stimulus (resp. (B) and (D)), for simulations with 2 fish in the weighted wall interaction (wei) condition.
Figure 3.16: Distributions of the (A) distance to the outer wall, (B) absolute value of the angle to the outer wall, (C) speed of the group and (D) polarisation of the group measured on all frames in experiments (bars) and simulations (curves) with 2 fish.
3.4 Discussion

Understanding what individuals perceive and how they combine this information to react are crucial questions to connect the individual behaviour to the collective behaviour in moving groups of animals. In this paper, we present a new method to investigate the integration of stimuli in fish from experimental trajectories and stochastic simulations. This method samples experimental data points after segmenting the trajectories in burst-and-coast. The quantitative reactions of agents to stimuli are thus not assumed (i.e., we did not measure interaction functions as we did in the previous chapter). However, we assumed that fish choose their orientation at the onset of the burst period and that only water frictions affect fish veloc-
Figure 3.17: A). Relative frequency of the selected stimulus for each condition of simulation (see Table 3.3). Distributions of the (B) distance to the nearest neighbour, (C) angular position of the nearest neighbour $\psi$ and (D) difference in heading between focal fish and nearest neighbour $\Delta \phi$ measured at the onset of bursts in experiments (bars) and simulations (curves) with 5 fish.
3.4. DISCUSSION

Figure 3.18: Distributions of the absolute value of the angle turned in experiments (bars) and simulations (curves) for groups of 2 (A) and 5 fish (B).

ity during the coast periods, in agreement with the previous chapter. In addition to these assumptions, we show from experiments that individuals control their speed in response to social and non-social information by adjusting 3 parameters: the duration of the burst and coast periods and the strength of the acceleration during the bursts. We use these findings in our simulations, where agents control their orientation as well as their speed.

We find that our method, with these new assumptions regarding the burst-and-coast behaviour of fish, succeeds to reproduce emergent quantities in experiments with 1 and 2 fish. The mechanisms assumed (namely that fish control the duration of the burst and coast phases and the strength of their acceleration) are thus compatible with empirical data, which helps to understand how fish control their speed in response to a changing environment, completing the findings of the previous chapter. We predict that these mechanisms should, at least to some extent, be similar among species with a burst-and-coast swimming behaviour, such as the zebrafish (Danio rerio) (Harpaz et al., 2017).

Although the trajectories simulated with our method are in a fairly good agreement with empirical data for 1 and 2 fish, we note that there is still some deviation with the experimental results (i.e. the fit could be better, even for simulations with 1 fish). Improving the binning of the sampled spaces might help to improve our results. Future research is required to find an efficient algorithm to make bins in 2 and 3-dimensional spaces with
Figure 3.19: Distributions of the distance to the outer wall (A), absolute value of the angle to the outer wall (B), speed of the group (C) and polarisation of the group (D) measured on all frames in experiments (bars) and simulations (curves) with 5 fish.
respect to the density of experimental points – as we tried here by using quantiles.

In this paper, we test the hypothesis that fish only react to the strongest stimulus among all the possible stimuli in their environment. Stimuli can be non-social (the walls of the ring-shaped tank) or social (the neighbours of the focal fish). Our method simulates trajectories that qualitatively resemble empirical data for individual fish and groups of 2, it suggests that all the key features of a swimming fish are covered. However, we note that in groups of 2 agents, the neighbour of the focal agent was much less often in front of it than in empirical data. In other words, the focal agent often selected social stimuli when their neighbour was behind them. This is even worse in groups of 5 agents. This questions our procedure of selection of the strongest stimulus. In future work we will try to adjust the weight of the potential stimuli by their relative position to the focal agent, in order to favour frontal stimuli, especially when they are social. In our simulations, agents perform more large turns than real fish. These U-turns might be performed in empirical data in response to a fish located behind. Thus, in addition to select more often a frontal neighbour (in agreement with empirical data), weighting the stimuli by their relative position to the focal agent may also help to decrease the too high number of large turning angles seen in simulations.

We note that groups of 2 and 5 agents are slower than real fish. This is explained by two distinct points. Firstly, we have used empirical data from experiments conducted in a circular tank and, in this tank, fish were slower than in the ring-shaped tank. This issue also concerns the turning angle, whose distribution may differ from empirical data in the ring-shaped tank. Secondly, in empirical data, we observe a greater swimming speed in groups of 5 fish than in pairs of fish, which is not reproduced by our simulations. Thus, our results suggest that both the ring-shaped tank and the number of individuals in the group affect the swimming speed in a way that is not grasped by our model. In other words, the shape of the tank and the number of group-members may change the stimulation level of fish. Thus, using data from experiments with 2 fish to simulate trajectories with 5 might not be possible using this method.

All the sets of parameters tested gave similar results. This suggests that our results are not sensitive to parameters. In particular, it is worth to note that the “inv” condition, where the inner wall is invisible, does not improve the results. It implies that the occlusion of the neighbour by the inner wall does not decrease the quality of the simulated trajectories. It suggests that fish do not need to remember that they are not alone in the
tank when their neighbour is occluded. Namely, we do not find evidence for a memory effect regarding the presence of a neighbour in the experiment.

Our simulations with 5 agents do not resemble empirical trajectories. Several explanations that are non-exclusive are possible. The assumption that fish would react only to the strongest stimulus does not seem satisfactory for groups larger than two fish, that is when fish need to react to more than one neighbour. We stress that *Hemigrammus rhodostomus* is a fish that makes very polarised groups, with all fish swimming in the same direction. Reacting only to one neighbour at each burst is not sufficient to make a cohesive group that swims in one direction because (i) agents react to only one neighbour at a time and (ii) they always pick the strongest stimulus and this seems to bias the turns sampled, that are on average larger than the experimental ones. Other hypotheses regarding the integration of stimuli may perform better, in particular if taking into account more than 1 fish at a time, such as a reaction averaged over the perceived neighbours. Nevertheless, our hypothesis might be relevant for other gregarious species that do not school, such as the zebrafish. In our simulations, we find that agents are in a shoaling state, that is reacting to their neighbour, mainly swimming at short distances from neighbours but not aligned in their heading to neighbours. These results shed light on the importance of the assumptions chosen to model how fish integrate stimuli: given the same interaction functions, different assumptions regarding the integration of stimuli lead to different collective states (shoaling or schooling).