HAL
open science

Experimental analysis and modelling of the behavioural interactions underlying the coordination of collective motion and the propagation of information in fish schools

Valentin Lecheval

To cite this version:

Valentin Lecheval. Experimental analysis and modelling of the behavioural interactions underlying the coordination of collective motion and the propagation of information in fish schools. Ecology, environment. Université Paul Sabatier - Toulouse III, 2017. English. NNT : 2017TOU30361 . tel-02287362

HAL Id: tel-02287362

https://theses.hal.science/tel-02287362v1

Submitted on 13 Sep 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*
Cotutelle internationale *University of Groningen*

Présentée et soutenue le *mardi 5 décembre 2017 (05/12/2017)* **par :**

Valentin Lecheval

Experimental analysis and modelling of the behavioural interactions underlying the coordination of collective motion and the propagation of information in fish schools

---

JURY

Nicolas Destainville	Professeur	Président du Jury
Simon Verhulst	Professeur	Membre du Jury
Rineke Verbrugge	Professeur	Membre du Jury
José Halloy	Professeur	Rapporteur
Christos C. Ioannou	Docteur	Rapporteur
Colin J. Torney	Docteur	Rapporteur

École doctorale et spécialité :
SEVAB : Écologie, biodiversité et évolution
Unité de Recherche :
Centre de Recherches sur la Cognition Animale (UMR 5169)
Directeur(s) de Thèse :
Charlotte K. Hemelrijk et *Guy Theraulaz*
Rapporteurs :
José Halloy, *Christos C. Ioannou* et *Colin J. Torney*

Acknowledgements

The research presented in this thesis is the result of a collective effort that has emerged from interactions between many individuals to whom I am extremely grateful.

First, I would like to warmly thank Guy Theraulaz and Charlotte Hemelrijk. They have initiated a fruitful collaboration between two research teams in Toulouse and Groningen, that investigate moving animal groups with different approaches and methods. I am very pleased for having been part of this project during three years, under their demanding and enriching supervision. I would also like to thank Clément Sire for all the work he dedicated to this thesis and for his inspired and illuminating assistance regarding analysis and modelling.

Thank you to all members of the research teams involved in this work, that is staff, researchers, students and friends of the Centre de Recherches sur la Cognition Animale (in particular the CAB[1] and IVEP[2] teams) in Toulouse as well as the BPE[3] and TRÊS[4] teams in Groningen. As for the technical support, I would like to stress and thank the very important contributions of Hanno Hildenbrandt, Patrick Arrufat, Stéphane Ferrere, Mathieu Moreau, Maud Combe and Gérard Latil who made all sorts of ideas possible. Thank you to the two master students who have been working with us in Toulouse, Pierre Tichit, Marlene Afflerbach and Alexandra Litchinko. Their dedication to this project has been amazing and their contribution substantial. I would also like to thank Jean-Baptiste Ferdy and Christophe Eloy for their time, ideas and inputs to this work, during my annual PhD committee. Thank you also to Li Jiang and Emanuele Crosato and their respective teams for their very interesting studies based on the empirical work of this thesis as well as for their friendship, that was of main importance during the talks we all gave for the first time in congresses.

The teaching dimension of my position as a PhD student is unfortunately missing from this thesis, but I would like to thank the ∼1000 students I have had over three years of what became an intensive and inspiring teaching duty, masterfully assisted by Sergine Ponsard and Laëtitia Buisson.

Finally, I would like to thank Jacques Gautrais, Richard Fournier, Stéphane Blanco and Anaïs Khuong for their supervision of my first research project, a few years ago. They have not only initiated my interests in collective behaviours and in modelling but also have deeply shaped my research and teaching practises.

[1] Collective Animal Behaviour
[2] Interindividual Variability and Emergent Plasticity
[3] Behavioural and Physiological Ecology
[4] Theoretical Research in Evolutionary Λife Sciences

Summary

Collective motion is ubiquitous in fish. Large groups may comprise thousands to millions of individuals. However, it is not yet understood what are the local interaction rules that underlie these collective patterns.

In this thesis, we investigate in a small gregarious fish, the rummy-nose tetra (*Hemigrammus rhodostomus*), the mechanisms underlying both the coordination of motion and the propagation of information in their schools. To discover the connection between individual behaviour and patterns at the collective level, we follow an approach that tightly combines experiments and modelling.

The first part of the thesis is dedicated to the behavioural mechanisms underlying the coordinated swimming in these schools. For this, we analysed the motion of a single individual and pairs of individuals swimming freely in a circular arena. *H. rhodostomus* has a burst-and-coast swimming style. This consists of sudden changes of heading combined with brief accelerations followed by quasi-passive, straight decelerations. We segmented their trajectories on the basis of this intermittent swimming mode. This segmentation determined the analysis of experimental data and the design of the model. We developed a new method to measure and disentangle the interactions between a fish and the wall and between pairs of fish. We tested these findings with a model derived from physical analogies and symmetry considerations. Our results support the presence of interactions among fish based on the coexistence of attraction and alignment.

We also investigated for fish swimming in a ring-shaped tank how an individual integrates the information coming from several other fish and from obstacles in its neighbourhood. For this, we develop a computational model based on maps of behavioural actions that were extracted from empirical data. We tested in this model whether the properties of schools at the global-level were reproduced when fish reacted only to the strongest stimulus they perceived. We rejected this for observations in groups of 5 fish.

The second part of the thesis is dedicated to the analysis of the propagation of information in reaction to internal and external perturbations, in schools of *H. rhodostomus*. The internal perturbations concern spontaneous collective U-turns occurring in a ring-shaped tank. The global properties of the propagation came from empirical data of group sizes with 1 to 20 fish. We formulate an Ising-spin model that integrates both asymmetrical interactions among fish and the tendency of individuals to follow the majority of their neighbours. The model shows that local social conformity may be underlying both the dynamics observed during the collective U-turns and the sharp decrease of the frequency of collective U-turns as the size of the group increases.

Finally, we develop a preliminary experimental method to induce controlled external perturbations with the objective to investigate the propagation of information during disturbances. We show that aversive conditioning (i) can be performed in this species, (ii) triggers collective escape reactions and (iii) transfers from the training condition to a new experimental set-up. These results come from a training in which a green light elicits an escape reaction in conditioned fish. Our findings suggest that the proportion of conditioned individuals in a group is critical in triggering collective escape reactions.

Résumé

Les déplacements collectifs sont omniprésents chez les poissons et peuvent comprendre des groupes de plusieurs milliers ou millions d'individus. Cependant, les interactions locales et les règles comportementales individuelles à l'origine de ces comportements collectifs posent encore question.

Dans cette thèse, nous étudions les mécanismes sous-jacents à la coordination du déplacement et à la propagation de l'information dans les bancs d'un petit poisson grégaire, le nez rouge (*Hemigrammus rhodostomus*). L'approche de cette thèse repose sur une étroite combinaison entre les méthodes expérimentales et de modélisation dans l'objectif de découvrir les liens entre les comportements individuels et les motifs observés à l'échelle collective.

La première partie du manuscrit est dédiée aux mécanismes comportementaux qui sous-tendent la coordination de la nage en banc chez *H. rhodostomus*. Nous y analysons le déplacement d'un individu solitaire ou d'une paire de poissons nageant librement dans une arène circulaire. *H. rhodostomus* a un comportement de nage intermittente. Ce mode de nage est constitué de changements de direction soudains combinés à de brèves saccades natatoires suivies d'une phase d'accostage durant laquelle le corps du poisson reste rigide. L'analyse des données expérimentales ainsi que le travail de modélisation reposent sur la segmentation des trajectoires des poissons sur la base du comportement de nage intermittente. Nous avons développé une nouvelle méthode pour mesurer et isoler les interactions d'un poisson avec un mur ou entre deux poissons. Ces mesures sont testées à l'aide d'un modèle inspiré par des considérations de physique et de symétrie. Nos résultats soutiennent la présence d'interactions reposant sur la coexistence d'attraction et d'alignement.

Nous étudions aussi comment les individus intègrent l'information issue de plusieurs congénères et d'obstacles dans leur voisinage. Nous formulons un modèle computationnel qui repose sur des cartes d'actions comportementales extraites des données expérimentales. Nous testons l'hypothèse selon laquelle les poissons réagissent uniquement au stimulus le plus important qu'ils perçoivent, en vérifiant que l'hypothèse permet de reproduire les propriétés du banc à l'échelle collective. Nous concluons que cette hypothèse n'est pas suffisante pour reproduire les caractéristiques de bancs formés dans des expériences réunissant 5 poissons dans un dispositif annulaire. La seconde partie du manuscrit est dédiée à l'analyse de la propagation d'information dans les bancs de *H. rhodostomus*, en réaction à des perturbations internes ou externes.

Nous analysons la propagation de l'information en réponse à des perturbations internes se produisant lors de demi-tours collectifs spontanés observés dans un dispositif annulaire, avec des groupes allant de 1 à 20 poissons. Les propriétés à l'échelle globale de la propagation sont déduites depuis des données expérimentales pour des tailles de groupes de 1 à 20 poissons. Nous formulons un modèle d'Ising qui intègre les interactions asymétriques entre les poissons ainsi que leur tendance à suivre la majorité de leurs voisins. Le modèle montre que la conformité sociale, appliquée localement, est un mécanisme possible pour expliquer à la fois les dynamiques observées durant les demi-tours collectifs ainsi que la diminution de la fréquence des ces événements quand la taille de groupe augmente.

Enfin, nous proposons une méthode expérimentale pour induire des perturbations externes contrôlées dans le but d'étudier la propagation de l'information au sein du banc dans ce contexte. Nous réalisons une étude préliminaire montrant que le conditionnement aversif peut être (i) réalisé avec cette espèce, (ii) qu'il peut déclencher des réactions de fuite collective et (iii) que l'apprentissage peut être transféré dans un autre dispositif. Ces résultats sont discutés dans le contexte de la propagation de l'information en réaction à un stimulus externe (ici, une lumière verte impliquant

une réaction de fuite chez les individus conditionnés). Nos travaux suggèrent que la proportion d'individus conditionnés est un paramètre critique au sein du banc pour qu'une fuite collective se produise en réponse à des stimuli externes.

Samenvatting

Gezamenlijke coördinatie van bewegingen is alomtegenwoordig in vissen. Grote visscholen kunnen duizenden tot miljoenen dieren bevatten. Het is echter onbekend welke locale gedragsregels deze collective gedragspatronen aansturen.

In dit proefschrift onderzoeken we in de roodneuszalm (*Hemigrammus rhodostomus*), welke mechanismen bijdragen aan de coördinatie van hun school en de transmissie van informatie. Om de samenhang tussen individueel gedrag en collectieve patronen te achterhalen, combineren we empirisch onderzoek en computermodellen in onze benadering. Het eerste deel van het proefschrift onderzoekt de gedragsmechanismen die ten grondslag liggen aan de coördinatie van een school in deze soort. Hiertoe analyseren we de voortbeweging van een of twee vissen als ze rondbewegen in een cylindervormige aquarium. Deze vissoort vertoont een "burst-en-coast" zwemstijl. Dit bestaat uit korte acceleraties met een verandering van richting gevolgd door een glijbeweging waarin de snelheid afneemt. We segmenteerden de trajectoriën overeenkomstig en baseerden ook het computermodel op deze zwemstijl. We ontwikkelden een nieuwe methode om de interacties tussen vissen en tussen een vis en de wand te meten en te ontrafelen. We baseerden ons model op natuurkundige analogieën en overwegingen van symmetrie. Resultaten laten zien dat de interacties in de school zowel aantrekking tussen vissen als het gelijkrichten van hun oriëntatie behelzen. We onderzochten voor vissen die zwemmen in een ringvormig aquarium ook hoe een individu verschillende typen informatie integreert, namelijk die van andere vissen en van nabije obstakels. Hiertoe ontwikkelden we een computermodel gebaseerd op kaarten van gedragsreacties die we extraheerden uit empirische data. We onderzochten in dit model of de globale eigenschappen van de school ontstonden als de individuen alleen maar reageerden op de sterkste stimulus die ze waarnamen. Deze hypothese moesten we verwerpen voor scholen van vijf vissen. Het tweede deel van mijn proefschrift wijd ik aan het onderzoek naar het doorgeven van informatie in scholen als reactie op interne en externe verstoringen. De transmissie in reactie op interne verstoringen betrof het voorkomen van spontane gemeenschappelijke omkeringen van de school (U-turns) in een ringvormig aquarium. Empirische data over de globale eigenschappen van de zwerm kwamen van scholen van 1 tot 20 indiviuen. We formuleren een Ising-spin model. Het integreert zowel asymetrische interacties tussen individuen onderling als de neidg van individual om de meerderheid van hun nabije buren te volgen. Het model toont dat locale sociale overeenstemming ten grondslag kan liggen aan zowel de dynamiek van de gezamenlijke omkeringen (U-turns) als hun sterke afname in frequentie bij toenemende groepsgrootte. Tenslotte, ontwikkelden we een voorlopige experimentele methode waarmee we gecontroleerd scholen verstoren met het doel om de doorgave van informatie tijdens deze verstoring te onderzoeken. We laten zien dat aversief conditioneren (1) succesvol is in deze soort, (2) collective ontsnappingsreacties opwekt, (3) overdraagbaar is van de training-sessie naar een nieuwe experimentele set-up. Deze resultaten zijn gebaseerd op een training waarin een groen licht een vluchtreactie opwekt. Onze resultaten tonen aan dat de proportie van gecondioneerde vissen in een groep doorslaggevend is voor het al of niet ontstaan van een gezamenlijke ontsnappingsreactie.

Contents

Acknowledgements **i**

Summary **iii**

Résumé **v**

Samenvatting **vii**

1 General introduction **1**
1.1 Collective motion in fish 1
1.2 Analysing collective motion in fish 2
1.3 Propagation of information in animal groups 4
1.4 Communication in fish schools 6
1.5 Thesis overview 7

I What are the individual-level interactions and behavioural rules that give rise to coordinated swimming? **11**

2 Disentangling and modelling interactions in fish with burst-and-coast swimming **13**
Daniel S. Calovi, Alexandra Litchinko, Valentin Lecheval, Ugo Lopez, Alfonso Pérez Escudero, Hugues Chaté, Clément Sire, Guy Theraulaz
2.1 Introduction 13
2.2 Results 15
2.3 Discussion and conclusion 24
Appendix 2.A Intelligent and dumb active matter 26
Appendix 2.B Experimental procedures and data collection 27
Appendix 2.C Data extraction and pre-processing 28
Appendix 2.D Analysis of the interactions 31
Appendix 2.E Parameter estimation and simulations 33

3 A data-driven method to investigate the integration of information in fish schools **37**
Valentin Lecheval, Hanno Hildenbrandt, Clément Sire, Guy Theraulaz and Charlotte K. Hemelrijk
3.1 Introduction 37
3.2 Material and methods 38
3.3 Results 46
3.4 Discussion 55

II How does information propagate in groups of fish in response to perturbations? **61**

4 Domino-like propagation of collective U-turns in fish schools **63**

Valentin Lecheval, Li Jiang, Pierre Tichit, Clément Sire, Charlotte K. Hemelrijk and Guy Theraulaz
4.1 Introduction . . . 63
4.2 Material and Methods . . . 64
4.3 Results . . . 68
4.4 Modelling collective U-turns . . . 69
4.5 Discussion . . . 71
Appendix 4.A Experimental procedures & data collection . . . 73
Appendix 4.B Supplemental figures . . . 76

5 Conditioning an avoidance response in groups of rummy-nose tetra (*Hemigrammus rhodostomus*) 89
Valentin Lecheval, Patrick Arrufat, Stéphane Ferrere, Charlotte K. Hemelrijk and Guy Theraulaz
5.1 Introduction . . . 89
5.2 Material and methods . . . 90
5.3 Results . . . 93
5.4 Discussion . . . 98

6 General discussion 101
6.1 Overview of the main results . . . 101
6.2 Outlook and future work . . . 104

Bibliography 107

Appendix A Informative and misinformative interactions in a school of fish 117
Emanuele Crosato, Li Jiang, Valentin Lecheval, Joseph T. Lizier, X. Rosalind Wang, Pierre Tichit, Guy Theraulaz, Mikhail Prokopenko

Appendix B Identifying influential neighbors in animal flocking 143
Li Jiang, Luca Giuggioli, Andrea Perna, Ramón Escobedo, Valentin Lecheval, Clément Sire, Zhangang Han, Guy Theraulaz

Chapter 1

General introduction

1.1 Collective motion in fish

Collective motion is ubiquitous in fish: it is assumed that 50% (i.e. approximately one fourth of all vertebrates species (IUCN, 2017)) of the 34,515 species of fish known nowadays (Eschmeyer et al., 2017) swim in groups at some point of their life (Shaw, 1978). Gregariousness in fish has even led to very large groups of thousands to millions of individuals (e.g. in herrings) yielding striking examples of collective motion such as bait balls (Figure 1.1).

The organisation of groups of fish is very diverse across species – and varies in time for some species (Tunstrøm et al., 2013). Groups are usually referred to as *shoals*, *swarms* or *schools*. All groups of fish that have aggregative tendencies can be termed as *shoals*. When the group members adopt the same orientation (i.e. they have a tendency to *polarise*) the group is called *a school*. In contrast to *schools*, groups that are loosely structured and whose members have random orientations although maintaining a significant degree of cohesion, are labelled *swarms* (Delcourt and Poncin, 2012, for a review).

In fish, it is commonly assumed that living in groups may improve the efficiency of individual motion (Hemelrijk et al., 2014), foraging (Pitcher et al., 1982) and, most of all, protection against predator threat (Krause and Ruxton, 2002). The latter is achieved thanks to several mechanisms, commonly termed as, among others, dilution of risk (Foster and Treherne, 1981), confusion effect (Ioannou et al., 2007), predator detection (Elgar, 1989) or attack-abatement effect (Turner and Pitcher, 1986). It has been found that predation threat increases the cohesion of fish shoals (Seghers, 1974; Herbert-Read et al., 2017). Still, it is unclear whether collective patterns reported in fish are adaptive or not – that is if they actually increase survival of individuals when a group is attacked for instance (Parrish and Edelstein-Keshet, 1999). In particular, since there is a great diversity of collective patterns, it is possible that some of them are evolutionarily neutral, or even pathological, as the rotational formation (so called *milling*) of army ants (Delsuc, 2003) that is also found in fish (Tunstrøm et al., 2013). Therefore, it seems essential to distinguish biologically relevant features from non-adaptive epiphenomena as well as to describe the causal links between mechanisms at the individual-level and group patterns to improve the understanding of fish aggregations in nature (Parrish et al., 2002).

The multiplicity of the levels at which groups of fish (and of animals in general) can be described makes them complex systems. The relations and feedback loops of the genes, the brain, and the social behaviour are entangled across several scales in time, from organismal development to evolutionary time, and space, from DNA molecules to groups of millions of individuals (Robinson et al., 2008). The study of these challenging systems thus requires to carefully define the extent of the time and spatial scales examined.

It seems necessary to characterise and quantify the interactions between individuals underlying the collective behaviours to make causal links from the neural and cognitive basis of individual behaviours to the collective behaviours in which these neuronal and cognitive processes are involved (Weitz et al., 2012). This thesis aims to investigate the behavioural mechanisms that are involved

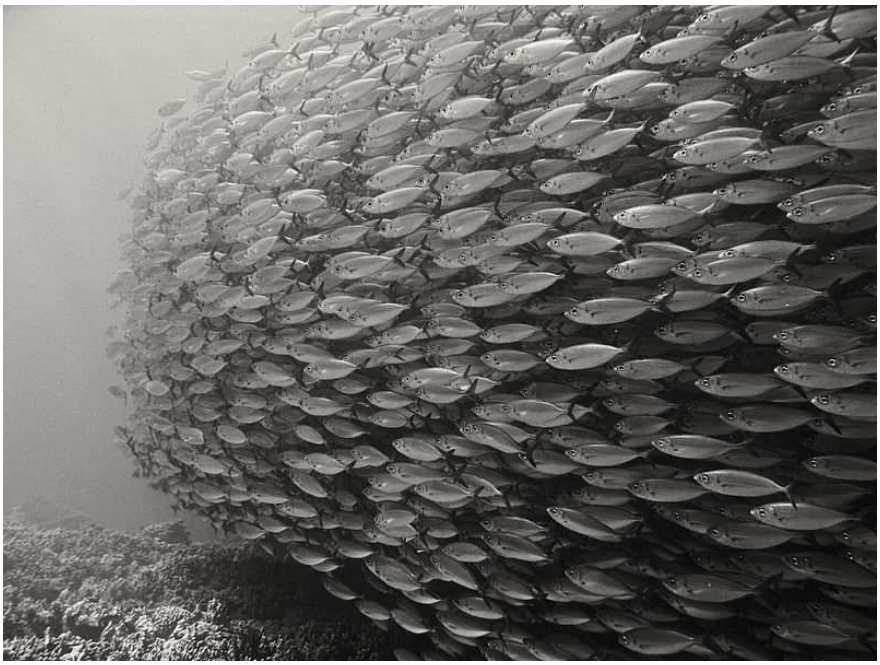


Figure 1.1: Edge of a *Caranx latus* bait ball. Work by Steve Dunleavy published on Flickr (under licence CC BY 2.0).

in the control of the coordination of motion and in the propagation of information within groups of fish.

This thesis consists of two main parts. The first part investigates how schooling behaviour emerges from the interactions between individuals. The second part examines how information propagates in groups of fish in response to internal and external perturbations. In what follows I present the framework of my thesis and a short review of the propagation of information among animals.

1.2 Analysing collective motion in fish

There have been recent developments in experimental and modelling methods in the field of collective behaviour in fish. They are shortly reviewed below.

1.2.1 Modelling methods

It is not necessary, and it is even impossible in large groups, for each fish swimming in shoals to have a complete knowledge of the group properties (such as the average orientation of the group members). It is commonly assumed that collective behaviour in fish is not choreographed by leaders but results from self-organization processes. In these systems, the collective patterns emerge from the local interactions among individuals that only have access to partial information (Bonabeau et al., 1997; Camazine et al., 2001). Models of collective motion in fish therefore investigate how the collective behaviour in a school emerges from assumed local interactions.

Seminal work of the late twentieth century has emphasized theoretical and general (i.e. not restricted to a taxon) mechanisms (see Lopez et al., 2012, for a review). Studies have mainly suggested theoretical rules of interactions between individuals, involving attraction, alignment and repulsion for most of them, and shown that group properties emerged from these rules. As pointed out by several authors, different quantitative combinations of these three basic rules can lead to the same properties observed at the collective scale (Weitz et al., 2012; Lopez et al., 2012). If several initial hypotheses can be compatible with the same properties at the group level, it is thus difficult to shed light on the actual individual mechanisms involved in the coordination of groups for a given

species by looking at the collective behaviour.

As claimed by Weitz et al. (2012) and Lopez et al. (2012), the methodological framework introduced by Gautrais et al. (2009, 2012) is relevant to overcome the difficulty mentioned above. In these two studies, a bottom-up data-driven method of modelling has been introduced. Data-driven modelling implies that every step of the modelling process, i.e. all hypotheses required in the formulation of the model, are validated against data. Bottom-up stands for starting first with a model of motion of a single fish swimming spontaneously, which is validated against experimental data (Gautrais et al., 2009). Then, they used this model and added terms of interaction with a second fish and more (Gautrais et al., 2012). These two studies, by closely combining experimental data and modelling approaches at each step, found mathematical formulations for the stimulus-response functions governing decisions of fish in response to the position and orientation of its neighbours. The authors have found that a gradual weighting between alignment (dominant at short distances) and attraction (dominant at large distances) best accounted for their experimental data. They also reported that the parameters governing these two types of interactions depend on the average speed of fish, leading to an increase in group polarization with swimming speed, a direct consequence of the predominance of alignment at high speed. However, here, the interactions of a fish with the walls of the tank or with other fish were only assumed to take phenomenological functional forms fitting well the experimental data. In other words, the mathematical equations of the interactions were not truly extracted from the experimental data nor derived from a theoretical framework. Thus, a natural question which arises is whether the fair quantitative agreement of the model with experiments actually constitutes an implicit validation of the assumed forms of the interactions. This question is addressed in the first chapter of this manuscript.

Other authors have also tried to measure the interactions in groups of fish from their experimental trajectories but without testing whether their findings could be used in a model to predict group properties (Katz et al., 2011; Herbert-Read et al., 2011).

Recently, other models have included a reconstruction of the visual fields of fish (Strandburg-Peshkin et al., 2013; Rosenthal et al., 2015; Collignon et al., 2016). Although these approaches are promising, they may suffer from a lack of specific experimental validation of the model of the sensory networks.

1.2.2 Data Collection

For a long time, studies of collective motion in fish have suffered from a lack of experimental data. These last decades have seen important improvements in computing efficiency as well as in data storage and quality of video recording. It is thus easier than before to obtain data of collective motion in animal groups and to run computer-intensive simulations of computational models. For instance, the reconstruction of the 3D positions of thousands (up to 2,600) of starlings (*Sturnus vulgaris*) has been done with a stereo-photography method (Ballerini et al., 2008), leading to trajectory reconstruction (Attanasi, Cavagna, Del Castello, Giardina, Grigera, Jelic, Melillo, Parisi, Pohl, Shen and Viale, 2014; Evangelista et al., 2017). This method has also been used with swarms of wild midges ranging in size between 100 and 600 individuals (Attanasi, Cavagna, Del Castello, Giardina, Melillo, Parisi, Pohl, Rossaro, Shen, Silvestri and Viale, 2014).

An important improvement in tracking has recently been achieved (Pérez-Escudero et al., 2014). Common multitracking systems calculate the most likely assignment of identities of individuals by taking into account the previous movements of the animals. These systems generally have problems when two or more individuals cross or touch because it can be difficult to find the correct identities after the point of overlap. The new algorithm suggested by Pérez-Escudero et al. (2014) works by extracting from the video a signature or fingerprint for each individual. These fingerprints are used to identify individuals in each frame, keeping the correct identities even after crossings or occlusions. In contrast to previous methods, this new feature makes the tracking of long videos (e.g. several hours) more efficient with respect to identity matching than before – even if issues of computational time still have to be addressed for large groups (when group size exceeds 20 individuals).

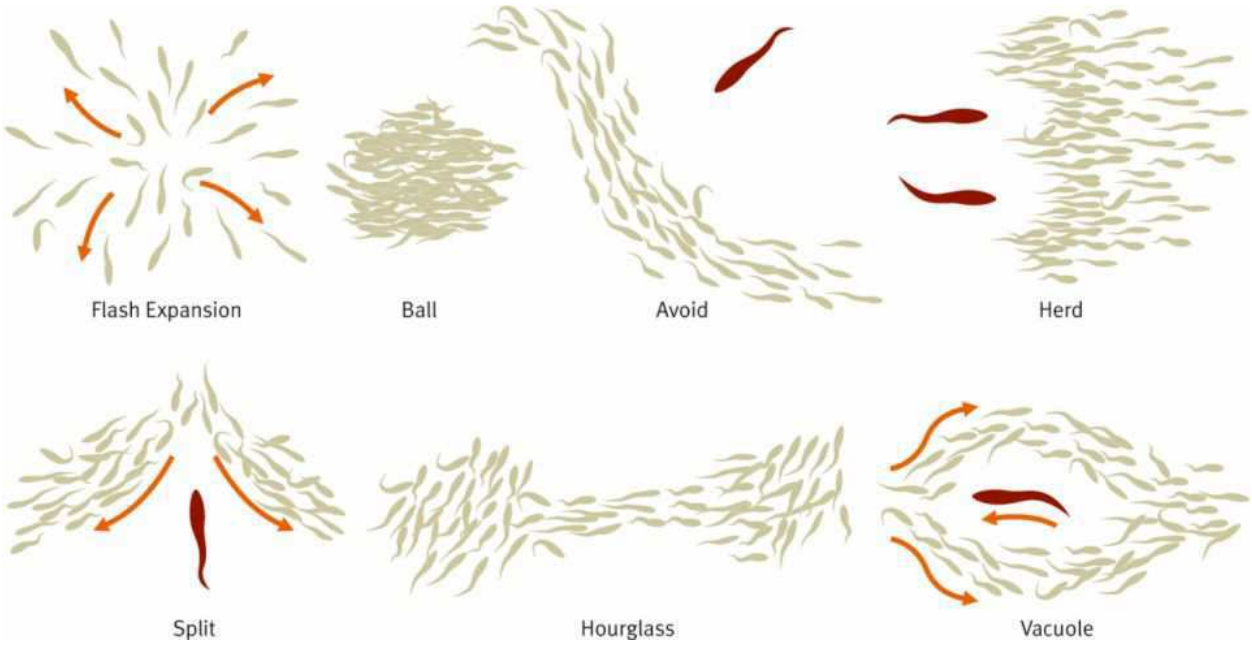


Figure 1.2: Collective responses of fish schools under predator threat (Pitcher and Parrish, 1993). Reproduced from Dugatkin (2013).

1.3 Propagation of information in animal groups

Several patterns of escape have been proposed as survival strategies when groups of fish are attacked by a predator (Pitcher and Wyche, 1983). Pitcher and Wyche (1983) report manoeuvres observed in schools of sand-eels (*Amodytes* sp.) in response to approaches of mackerel (*Scomber scombrus*). These patterns were called *avoid*, *herd*, *vacuole*, *hourglass*, *split - join* and *flash expansion* (see Figure 1.2). They have also been observed in other species such as herrings (*Clupea harengus*) (Pitcher et al., 1996; Nøttestad and Axelsen, 1999; Axelsen et al., 2001). All these patterns were reproduced by computer-simulations (Inada and Kawachi, 2002; Zheng et al., 2005; Lett et al., 2014). The difficulty in studying collective behaviour under predator threat is not only to explain which collective patterns minimize risks of individuals (assuming that the patterns are not all different outcomes of the same behaviour, as suggested by Axelsen et al. (2001) and Inada and Kawachi (2002)) but also to understand how individuals make their choices according to the local information. This requires to investigate how information is propagated, i.e. which cues are shared, which are the individuals sharing information in a school and to describe how fish react to this information.

Results from laboratory experiments showed that, when a perturbation external to the group is applied (an artificial sound stimulus), schools of herring escape by being aligned with their neighbours and going away from the perturbation (Domenici and Batty, 1994). Two modes occur in the distribution of lags between the emission of the stimulus and the reaction of fish: a short lag for fish close to the stimulus and a long lag for fish distant from it. It was also found that the responses with a long latency were more accurate in responding away from the stimulus. The hypothesis of Domenici and Batty (1994) is that the short latency escapes are responses to the sound stimulus and that the long latency escapes are responses to startled neighbours. As for the latter, individuals can integrate the information from both the sound stimulus and startled neighbours and therefore increase the accuracy of their response by adding to the sensory information received by the sound stimulus the swimming direction of startled neighbours. It seems that, in this case, besides the direct emission of the stimulus, social information is also very important for individuals to make accurate decisions and react collectively. Several authors have investigated the individual-level mechanisms that underlie information transfer. Besides differences in their results considering the relative importances of alignment, repulsion and attraction forces (which could be dependent

on the species, the experimental set-ups or the methods of analysis used), all recent studies agree that the speed of individuals is a key element in the information flow of the undisturbed groups (Katz et al., 2011; Herbert-Read et al., 2011; Gautrais et al., 2012) as well as disturbed ones (Herbert-Read et al., 2015). The interactions between fish and thus the properties of the group change with the ecological context (e.g. feeding vs predator threat) (Schaerf et al., 2017). The information flow in a shoal can also be altered by the composition of the group which depends on parameters such as the age, the sex or the numbers of congeners (Hoare and Krause, 2003; Ward et al., 2017). For instance, in adult guppies, it has been found that novel foraging information spreads at a significantly faster rate through subgroups of females than subgroups of males (Reader and Laland, 2000).

When a flotilla of ocean skaters (*Halobates robustus*) is attacked, individuals increase velocities and rate of turning (Treherne and Foster, 1981). This results in a transition from a state where individuals are aligned and moving slowly to a state where individuals are moving rapidly and randomly. This transition of collective behaviour in reaction to predator threat is thought to have two consequences: confusion of the predator because of unpredictable (*protean*) behaviour and fast and synchronised dispersal of the flotilla. This transmission of predator avoidance within the group was faster than the speed of the approach of the predator. Treherne and Foster (1981) called this fast transfer of information the *Trafalgar effect*.

Social waves called shimmering waves also occur in other groups of animals such as the giant honeybees *Apis cerana*, *Apis florea* and *Apis dorsata*. Hundreds of giant honeybees at the surface of their nest (the *bee curtain*) flip their abdomens upward resulting in impressive waves. This behaviour has been linked to a behaviour of defence against attacks by wasps in the species *Apis dorsata* (Kastberger et al., 2008). Two different effects have been shown by Kastberger et al. (2008): repellence of wasps at a distance of at most 50 cm from the nest and confusion of wasps very close to the nest. The fast propagation of the wave within all layers of the bee curtain is achieved thanks to several mechanisms (Kastberger et al., 2014). Most of the shimmering-active bees were acting in a *bucket bridging*-like manner that is receiving information from a close neighbour at one side and transferring it to a close by neighbour at the other side. A small part (about 15%) of the shimmering-active bees elicits abdominal flipping before any bucket-bridging activity can be detected in their neighbourhood, contributing to a *saltatoric* propagation of the wave by creating a *daughter* wave. The result of the saltatoric process is to speed up the propagation as well as to facilitate changes of direction. Waves can also occur without predator attack but this results in short waves only (Kastberger et al., 2008).

Waves in presence of predators can also be termed agitation waves when they involve a sudden change of direction from the group motion. Such waves have for instance been described in birds (Procaccini et al., 2011). As for fish, it has been investigated by monitoring anchovy school (*Engraulis ringens*) movements and their reactions to sea-lion (*Arctocephalus australis* and *Otaria byronia*) attacks in Peruvian waters (Gerlotto et al., 2006). The attacks of sea-lions result in waves of agitation expanding in concentric circles around the sea-lions. Gerlotto et al. (2006) show that the signal of these waves is not damped so that the same information (i.e. the direction of the predator) is transmitted through the entire school, resulting in a reorganized collective structure. Although these collective patterns have been reported independently in the field for several species (see for instance Radakov (1973) and Axelsen et al. (2001) who described a pattern called *density propagation* in herrings (*Clupea harengus*)), the behavioural mechanisms used by individuals in fish schools to propagate these signals are poorly understood. Velocity changes of individuals in response to stimuli (i.e. their speed and their direction) without centralised control are assumed to be essential to propagate escape waves (Herbert-Read et al., 2015).

The principles of the social waves described for flotilla of ocean skaters and giant honeybees as a collective pattern emerging from the local interactions between the agents of a system (that is as a self-organizing pattern) have been modelled in many different situations involving a wide range of living systems. In starling flocks, a model suggests that the agitation waves result from the successive changes of orientation of birds performing escape manoeuvres and not from density

waves (Hemelrijk et al., 2015). In emperor penguin, a model has been used to describe the waves observed in penguins huddles occurring when penguins form dense clusters of thousands of individuals to protect themselves against cold temperatures and wind (Gerum et al., 2013). In this work, the model assumes very simple interactions between individuals: each individual has a preferred distance from its close neighbours that they are trying to maintain. When a perturbation occurs (e.g. a bird moving forward), it triggers a disordering of the group, each individual moving to recover its preferred distance from neighbours, in the same way drivers behave in traffic jams. This model was able to reproduce the collective properties of the waves observed in the field, namely the propagation in all directions, suggesting a mechanism that could make huddles merging.

In many cases, a social wave occurs in a group when individuals exhibit a transition from a state **A** to a state **B** (e.g. the direction of motion of the group that changes during a collective U-turn). A simple and common example of such propagation of information is the Mexican wave, "La Ola", that can be seen, for instance, in many stadiums during sport events (Farkas et al., 2002, 2003). These two papers respectively address the propagation and the initiation of these waves by presenting a model combining local and global interactions unfortunately not derived from a fine-grained analysis of empirical data, i.e. not validating the model at each scale of description. In this example, individuals are modelled as transiting from an inactive state (e.g. people watching the game) to an active state (people standing up being involved in the wave). The model shows that triggering a Mexican wave requires a critical mass of initiators. Other biological examples of state transition leading to a social wave are the landing process in birds (Bhattacharya and Vicsek, 2010), the stop-and-go behaviour of sheep (Pillot et al., 2011; Toulet et al., 2015) or the striking synchronized flashing among fireflies such as *Pteroptyx cribellata* (Camazine et al., 2001, chap. 10, for a review).

1.4 Communication in fish schools

When communication and information transfer in fish groups are investigated, it is important to have some idea about what kind of information a fish perceives, for instance, information about the number and identity of the neighbouring fish that can actually interact with a focal fish. The use of social information enables individuals to coordinate their motion as well as to respond to threats without having to verify the presence of danger independently. What follows is a general picture of how environmental cues may be used by fish when they share information. It is likely to vary from one species to another and to depend on the ecological conditions (light exposure, turbidity, presence of obstacles, etcetera) (Hartman and Abrahams, 2000). The internal mechanisms in the fish brains involved in the interactions with congeners (namely the neuronal scale) are beyond the scope of this thesis.

Fish communicate through various signals related to different sensory systems which can be classified as follows, according to Helfman et al. (2009, chap. 6):

1. Mechanoreception
2. Chemoreception
3. Vision
4. Electroreception
5. Magnetic reception

Mechanoreception involves the lateral line and the inner ear. The lateral line permits the fish to detect disturbances in the water such as currents, prey, predators, congeners and obstacles. It is of main importance when considering predator–prey interactions and fish coordinating in a shoal (Partridge and Pitcher, 1980; Faucher et al., 2010; Polverino et al., 2013). The inner ear detects sound in water.

A **B**

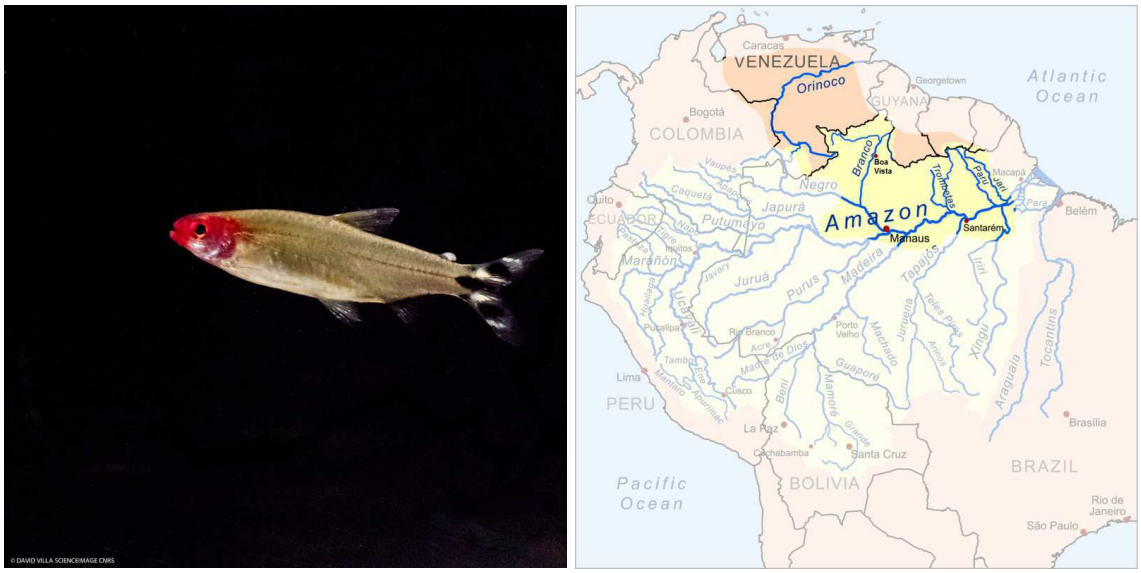


Figure 1.3: A). Photograph of a rummy-nose tetra (*Hemigrammus rhodostomus*) kept in our laboratory. Credits to David Villa ScienceImage/CBI/CNRS, Toulouse. B). Map of the distribution of the rummy-nose tetra (highlighted regions, that correspond to the Orinoco river basin and to the lower Amazon river basin). Adapted from a map made by the user Kmusser on Wikipedia and shared with a CC BY-SA 3.0 licence.

Fish, when inspecting for predators, also rely on chemical substances and visual cues either emitted by the environment (e.g. odour of the predator or visual detection of the predator), or shared (intentionally or not) by congeners (e.g. the chemical alarm substance diffusing from an injured fish or a fish escaping some undetected stimuli with a strong flight behaviour). For instance, although the three-spined stickleback has been classified as a *microsmatic* species that is as a species relying more on vision than on olfaction (Teichmann, 1954; Honkanen and Ekström, 1992), it seems that chemical cues are involved in several processes such as recognition of congeners and foraging (Ward, 2004; Webster et al., 2007). Unlike visual cues, chemical substances might be hard to manipulate for a predator and therefore may be more reliable information for prey (Brown, 2003). However, visual cues, as well as hydrodynamical signals perceived by the lateral line system, are likely to propagate much faster than chemical cues through a shoal (Hunter, 1969; Brown and Laland, 2003). Therefore, it is commonly suggested that the key systems actually used by fish to coordinate their motion are Mechanoreception and Vision.

1.5 Thesis overview

In this thesis, I have investigated the behavioural mechanisms underlying the coordination of motion and the propagation of information in schools of a gregarious fish, the rummy-nose tetra (*Hemigrammus rhodostomus*). This small freshwater fish (mean body-length of $\sim 3\,\mathrm{cm}$) lives in the Lower Amazon River basin in Pará State (in Brazil) and Orinoco River basin in Venezuela (Reis et al., 2003) (Figure 1.3). The *Hemigrammus* taxon is assumed to be non-monophyletic (Marcos Mirande, 2009) and includes 51 species throughout South America (Carvalho et al., 2010). Little has became known about this species since its discovery in 1924 (Ahl, 1924), especially regarding its ecology, despite its success for aquarists. This success in fishkeeping is likely to be the result of the coordination seen in the schools of *H. rhodostomus* (Figure 1.4). Only a few papers have studied the *Hemigrammus* taxon beyond taxonomy and phylogeny. It has been shown in *Hemigrammus bleheri* that the lateral line was essential to the shoaling behaviour (Faucher et al., 2010). How *H. bleheri* swims in pairs or in trios when facing a water flow has also been investigated (Ashraf et al., 2016).

Figure 1.4: A polarised school of swimming rummy-nose tetras. Credits to David Villa ScienceImage/CBI/CNRS, Toulouse.

The choice of *H. rhodostomus* as a model species in our research is supported by (i) their schooling behaviour being obligate, (ii) the ease of buying them (from standard pet shops) and (iii) the ease of keeping them in our facilities.

This thesis will follow an approach based on a tight combination between experiments and modelling to connect individual and collective levels (Camazine et al., 2001; Sumpter et al., 2012; Weitz et al., 2012) that was already initiated by the team in Toulouse (Gautrais et al., 2012; Lopez, 2015). This methodology consists in, given a global pattern, to first focus on experimental observations at the individual level. The findings that, for instance, concern the interactions between animals, are incorporated into data-driven models whose predictions are tested against experimental data at the collective level.

1.5.1 Part I: What are the individual-level interactions and behavioural rules that give rise to coordinated swimming

Part I (Chapters 2 and 3) is dedicated to the behavioural mechanisms that underlie the coordinated swimming in schools of *H. rhodostomus*.

In Chapter 2, we focus on the motion of a single individual and pairs of individuals swimming freely in a circular arena. Fish have been monitored while swimming in circular arenas of different radii. *Hemigrammus rhodostomus* has a burst-and-coast swimming behaviour. This swimming behaviour consists of cyclic bursts of swimming followed by a coast phase in which the body is kept motionless and straight. It is thought to provide hydrodynamic efficiency (Weihs, 1974; Videler and Weihs, 1982). The discretisation of the trajectories on the basis of this intermittent swimming mode drove the analysis of experimental data and the modelling. We developed a new method to measure and disentangle the interactions between a fish and the wall and between pairs of fish and tested these findings in a model. In particular, our findings strongly support the presence of an explicit alignment interaction.

Chapter 3 addresses specifically the question of the integration of information from multiple sources. This issue has rarely been explored in previous studies and current models usually assume reactions averaged over pairwise reactions computed with respect to each separate stimulus, possibly weighted, e.g., by the distance to the stimulus – at the notable exception of (Collignon et al., 2016). As for the latter, the authors develop an interesting hypothesis where fish react by sampling one turning angle from the sum of the probability density functions of turning angles measured in

reaction to each stimulus. Unfortunately, their model does not test the hypothesis specifically and many assumptions (with possible confounding effects) are tested at the same time. We develop a method based on experimental data to test hypotheses regarding the integration of stimuli from multiple sources and we investigate a simple hypothesis in which fish react only to the strongest stimulus we assume they perceive. The method is tested with experimental data in a ring-shaped tank with non-social stimuli (the walls of the corridor) and social stimuli (in groups of 2 and 5 fish). We find that the hypothesis that fish would react only to the strongest stimulus is not sufficient to reproduce the global properties found in experiments with 5 fish, suggesting that fish integrate more information.

1.5.2 Part II: How does information propagate in groups of fish in response to perturbations?

Part II (Chapters 4 and 5) aims to analyse and characterise the propagation of information in schools of *Hemigrammus rhodostomus*, in reaction to *internal* and *external* perturbations. *Internal* and *external* perturbations here refer to whether the stimuli are respectively elicited by a group member or not (e.g. a green light).

In Chapter 4, we analyse and model the propagation of information in response to internal perturbations, i.e. spontaneous collective U-turns occurring in a ring-shaped tank. The global properties of the propagation are characterised from experimental data in group sizes ranging from 1 to 20 fish. We formulate a theory-driven local to global model to explain the main properties of the collective patterns observed. Our model is inspired by the Ising model first suggested in statistical physics to describe ferromagnetism – and one of the simplest statistical models to show a phase transition in 2D (Brush, 1967). The main interest of the model is to show that social conformity is a possible mechanism to explain both the dynamics observed during the collective U-turns and the effect of the group size on the frequency of the collective U-turns.

Chapter 4 is thus a benchmark of the spatio-temporal dynamics of the propagation of information in response to internal and spontaneous perturbations in *Hemigrammus rhodostomus*. In Chapter 5, we develop an experimental method to induce controlled external perturbations in order to investigate the propagation of information in this context. In particular, we conduct a preliminary study showing that aversive conditioning can (i) be used in this species, (ii) trigger collective escape reactions and (iii) be transferred from the conditioning set-up to another experimental set-up. We characterise the aversive conditioning and discuss long-term habituation and forgetting. We discuss these preliminary results in the context of propagation of information in reaction to external stimuli (here, a green light that elicits an escape reaction in conditioned fish). Our findings suggest that the proportion of conditioned individuals in a group is critical to trigger collective escape reactions in response to external stimuli. Our experimental results open promising possibilities regarding the use of conditioning experiments to investigate collective behaviour in fish and the propagation of information within groups in response to perturbations mimicking predatory perturbations in particular.

1.5.3 Appendices

The experimental work conducted in this thesis has been used in two other collaborations summarised in the appendices of the manuscript. In Appendix A, a framework based on Information Theory is used to quantify the dynamics of information transfer in school of fish. This method measures informative and misinformative flows and their spatio-temporal properties during the collective U-turns that occur in the ring-shaped tank. In Appendix B, the identity and respective influences of the neighbours of a focal fish are analysed by studying the short-term directional correlations between their trajectories.

Part I

What are the individual-level interactions and behavioural rules that give rise to coordinated swimming?

Chapter 2

Disentangling and modelling interactions in fish with burst-and-coast swimming

Daniel S. Calovi, Alexandra Litchinko, Valentin Lecheval, Ugo Lopez, Alfonso Pérez Escudero, Hugues Chaté, Clément Sire, Guy Theraulaz

Article published in Calovi, DS et al. 2018. "Disentangling and Modeling Interactions in Fish with Burst-and-Coast Swimming Reveal Distinct Alignment and Attraction Behaviors." PLOS Computational Biology 14 (1): 1–28. https://doi.org/10.1371/journal.pcbi.1005933.

Abstract

We combine extensive data analyses with a modelling approach to measure, disentangle, and reconstruct the actual functional form of interactions involved in the coordination of swimming in Rummy-nose tetra (*Hemigrammus rhodostomus*). This species of fish performs burst-and-coast swimming behaviour that consists of sudden heading changes combined with brief accelerations followed by quasi-passive, straight decelerations. We quantify the spontaneous stochastic behaviour of a fish and the interactions that govern wall avoidance and the attraction and alignment to a neighbouring fish, the latter by exploiting general symmetry constraints for the interactions. In contrast with previous experimental works, we find that both attraction and alignment behaviours control the reaction of fish to a neighbour. We then exploit these results to build a model of spontaneous burst-and-coast swimming and interactions of fish, with all parameters being estimated or directly measured from experiments. This model quantitatively reproduces the key features of the motion and spatial distributions observed in experiments with a single fish and with two fish. This demonstrates the power of our method that exploits large amounts of data for disentangling and fully characterizing the interactions that govern collective behaviours in animals groups. Moreover, we introduce the notions of "dumb" and "intelligent" active matter and emphasize and clarify the strong differences between them.

Contribution of authors

C.S. and G.T. designed research; D.S.C., V.L., U.L., and G.T. performed research; D.S.C., A.L., and C.S. developed the model; D.S.C., A.L., V.L., U.L., H.C., C.S., and G.T. analysed data; A.P.E. contributed new reagents/analytic tools; V.L., C.S., and G.T. wrote the paper.

2.1 Introduction

The study of physical or living self-propelled particles – active matter – has certainly become a booming field, notably involving biologists and physicists, often working together. Physical examples of active matter include self-propelled Janus colloids (Brown and Poon, 2014; Walther and Muller, 2008; Howse et al., 2007; Theurkauff et al., 2012; Palacci et al., 2010, 2013; Buttinoni et al., 2012; Ginot et al., 2015), vibrated granular matter (Narayan et al., 2007; Kudrolli et al., 2008; Deseigne et al., 2010), or self-propulsion mediated by hydrodynamical effects (Thutupalli et al., 2011; Bricard et al., 2013), whereas biological examples are

obviously ubiquitous: bacteria, cells, and simply speaking, most animals. In both physical and biological contexts, active matter can organize into rich collective phases. For instance, fish schools can be observed in a disordered swarming phase, or ordered schooling and vortex/milling phases (Tunstrøm et al., 2013; Calovi et al., 2014).

Yet, there are important differences between "dumb" and "intelligent" active matter (see the Appendix 2.A for a more formal definition and discussion). For the former class, which concerns most physical self-propelled particles, but also, in some context, living active matter, interactions with other particles or obstacles do not modify the intrinsic or "desired" velocity of the particles but exert forces whose effect *adds up to its intrinsic velocity*. Intelligent active matter, like fish, birds, or humans, can also interact through physical forces (a human physically pushing another one or bumping into a wall) but mostly interact through "social forces". For instance, a fish or a human wishing to avoid a physical obstacle or another animal will *modify its intrinsic velocity* in order to never actually touch it. Moreover, a physical force applied to an intelligent active particle, in addition to its direct impact, can elicit a response in the form of a change in its intrinsic velocity (for instance, a human deciding to escape or resist another human physically pushing her/him). Social forces strongly break the Newtonian law of action and reaction: a fish or a human avoiding a physical obstacle obviously does not exert a contrary force on the obstacle. In addition, even between two animals 1 and 2, the force exerted by 1 on 2 is most often not the opposite of the force exerted by 2 on 1, since social forces commonly depend on stimuli (vision, hearing...) associated to an anisotropic perception: a human will most often react more to another human ahead than behind her/him. Similarly, social forces between two fish or two humans will also depend on their relative velocities or orientations: the need to avoid another animal will be in general greater when a collision is imminent than if it is unlikely, due to the velocity directions.

Hence, if the understanding of the social interactions that underlie the collective behaviour of animal groups is a central question in ethology and behavioural ecology (Camazine et al., 2001; Giardina, 2008), it has also a clear conceptual interest for physicists, since social and physical forces play very different roles in the dynamics of an active matter particle (see Appendix for details).

These social interactions play a key role in the ability of group members to coordinate their actions and collectively solve a wide range of problems, thus increasing their fitness (Sumpter, 2010; Krause and Ruxton, 2002). In the past few years, the development of new methods based on machine learning algorithms for automating tracking and behaviour analyses of animals in groups has improved to unprecedented levels the precision of available data on social interactions (Branson et al., 2009; Pérez-Escudero et al., 2014; Dell et al., 2014). A wide variety of biological systems have been investigated using such methods, from swarms of insects (Buhl et al., 2006; Attanasi, Cavagna, Del Castello, Giardina, Melillo, Parisi, Pohl, Rossaro, Shen, Silvestri and Viale, 2014; Schneider and Levine, 2014) to schools of fish (Katz et al., 2011; Herbert-Read et al., 2011; Gautrais et al., 2012; Mwaffo et al., 2015), flocks of birds (Ballerini et al., 2008; Nagy et al., 2010; Bialek et al., 2014), groups of mice (de Chaumont et al., 2012; Shemesh et al., 2013), herds of ungulates (Ginelli et al., 2015; King et al., 2012), groups of primates (Strandburg-Peshkin et al., 2015; Ballesta et al., 2014), and human crowds (Moussaïd et al., 2011; Gallup et al., 2012), bringing new insights on behavioural interactions and their consequences on collective behaviour.

The fine-scale analysis of individual-level interactions opens up new perspectives to develop quantitative and predictive models of collective behaviour. One major challenge is to accurately identify the contributions and combination of each interaction involved at individual-level and then to validate with a model their role in the emergent properties at the collective level (Lopez et al., 2012; Herbert-Read, 2016). Several studies on fish schools have explored ways to infer individual-level interactions directly from experimental data. The force-map technique (Katz et al., 2011) and the non-parametric inference technique (Herbert-Read et al., 2011) have been used to estimate from experiments involving groups of two fish the effective turning and speeding forces experienced by an individual. In the force-map approach, the implicit assumption considers that fish are particles on which the presence of neighbouring fish and physical obstacles exert "forces". Visualizing these effective forces that capture the coarse-grained regularities of actual interactions has been a first step to characterize the local individual-level interactions (Katz et al., 2011; Herbert-Read et al., 2011). However, none of these works incorporate or characterize the intrinsic stochasticity of individual behaviour and nor do they attempt to validate their findings by building trajectories from a model.

On the other hand, only a few models have been developed to connect a detailed quantitative description of individual-level interactions with the emergent dynamics observed at a group level (Herbert-Read et al., 2011; Gautrais et al., 2012; Mwaffo et al., 2015). The main difficulty to build such models comes from the entanglement of interactions between an individual and its physical and social environment. To overcome this problem, Gautrais et al. (2012) have introduced an incremental approach that consists in first building from the experiments a model for the spontaneous motion of an isolated fish (Gautrais et al., 2009). This model is then used as a dynamical framework to include the effects of interactions of that fish with the physical

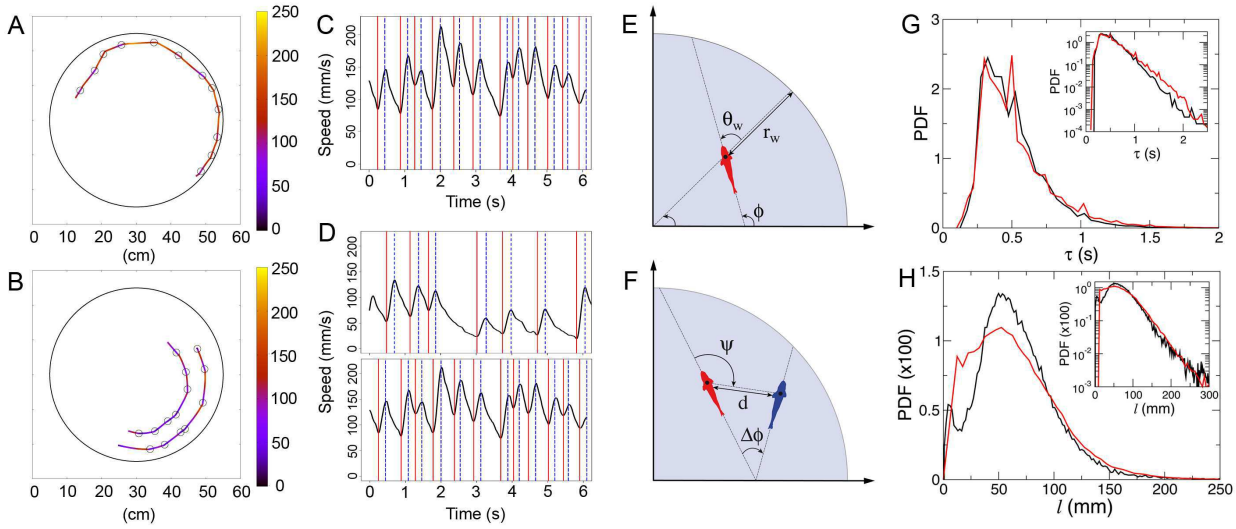


Figure 2.1: Trajectories along with the bursts (circles) of a fish swimming alone (A) and a group of 2 fish (B). The colour of trajectories indicates instantaneous speed. The corresponding speed time series are shown in C and D, along with the acceleration/burst phase delimited by red and blue vertical lines. E defines the variables r_{w} and θ_{w} (distance and relative orientation to the wall) in order to describe the fish interaction with the wall. F defines the relevant variables d, ψ, and $\Delta\phi$ (distance, viewing angle, relative orientation of the focal fish with respect to the other fish) in order to describe the influence of the blue fish on the red one. G and H show respectively the probability distribution function (PDF) of the duration and distance travelled between two kicks as measured in the one (black) and two (red) fish experiments (tank of radius $R = 250\,\mathrm{mm}$). Insets show the corresponding graphs in semi-log scale.

environment and with a neighbouring fish. The validation of the model is then based on the agreement of its predictions with experiments on several observables in different conditions and group sizes.

In the present work, we use but improve and extend this approach to investigate the swimming behaviour and interactions in the red nose fish *Hemigrammus rhodostomus*. This species performs a burst-and-coast type of swimming that makes it possible to analyse a trajectory as a series of discrete behavioural decisions in time and space. This discreteness of trajectories is exploited to characterize the spontaneous motion of a fish, to identify the candidate stimuli (*e.g.* the distance, the orientation and velocity of a neighbouring fish, or the distance and orientation of the tank wall), and to measure their effects on the behavioural response of a fish. We assume rather general forms for the expected repulsive effect of the tank wall and for the repulsive/attractive and alignment interactions between two fish. These forms take into account the fish anisotropic perception of its physical and social environment and must satisfy some specific symmetry constraints which help us to differentiate these interactions and disentangle their relative contributions. The amount and precision of data accumulated in this work and this modelling approach allow us to reconstruct the actual functional form of the response functions of fish governing their heading changes as a function of the distance, orientation, and angular position relative to an obstacle or a neighbour. We show that the implementation of these interactions in a stochastic model of spontaneous burst-and-coast swimming quantitatively reproduces the motion and spatial distributions observed in experiments with a single fish and with two fish.

2.2 Results

2.2.1 Characterization of individual swimming behaviour

Hemigrammus rhodostomus fish have been monitored swimming alone and freely in shallow water in three different circular tanks of radius $R = 176,\ 250,\ 353\,\mathrm{mm}$ (see Supplementary Information (SI) for details). This species performs a burst-and-coast type of swimming characterized by sequences of sudden increase in speed followed by a mostly passive gliding period. This allows the analysis of a trajectory as a series of discrete decisions in time. One can then identify the candidate stimuli (*e.g.* the distance, the orientation and velocity of a neighbouring fish, or the distance and orientation of an obstacle) that have elicited a fish

response and reconstruct the associated stimulus-response function. Most changes in fish heading occur exactly at the onset of the acceleration phase. We label each of these increases as a "kick".

Figures 2.1A and 2.1B show typical trajectories of *H. rhodostomus* swimming alone or in groups of two fish. After the data treatment (see SI and Figure S1 and S2 there), it is possible to identify each kick (delimited by vertical lines in Figures 2.1C and 2.1D), which we use to describe fish trajectories as a group of straight lines between each of these events. While the average duration between kicks is close to 0.5 s for experiments with one or two fish (Figure 2.1G), the mean length covered between two successive kicks is slightly lower for two fish (Figure 2.1H). The typical velocity of the fish in their active periods (see SI) is of order 140 mm/s.

2.2.2 Quantifying the effect of the interaction of a single fish with the wall

Figure 2.2A shows the experimental probability density function (PDF) of the distance to the wall r_{w} after each kick, illustrating that the fish spends most of the time very close to the wall. We will see that the combination of the burst-and-coast nature of the trajectories (segments of average length ~ 70 mm, but smaller when the fish is very close to the wall) and of the narrow distribution of angle changes between kicks (see Figure 2.2D) prevent a fish from efficiently escaping the curved wall of the tank. Figure 2.2C shows the PDF of the relative angle of the fish to the wall θ_{w}, centred near, but clearly below 90°, as the fish remains almost parallel to the wall and most often goes toward it.

In order to characterize the behaviour with respect to the walls, we define the signed angle variation $\delta\phi_+ = \delta\phi \times \mathrm{Sign}(\theta_{\mathrm{w}})$ after each kick, where $\delta\phi$ is the measured angle variation. Therefore, $\delta\phi_+$ is positive when the fish goes away from the wall and negative when the fish is heading towards it. The PDF of $\delta\phi_+$ is wider than a Gaussian and is clearly centred at a positive $\delta\phi_+ \approx 15°$ (tank of radius $R = 353$ mm), illustrating that the fish works at avoiding the wall (Figure 2.2D). When one restricts the data to instances where the fish is at a distance $r_{\mathrm{w}} > 60$ mm from the wall, for which its influence becomes negligible (see Figure 2.4A and the discussion hereafter), the PDF of $\delta\phi_+$ indeed becomes symmetric, independent of the tank in which the fish swims, and takes a quasi Gaussian form of width of order 20° (inset of Figure 2.2D). The various quantities displayed in Figure 2.2 will ultimately be used to calibrate and test the predictions of our model.

2.2.3 Modelling and direct measurement of fish interaction with the wall

We first define a simple model for the spontaneous burst-and-coast motion of a single fish without any wall boundaries, and then introduce the fish-wall interaction, before considering the interaction between two fish in the next subsection. The large amount of data accumulated (more than 300000 recorded kicks for 1 fish, and 200000 for 2 fish; see SI) permits us to not only precisely characterize the interactions, but also to test the model by comparing its results to various experimental quantities which would be very sensitive to a change in model and/or parameters (*e.g.* the full fish-wall and fish-fish distance and angle distributions instead of simply their mean).

Swimming dynamics without any interaction

We model the burst-and-coast motion by a series of instantaneous kicks each followed by a gliding period where fish travel in straight lines with a decaying velocity. At the n-th kick, the fish located at $\vec{x}_n$ at time t_n with angular direction ϕ_n randomly selects a new heading angle ϕ_{n+1}, a start or peak speed v_n, a kick duration τ_n, and a kick length l_n. During the gliding phase, the speed is empirically found to decrease quasi exponentially to a good approximation, as shown in Figure 2.3, with a decay or dissipation time $\tau_0 \approx 0.80$ s, so that knowing v_n and τ_n or v_n and l_n, the third quantity is given by $l_n = v_n \tau_0 (1 - \exp[-\frac{\tau_n}{\tau_0}])$. At the end of the kick, the position and time are updated to

$$\vec{x}_{n+1} = \vec{x}_n + l_n \vec{e}(\phi_{n+1}), \quad t_{n+1} = t_n + \tau_n, \tag{2.1}$$

where $\vec{e}(\phi_{n+1})$ is the unit vector along the new angular direction ϕ_{n+1} of the fish. In practice, we generate v_n and l_n, and hence τ_n from simple model bell-shaped probability density functions (PDF) consistent with the experimental ones shown in Figures 2.1G and 2.1H. In addition, the distribution of $\delta\phi_{\mathrm{R}} = \phi_{n+1} - \phi_n$ (the R subscript stands for "random") is experimentally found to be very close to a Gaussian distribution when the fish is located close to the centre of the tank, *i.e.* when the interaction with the wall is negligible (see the inset of Figure 2.2D). The random variable $\delta\phi_{\mathrm{R}}$ describes the spontaneous decisions of the fish to change its heading:

$$\phi_{n+1} = \phi_n + \delta\phi_{\mathrm{R}} = \phi_n + \gamma_{\mathrm{R}}\, g, \tag{2.2}$$

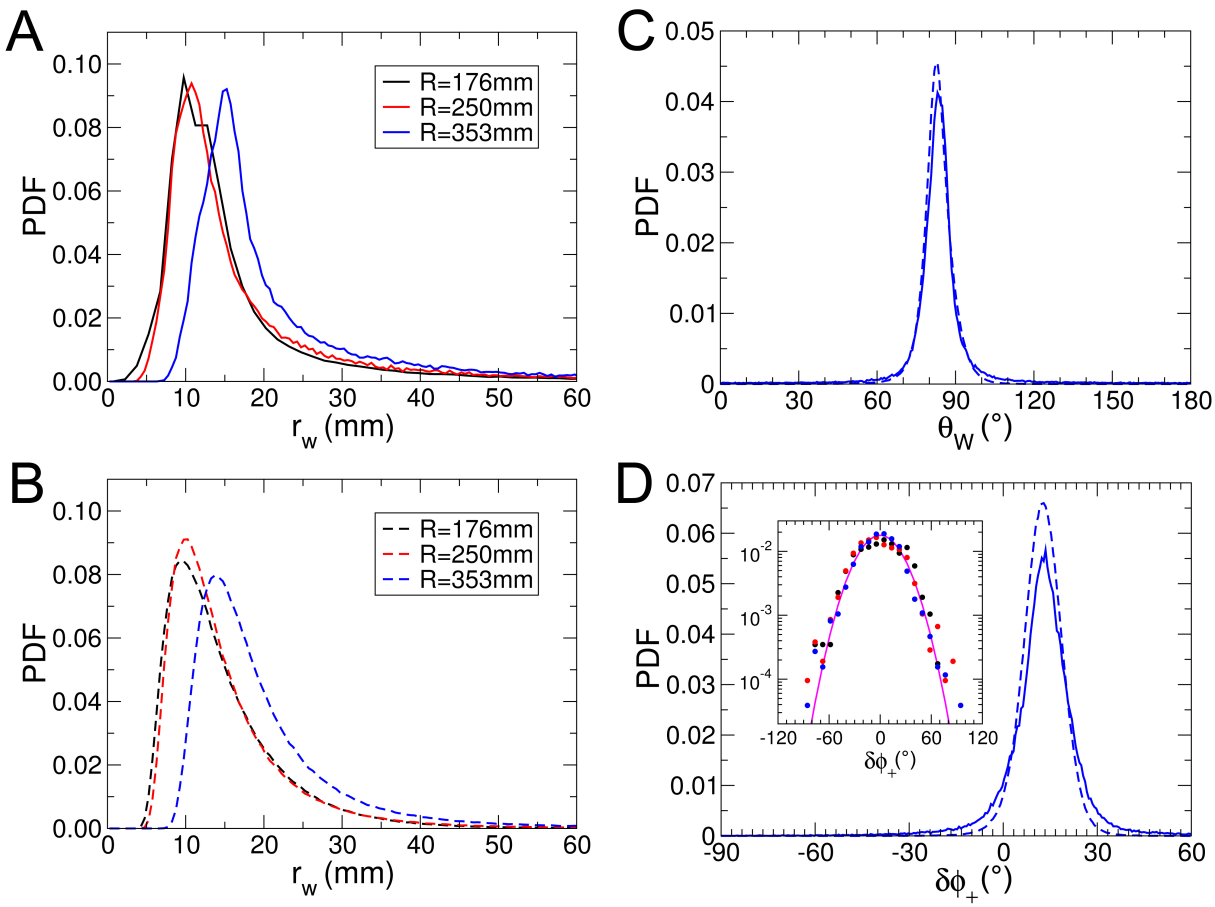


Figure 2.2: Quantification of the spatial distribution and motion of a fish swimming alone. Experimental (A; full lines) and theoretical (B; dashed lines) PDF of the distance to the wall r_{w} after a kick in the three arenas of radius $R = 176$, 250, 353 mm. C: experimental (full line) and theoretical (dashed line) PDF of the relative angle of the fish with the wall θ_{w} ($R = 353$ mm). D: PDF of the signed angle variation $\delta\phi_+ = \delta\phi \times \mathrm{Sign}(\theta_{\mathrm{w}})$ after each kick ($R = 353$ mm). The inset shows the distribution of $\delta\phi_+$ when the fish is near the centre of the tank ($r_{\mathrm{w}} > 60$ mm), for $R = 176$, 250, 353 mm (coloured dots), which becomes centred at $\delta\phi_+ = 0°$ and Gaussian of width $\approx 20°$ (full line).

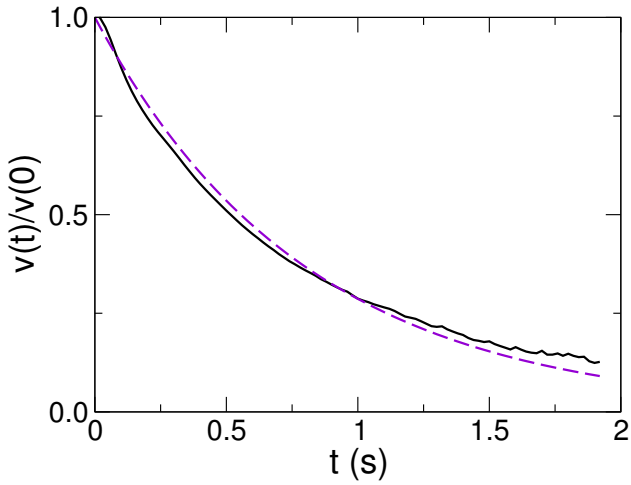


Figure 2.3: Average decay of the fish speed right after a kick (black line), which can be reasonably described by an exponential decay with a relaxation time $\tau_0 \approx 0.80\,\mathrm{s}$ (violet dashed line)

where g is a Gaussian random variable with zero average and unit variance, and γ_R is the intensity of the heading direction fluctuation, which is found to be of order 0.35 radian ($\approx 20°$) in the three tanks.

By exploiting the burst-and-coast dynamics of *H. rhodostomus*, we have defined an effective kick dynamics, of length and duration l_n and τ_n. However, it can be useful to generate the full continuous time dynamics from this discrete dynamics. For instance, such a procedure is necessary to produce "real-time" movies of fish trajectories obtained from the model. As already mentioned, during a kick, the speed is empirically found to decrease exponentially to a good approximation (see Figure 2.3), with a decay or dissipation time $\tau_0 \approx 0.80\,\mathrm{s}$. Between the time t_n and $t_{n+1} = t_n + \tau_n$, the viscous dynamics due to the water drag for $0 \leq t \leq \tau_n$ leads to

$$\vec{x}(t_n + t) = \vec{x}_n + l_n \frac{1 - \exp[-\frac{t}{\tau_0}]}{1 - \exp[-\frac{\tau_n}{\tau_0}]} \vec{e}(\phi_{n+1}), \tag{2.3}$$

so that one recovers $\vec{x}(t_n + \tau_n) = \vec{x}(t_{n+1}) = \vec{x}_n + l_n \vec{e}(\phi_{n+1}) = \vec{x}_{n+1}$.

Fish interaction with the wall

In order to include the interaction of the fish with the wall, we introduce an extra contribution $\delta\phi_\mathrm{W}$

$$\delta\phi = \delta\phi_\mathrm{R}(r_\mathrm{w}) + \delta\phi_\mathrm{W}(r_\mathrm{w}, \theta_\mathrm{w}), \tag{2.4}$$

where, due to symmetry constraints in a circular tank, $\delta\phi_\mathrm{W}$ can only depend on the distance to the wall r_w, and on the angle θ_w between the fish angular direction ϕ and the normal to the wall (pointing from the tank centre to the wall; see Figure 2.1E). We did not observe any statistically relevant left/right asymmetry, which imposes the symmetry condition

$$\delta\phi_\mathrm{W}(r_\mathrm{w}, -\theta_\mathrm{w}) = -\delta\phi_\mathrm{W}(r_\mathrm{w}, \theta_\mathrm{w}). \tag{2.5}$$

The random fluctuations of the fish direction are expected to be reduced when it stands near the wall, as the fish has less room for large angles variations (compare the main plot and the inset of Figure 2.2D), and we now define

$$\delta\phi_\mathrm{R}(r_\mathrm{w}) = \gamma_\mathrm{R}[1 - \alpha f_\mathrm{w}(r_\mathrm{w})]g. \tag{2.6}$$

$f_\mathrm{w}(r_\mathrm{w}) \to 0$, when $r_\mathrm{w} \gg l_\mathrm{w}$ (where l_w sets the range of the wall interaction), recovering the free spontaneous motion in this limit. In addition, we define $f_\mathrm{w}(0) = 1$ so that the fluctuations near the wall are reduced by a factor $1 - \alpha$, which is found experimentally to be close to 1/3, so that $\alpha \approx 2/3$.

If the effective "repulsive force" exerted by the wall on the fish (first considered as a physical particle) tends to make it go toward the centre of the tank, it must take the form $\delta\phi_\mathrm{W}(r_\mathrm{w}, \theta_\mathrm{w}) = \gamma_\mathrm{W} \sin(\theta_\mathrm{w}) f_\mathrm{w}(r_\mathrm{w})$, where the term $\sin(\theta_\mathrm{w})$ is simply the projection of the normal to the wall (*i.e.* the direction of the repulsion "force" due to the wall) on the angular acceleration of the fish (of direction $\phi + 90°$). For the sake of simplicity, $f_\mathrm{w}(r_\mathrm{w})$ is taken as the same function as the one introduced in Equation (2.6), as it satisfies the same limit behaviours. In fact, a fish does not have an isotropic perception of its environment. In order to take into account this important effect in a phenomenological way, we introduce $\epsilon_\mathrm{w}(\theta_\mathrm{w}) = \epsilon_{w,1} \cos(\theta_\mathrm{w}) +$

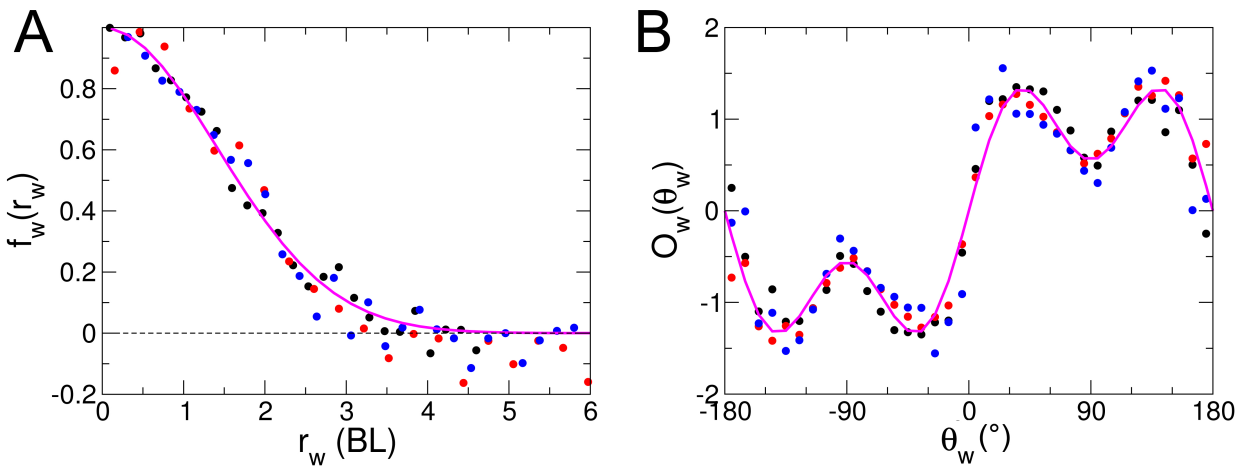


Figure 2.4: Interaction of a fish with the tank wall as a function of its distance r_w (A) and its relative orientation to the wall θ_w(B) as measured experimentally in the three tanks of radius $R = 176\,\mathrm{mm}$ (black), $R = 250\,\mathrm{mm}$ (blue), $R = 353\,\mathrm{mm}$ (red). The full lines correspond to the analytic forms of $f_\mathrm{w}(r_\mathrm{w})$ and $O_\mathrm{w}(\theta_\mathrm{w})$ given in the text. In particular, $f_\mathrm{w}(r_\mathrm{w})$ is well approximated by a Gaussian of width $l_\mathrm{w} \approx 2\,\mathrm{BL}\sim 60\,\mathrm{mm}$.

$\epsilon_{w,2}\cos(2\theta_\mathrm{w}) + ...$, an even function (by symmetry) of θ_w, which, we assume, does not depend on r_w, and finally we define

$$\delta\phi_\mathrm{W}(r_\mathrm{w}, \theta_\mathrm{w}) = \gamma_\mathrm{W}\sin(\theta_\mathrm{w})[1 + \epsilon_\mathrm{w}(\theta_\mathrm{w})]f_\mathrm{w}(r_\mathrm{w}), \tag{2.7}$$

where γ_W is the intensity of the wall repulsion.

Once the displacement l and the total angle change $\delta\phi$ have been generated as explained above, we have to eliminate the instances where the new position of the fish would be outside the tank. More precisely, and since $\vec{x}$ refers to the position of the centre of mass of the fish (and not of its head) before the kick, we introduce a "comfort length" l_c, which must be of the order of one body length (BL; $1\,\mathrm{BL} \sim 30\,\mathrm{mm}$; see SI), and we reject the move if the point $\vec{x}+(l+l_c)\vec{e}(\phi+\delta\phi)$ is outside the tank. When this happens, we regenerate l and $\delta\phi$ (and in particular, its random contribution $\delta\phi_\mathrm{R}$), until the new fish position is inside the tank. Note that in the rare cases where such a valid couple is not found after a large number of iterations (say, 1000), we generate a new value of $\delta\phi_\mathrm{R}$ uniformly drawn in $[-\pi, \pi]$ until a valid solution is obtained. Such a large angle is for instance necessary (and observed experimentally), when the fish happens to approach the wall almost perpendicularly to it ($\delta\phi \sim 90°$ or more).

In order to measure experimentally $\epsilon_\mathrm{w}(\theta_\mathrm{w})$ and $f_\mathrm{w}(r_\mathrm{w})$, and confirm the functional form of Equation (2.7), we define a fitting procedure which is explicitly described in SI, by minimizing the error between the experimental $\delta\phi$ and a general product functional form $\delta\phi_\mathrm{W}(r_\mathrm{w}, \theta_\mathrm{w}) = f_\mathrm{w}(r_\mathrm{w})O_\mathrm{w}(\theta_\mathrm{w})$, where the only constraint is that $O_\mathrm{w}(\theta_\mathrm{w})$ is an *odd* function of θ_w (hence the name O), in order to satisfy the symmetry condition of Equation (2.5). Since multiplying O_w by an arbitrary constant and dividing f_w by the same constant leaves the product unchanged, we normalize O_w (and all angular functions appearing below) such that its average square is unity: $\frac{1}{2\pi}\int_{-\pi}^{+\pi} O_\mathrm{w}^2(\theta_\mathrm{w})\,d\theta_\mathrm{w} = 1$.

For each of the three tanks, the result of this procedure is presented as a scatter plot in Figures 2.4A and 2.4B respectively, along with the simple following functional forms (solid lines)

$$O_\mathrm{w}(\theta_\mathrm{w}) \propto \sin(\theta_\mathrm{w})[1 + 0.7\cos(2\theta_\mathrm{w})], \tag{2.8}$$

$$f_\mathrm{w}(r_\mathrm{w}) = \exp\left[-\left(r_\mathrm{w}/l_\mathrm{w}\right)^2\right], \text{ with } l_\mathrm{w} \approx 2\,\mathrm{BL}. \tag{2.9}$$

Hence, we find that the range of the wall interaction is of order $l_\mathrm{w} \approx 2\,\mathrm{BL} \sim 60\,\mathrm{mm}$, and is strongly reduced when the fish is parallel to the wall (corresponding to a "comfort" situation), illustrated by the deep (*i.e.* lower response) observed for $\theta_\mathrm{w} \approx 90°$ in Figure 2.4B ($\cos(2\theta_\mathrm{w}) \approx -1$). Moreover, we do not find any significant dependence of these functional forms with the radius of the tank, although the interaction strength γ_W is found to decrease as the radius of the wall increases (see Table S3). The smaller the tank radius (of curvature), the more effort is needed by the fish to avoid the wall.

Note that the fitting procedure used to produce the results of Figure 2.4 (described in detail in the SI) does not involve any regularization scheme imposing the scatter plots to fall on actual continuous curves.

The fact that they actually do describe such fairly smooth curves (as we will also find for the interaction functions between two fish; see Figure 2.6) is an implicit validation of our procedure.

In Figure 2.2, and for the three tank radii considered, we compare the distribution of distance to the wall r_w, relative angle to the wall θ_w, and angle change $\delta\phi$ after each kick, as obtained experimentally and in extensive numerical simulations of the model, finding an overall satisfactory agreement. On a more qualitative note, the model fish dynamics mimics fairly well the behaviour and motion of a real fish.

2.2.4 Quantifying the effect of interactions between two fish

Experiments with two fish were performed using the tank of radius $R = 250\,\mathrm{mm}$; and a total of around 200000 kicks were recorded (see SI for details).

In Figure 2.5, we present various experimental PDF which characterize the swimming behaviour of two fish resulting from their interaction, and which will permit to calibrate and test our model. Figure 2.5A shows the PDF of the distance to the wall, for the geometrical "leader" and "follower" fish. The geometrical leader is defined as the fish with the largest viewing angle $|\psi| \in [0, 180°]$ (see Figure 2.1F where the leader is the red fish), that is, the fish which needs to turn the most to directly face the other fish. Note that the geometrical leader is not always the same fish, as they can exchange role. We find that the geometrical leader is much closer to the wall than the follower, as the follower tries to catch up and hence hugs the bend. Still, both fish are farther from the wall than an isolated fish is (see Figure 2.2A). The inset of Figure 2.5A shows the PDF of the distance d between the two fish, illustrating the strong attractive interaction between them.

Figure 2.5C shows the PDF of θ_w for the leader and follower fish, which are again much wider than for an isolated fish (see Figure 2.2C). The leader, being closer and hence more parallel to the wall, displays a sharper distribution than the follower. Figure 2.5B shows the PDF of the relative orientation $\Delta\phi = \phi_2 - \phi_1$ between the two fish, illustrating their tendency to align, along with the PDF of the viewing angle ψ of the follower. Both PDF are found to be very similar and peaked at 0°. Finally, Figure 2.5D shows the PDF (averaged over both fish) of the signed angle variation $\delta\phi_+ = \delta\phi \times \mathrm{Sign}(\theta_\mathrm{w})$ after each kick, which is again much wider than for an isolated fish (Figure 2.2D). Due to their mutual influence, the fish swim farther from the wall than an isolated fish, and the wall constrains less their angular fluctuations.

2.2.5 Modelling and direct measurement of interactions between two fish

In the presence of another fish, the total heading angle change now reads

$$\begin{aligned} \delta\phi \quad = \quad & \delta\phi_\mathrm{R}(r_\mathrm{w}) + \delta\phi_\mathrm{W}(r_\mathrm{w}, \theta_\mathrm{w}) + \\ & \delta\phi_\mathrm{Att}(d, \psi, \Delta\phi) + \delta\phi_\mathrm{Ali}(d, \psi, \Delta\phi), \end{aligned} \tag{2.10}$$

where the random and wall contributions are given by Eqs. (2.6,2.7,2.8,2.9), and the two new contributions result from the expected attraction (Att) and alignment (Ali) interactions between fish. The distance between fish d, the relative position or viewing angle ψ, and the relative orientation angle $\Delta\phi$ are all defined in Figure 2.1F. By mirror symmetry already discussed in the context of the interaction with the wall, one has the exact constraint

$$\delta\phi_{\mathrm{Att,\,Ali}}(d, -\psi, -\Delta\phi) = -\delta\phi_{\mathrm{Att,\,Ali}}(d, \psi, \Delta\phi), \tag{2.11}$$

meaning that a trajectory of the two fish observed from above the tank has the same probability of occurrence as the same trajectory as it appears when viewing it from the bottom of the tank. We hence propose the following product expressions

$$\delta\phi_\mathrm{Att}(d, \psi, \Delta\phi) = F_\mathrm{Att}(d) O_\mathrm{Att}(\psi) E_\mathrm{Att}(\Delta\phi), \tag{2.12}$$

$$\delta\phi_\mathrm{Ali}(d, \psi, \Delta\phi) = F_\mathrm{Ali}(d) O_\mathrm{Ali}(\Delta\phi) E_\mathrm{Ali}(\psi), \tag{2.13}$$

where the functions O are odd, and the functions E are even. For instance, O_Att must be odd as the focal fish should turn by the same angle (but of opposite sign) whether the other fish is at the same angle $|\psi|$ to its left or right. Like in the case of the wall interaction, we normalize the four angular functions appearing in Eqs. (2.12,2.13) such that their average square is unity. Both attraction and alignment interactions clearly break the law of action and reaction, as briefly mentioned in the Introduction and discussed in the Appendix. Although the heading angle difference perceived by the other fish is simply $\Delta\phi' = -\Delta\phi$, its viewing angle ψ' is in general not equal to $-\psi$ (see Figure 2.1F).

As already discussed in the context of the wall interaction, an isotropic radial attraction force between the two fish independent of the relative orientation, would lead exactly to Equation (2.12), with $O_\mathrm{Att}(\psi) \sim \sin(\psi)$ and $E_\mathrm{Att}(\Delta\phi) = 1$. Moreover, an alignment force tending to maximize the scalar product, *i.e.* the alignment,

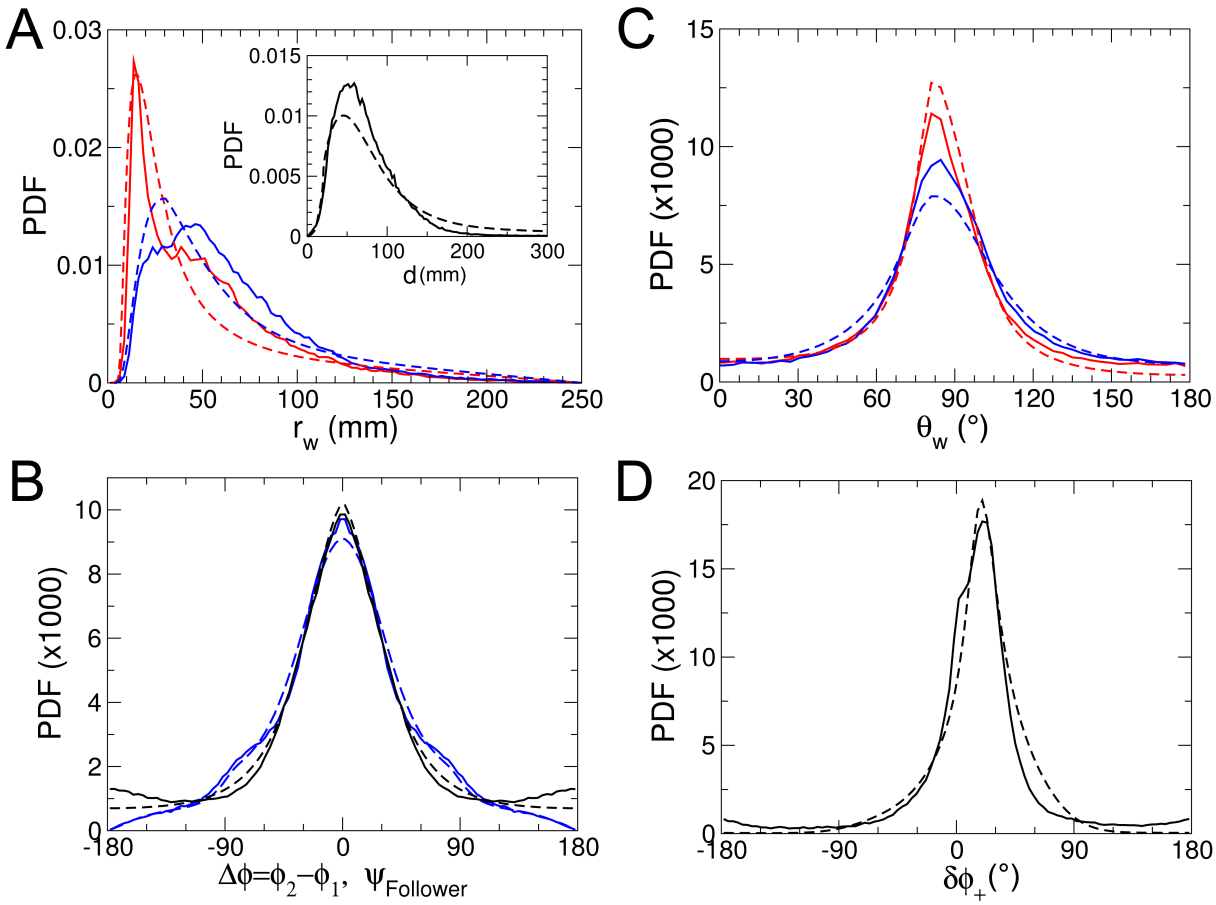


Figure 2.5: Quantification of the spatial distribution and motion in groups of two fish. In all graphs, full lines correspond to experimental results and dashed lines to numerical simulations of the model. A: PDF of the distance to the wall, for the geometrical leader (red) and follower (blue) fish; the inset displays the PDF of the distance d between the two fish. B: PDF of the relative orientation $\Delta\phi = \phi_2 - \phi_1$ between the two fish (black) and PDF of the viewing angle ψ of the follower (blue). C: PDF of the relative angle to the wall θ_{w} for the leader (red) and follower fish (blue). D: PDF (averaged over both fish) of the signed angle variation $\delta\phi_+ = \delta\phi \times \mathrm{Sign}(\theta_{\mathrm{w}})$ after each kick.

between the two fish headings takes the natural form $O_{\rm Ali}(\Delta\phi) \sim \sin(\Delta\phi)$, similar to the one between two magnetic spins, for which one has $E_{\rm Ali}(\psi) = 1$. However, we allow here for more general forms satisfying the required parity properties, due to the fish anisotropic perception of its environment, and to the fact that its behaviour may also be affected by its relative orientation with the other fish. For instance, we anticipate that $E_{\rm Ali}(\psi)$ should be smaller when the other fish is behind the focal fish ($\psi = 180°$; bad perception of the other fish direction) than when it is ahead ($\psi = 0°$).

As for the dependence of $F_{\rm Att}$ with the distance between fish d, we expect $F_{\rm Att}$ to be negative (repulsive interaction) at short distance $d \leq d_0 \sim 1\,{\rm BL}$, and then to grow up to a typical distance $l_{\rm Att}$, before ultimately decaying above $l_{\rm Att}$. Note that if the attraction force is mostly mediated by vision at large distance, it should be proportional to the 2D solid angle produced by the other fish, which decays like $1/d$, for large d. These considerations motivate us to introduce an explicit functional form satisfying all these requirements:

$$F_{\rm Att}(d) \propto \frac{d - d_0}{1 + (d/l_{\rm Att})^2}. \tag{2.14}$$

$F_{\rm Ali}$ should be dominant at short distance, before decaying for d greater than some $l_{\rm Ali}$ defining the range of the alignment interaction. For large distance d, the alignment interaction should be smaller than the attraction force, as it becomes more difficult for the focal fish to estimate the precise relative orientation of the other fish than to simply identify its presence.

Figure 2.6A shows strong evidence for the existence of an alignment interaction. Indeed, we plot the average signed angle change after a kick $\delta\phi_+ = \delta\phi{\times}{\rm Sign}(\psi)$ *vs* $\Delta\phi{\times}{\rm Sign}(\psi)$ and $\delta\phi_+ = \delta\phi{\times}{\rm Sign}(\Delta\phi)$ *vs* $\psi{\times}{\rm Sign}(\Delta\phi)$. In accordance with Eqs. (2.12,2.13), a strong positive $\delta\phi_+$ when the corresponding variable is positive indicates that the fish changes more its heading if it favours mutual alignment (reducing $\Delta\phi$), for the same viewing angle ψ.

As precisely explained in SI (section 2.D), we have determined the six functions appearing in Eqs. (2.12,2.13) by minimizing the error with the measured $\delta\phi$, only considering kicks for which the focal fish was at a distance $r_{\rm w} > 2\,{\rm BL}$ from the wall, in order to eliminate its effect (see Figure 2.4A). This procedure leads to smooth and well behaved measured functions displayed in Figure 2.6. As shown in Figure 2.6B, the functional form of Equation (2.14) adequately describes $F_{\rm Att}(d)$, with $l_{\rm Att} \approx 200\,{\rm mm}$, and with an apparent repulsive regime at very short range, with $d_0 \approx 30\,{\rm mm} \sim 1\,{\rm BL}$. The crossover between a dominant alignment interaction to a dominant attraction interaction is also clear. The blue full line in Figure 2.6B, a guide to the eye reproducing appropriately $F_{\rm Ali}(d)$, corresponds to the phenomenological functional form

$$F_{\rm Ali}(d) \propto (d + d_0')\exp[-(d/l_{\rm Ali})^2], \tag{2.15}$$

with $l_{\rm Ali} \approx 200\,{\rm mm}$. Note that $F_{\rm Att}(d)$ and $F_{\rm Ali}(d)$ cannot be properly measured for $d > 280\,{\rm mm}$ due to the lack of statistics, the two fish remaining most of the time close to each other (see the inset of Figure 2.6A; the typical distance between fish is $d \sim 75\,{\rm mm}$).

Figure 2.6C shows $O_{\rm Att}(\psi) \propto \sin(\psi)[1{+}\epsilon_{\rm Att,1}\cos(\psi){+}...]$ (odd function) and $E_{\rm Att}(\Delta\phi) \propto 1{+}\eta_{\rm Att,1}\cos(\Delta\phi){+}$ ... (even function) along with fits involving no more than 2 non zero Fourier coefficients (and often only one; see SI (section 2.D.2) for their actual values). $E_{\rm Att}(\Delta\phi)$ has a minimum for $\Delta\phi = 0$ indicating that the attraction interaction is reduced when both fish are aligned. Similarly, Figure 2.6D shows $O_{\rm Ali}(\Delta\phi)$ and $E_{\rm Ali}(\psi)$ and the corresponding fits. As anticipated, the alignment interaction is stronger when the influencing fish is ahead of the focal fish ($|\psi| < 90°$), and almost vanishes when it is behind ($\psi = \pm 180°$).

In Figure 2.5, we compare the results of extensive numerical simulations of the model including the interactions between fish to experimental data, finding an overall qualitative and quantitative agreement.

As a conclusion of this section, we would like to discuss the generality of the product functional forms of Eqs. (2.12,2.13) for the interaction between fish, or of Equation (2.7) in the context of the wall interaction. As already briefly mentioned, for a physical point particle interacting through a physical force like gravity, the angle change $\delta\phi_{\rm Att}(d,\psi)$ would be the projection of the radial force onto the angular acceleration (normal to the velocity of angular direction ψ relative to the vector between the two particles) and would then exactly take the form $F_{\rm Att}(d){\times}\sin(\psi)$. Hence, Equation (2.12) (resp. Equation (2.7), for the wall interaction) is the simplest generalization accounting for the fish anisotropic perception of its environment, while keeping a product form and still obeying the left/right symmetry condition of Equation (2.11) (resp. of Equation (2.5)). In principle, $\delta\phi_{\rm Att}(d,\psi,\Delta\phi)$ should be written most generally as an expansion $\sum_i F_{{\rm Att},i}(d)O_{{\rm Att},i}(\psi)E_{{\rm Att},i}(\Delta\phi)$. However, as the number of terms of this expansion increases, we run the risk of overfitting the experimental data by the procedure detailed in the SI, section 2.D. In addition, the leading term of this expansion would still capture the main behavioural effects of the interaction and should be very similar to the results of Figure 2.6, while the weaker remaining terms would anyway be difficult to interpret. Note that the same argument applies to the alignment interaction, when exploiting the analogy with the magnetic alignment force between two spins. Equation (2.13) is the simplest generalization of the

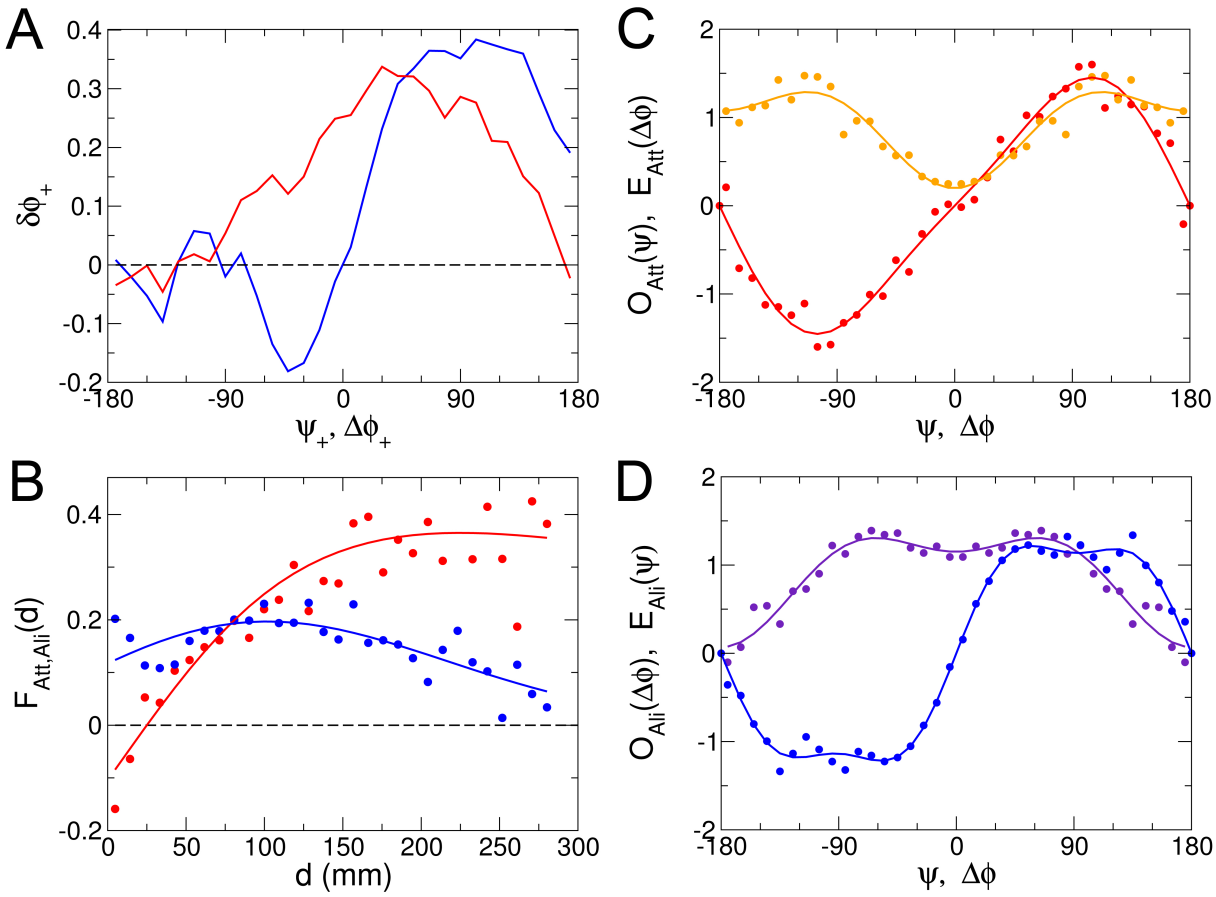


Figure 2.6: Quantification and modelling of interactions between pairs of fish. A: we plot the average signed angle change after a kick $\delta\phi_+ = \delta\phi \times \mathrm{Sign}(\psi)$ *vs* $\Delta\phi \times \mathrm{Sign}(\psi)$ (red) and $\delta\phi_+ = \delta\phi \times \mathrm{Sign}(\Delta\phi)$ *vs* $\psi \times \mathrm{Sign}(\Delta\phi)$ (blue) (see text). B: dependence of the attraction ($F_{\mathrm{Att}}(d)$ in red) and alignment ($F_{\mathrm{Ali}}(d)$ in blue) interactions with the distance d between fish. The full lines correspond to the physically motivated form of Equation (2.14) (red), and the fit proposed in the text for $F_{\mathrm{Ali}}(d)$ (blue). C: $O_{\mathrm{Att}}(\psi)$ (odd function in red) and $E_{\mathrm{Att}}(\Delta\phi)$ (even function in orange) characterize the angular dependence of the attraction interaction, and are defined in Equation (2.12). D: $O_{\mathrm{Ali}}(\Delta\phi)$ (odd function in blue) and $E_{\mathrm{Ali}}(\psi)$ (even function in violet), defined in Equation (2.13), characterize the angular dependence of the alignment interaction. Dots in B, C, and D correspond to the results of applying the procedure explained in SI to extract the interaction functions from experimental data.

interaction $\delta\phi_{\mathrm{Ali}}(d, \Delta\phi) = F_{\mathrm{Ali}}(d)\sin(\Delta\phi)$ obtained in this case, while preserving the left/right symmetry and product form. Considering the fact that no regularization or smoothing procedure was used in our data analysis (see SI), the quality (low noise, especially for angular functions) of the results presented in Figures 2.4 and 2.6 strongly suggests that the generalized product forms used here capture most of the features of the actual experimental angle change.

2.3 Discussion and conclusion

Characterizing the social interactions between individuals as well as their behavioural reactions to the physical environment is a crucial step in our understanding of complex collective dynamics observed in many group-living species and their impact on individual fitness (Camazine et al., 2001; Krause and Ruxton, 2002). In the present work, we have analysed the behavioural responses of a fish to the presence in its neighbourhood of an obstacle and to a conspecific fish. In particular, we used the discrete decisions (kicks) of *H. rhodostomus* to control its heading during burst-and-coast swimming as a proxy to measure and model individual-level interactions. The large amount of data accumulated allowed us to disentangle and quantify the effects of these interactions on fish behaviour with a high level of accuracy.

We have quantified the spontaneous swimming behaviour of a fish and modelled it by a kick dynamics with Gaussian distributed angle changes. We found that the interactions of fish with an obstacle and a neighbouring fish result from the combination of four behavioural modes:

1. wall avoidance, whose effect starts to be effective when the fish is less than 2 BL from a wall;
2. short-range repulsion between fish, when inter-individual distance is less than 30 mm ($\sim$ 1 BL);
3. attraction to the neighbouring fish, which reaches a maximum value around 200 mm ($\sim$ 6 to 7 BL) in our experimental conditions;
4. alignment to the neighbour, which saturates around 100 mm ($\sim$ 3 BL).

In contrast to previous phenomenological models, these behavioural modes are not fixed to discrete and somewhat arbitrary zones of distances in which the neighbouring fish are found (Aoki, 1982; Huth and Wissel, 1992; Couzin et al., 2002). Instead, there is a continuous combination of attraction and alignment as a function of the distance between fish. Alignment dominates attraction up to $\sim$ 75 mm ($\sim$ 2.5 BL) while attraction becomes dominant for larger distances. As distance increases even more, attraction must decrease as well. However, the limited size of the experimental tanks and the lack of sufficient data for large d prevented us from measuring this effect, suggesting the long-range nature of the attraction interaction mediated by vision. Note that a cluster of fish can elicit a higher level of attraction, proportional to the 3D solid angle of the fish group as seen by the focal fish, as suggested by models based on visual perception (Pita et al., 2015; Collignon et al., 2016), and as captured by the power-law decay proposed in Equation (2.14). Designing experiments to test and quantify the long-range nature of the attraction interaction between fish would be of clear interest.

Moreover, the behavioural responses are strongly modulated by the anisotropic perception of fish. The wall repulsion effect is maximum when the orientation of the fish with regards to the wall is close to 45° and minimum when the fish is parallel to the wall. Likewise, the maximum amplitude alignment occurs when a neighbouring fish is located on the front left or right and vanishes as its position around the focal fish moves towards the back.

To quantify separately the effects of attraction and alignment, we exploited physical analogies and symmetry considerations to extract the interactions between a focal fish and the wall and with another fish. Previous studies have shown that in the Golden shiners (Katz et al., 2011) and the Mosquito fish (Herbert-Read et al., 2011), there was no clear evidence for an explicit matching of body orientation. In these species, the alignment between fish was supposed to results from a combination of attraction and repulsion. However, at least in the Mosquito fish, it is likely that the strength of alignment could have been underestimated because the symmetry constraints on alignment and attraction were not taken into consideration. In the Rummy-nose tetra, we find strong evidence for the existence of an explicit alignment.

The characterization and the measurement of burst-and-coast swimming and individual interactions were then used to build and calibrate a model that quantitatively reproduces the dynamics of swimming of fish alone and in groups of two and the consequences of interactions on their spatial and angular distributions. The model shows that the wall avoidance behaviour coupled with the burst-and-coast motion results in an unexpected concentration of fish trajectories close to the wall, as observed in our experiments. In fact, this phenomenon is well referenced experimentally for run-and-tumble swimming (for instance, in sperm cells (Elgeti et al., 2010) or bacteria (Vladescu et al., 2014)). It can be explained theoretically and reproduced in simple models (Tailleur and Cates, 2009; Elgeti and Gompper, 2015), as the effective discreteness of the

trajectories separated in bursts or tumbles prevents the individuals from escaping the wall. Our model also reproduces the alternation of temporary leaders and followers in groups of two fish, the behaviour of the temporary leader being mostly governed by its interactions with the wall, while the temporary follower is mostly influenced by the behaviour of the temporary leader.

This validated model can serve as a basis for testing hypotheses on the combination of influence exerted by multiples neighbours on a focal fish in tanks of arbitrary shape. Moreover, it would certainly be interesting to study theoretically the dynamics of many fish swimming without any boundary and according to the found interactions. The study of the phase diagram as a function of the strength of the attraction and alignment interactions (and possibly their range) should show the emergence of various collective phases (schooling phase, vortex phase...) (Tunstrøm et al., 2013; Calovi et al., 2014).

Finally, our method has proved successful in disentangling and fully characterizing the interactions that govern the behaviour of pairs of animals when large amounts of data are available. Hence, it could be successfully applied to collective motion phenomena occurring in various biological systems at different scales of organization.

Supplementary Information

2.A Intelligent and dumb active matter

A rather general equation describing the dynamics of a standard physical particle moving in a thermal bath (or a medium inducing a friction and a random stochastic force, like a gas) and submitted to physical external forces $\vec{F}_{\text{Phys}}(\vec{x})$ (due to other particles and/or external fields) reads

$$\frac{d\vec{v}}{dt} = -\frac{\vec{v}}{\tau} + \vec{F}_{\text{Phys}} + \sqrt{\frac{2T}{\tau}}\vec{\eta}, \tag{2.16}$$

where $\vec{v} = \frac{d\vec{x}}{dt}$ is the particle velocity, T is the temperature, and $\vec{\eta}(t)$ is a stochastic Gaussian noise, delta-correlated in time, $\langle\vec{\eta}(t)\vec{\eta}(t')\rangle = \delta(t-t')$. In particular, if the physical force is conservative and hence is the gradient of a potential $V_{\text{Phys}}(\vec{x})$, the stationary velocity and position probability distribution of the particle produced by this equation is well known to be the Boltzmann distribution,

$$P(\vec{x},\vec{v}) = \frac{1}{Z}\exp\left(-\frac{E}{T}\right), \tag{2.17}$$

where $E = \frac{v^2}{2} + V_{\text{Phys}}$ is the energy, and Z is a normalization constant.

2.A.1 Dumb active matter

An active particle is characterized by its intrinsic or desired velocity $\vec{u}$. Its actual velocity $\vec{v} = \frac{d\vec{x}}{dt}$ rather generally obeys an equation similar to Equation (2.16):

$$\frac{d\vec{v}}{dt} = -\frac{\vec{v}-\vec{u}}{\tau} + \vec{F}_{\text{Phys}} + \sqrt{\frac{2T}{\tau}}\vec{\eta}, \tag{2.18}$$

where the first term on the right-hand side tends to make the actual velocity go to the intrinsic velocity. Equation (2.18) has to be supplemented with a specific equation for the intrinsic velocity. Here, for the sake of simplicity, we assume that $\vec{u}$ is a simple Ornstein-Uhlenbeck stochastic process,

$$\frac{d\vec{u}}{dt} = -\frac{\vec{u}}{\tau'} + \sqrt{\frac{2T'}{\tau}}\vec{\eta'}, \tag{2.19}$$

where $\vec{\eta'}$ is an other stochastic Gaussian noise, uncorrelated with $\vec{\eta}$, and τ' is some correlation time, a priori unrelated to τ. In general, the stationary distribution $P(\vec{x},\vec{v},\vec{u})$ is not known, although some analytical results can be obtained in some limits (for instance, large friction, and separation of the time scales τ and τ') (Jung and Hänggi, 1987; Fox and Roy, 1987).

A first limiting case of this equation is the strong friction limit (small τ), where the inertial term in Equation (2.18) becomes negligible, leading to

$$\frac{d\vec{x}}{dt} = \vec{v} = \vec{u} + \tau\vec{F}_{\text{Phys}} + \sqrt{2T\tau}\vec{\eta}. \tag{2.20}$$

In the limit of a "cold" medium, where the stochastic force is absent or negligible, we obtain

$$\vec{v} = \vec{u} + \tau\vec{F}_{\text{Phys}}. \tag{2.21}$$

Note that in Eqs. (2.18,2.19,2.20,2.21), the physical force directly impacts the final velocity $\vec{v}$, but *not* the intrinsic velocity $\vec{u}$ of the active particle. This very property constitutes our definition of a "dumb" active particle.

2.A.2 Intelligent active matter

As explained in the Introduction, animals can not only be submitted to physical forces $\vec{F}_{\text{Phys}}$ (*e.g.* a human physically pushing another one), but mostly react to "social forces" $\vec{F}_{\text{Soc}}$. These social interactions *directly affect the intrinsic velocity* of the active particle, which constitutes our definition of an "intelligent" active

particle. In the "cold" limit relevant for fish or humans (the substrate in which they move does not exert any noticeable random force), the system of equations Eqs. (2.18,2.19) becomes

$$\frac{d\vec{v}}{dt} = -\frac{\vec{v}-\vec{u}}{\tau} + \vec{F}_{\text{Phys}}, \tag{2.22}$$

$$\frac{d\vec{u}}{dt} = -\frac{\vec{u}}{\tau'} + \vec{F}_{\text{Soc}} + \sqrt{\frac{2T'}{\tau}}\vec{\eta}', \tag{2.23}$$

In the context of animal and intelligent active matter, the stochastic noise $\vec{\eta}'$ models the spontaneous motion – the "free will" – of the animal (see Equation (2.2), for *Hemigrammus rhodostomus*).

Moreover, we also already mentioned that these social forces are in general non conservative and hence strongly break the action-reaction law, as they generally depend not only on the positions of the particles, but also on their velocities (their relative direction $\Delta\phi$ and the viewing angle ψ and θ_{w} defined in Figure 2.1). In the present work, we have for instance shown how the interaction of *Hemigrammus rhodostomus* with a circular wall depends not only on the distance to the wall, but also on the viewing angle θ_{w} between the fish heading and the normal to the wall (see Figure 2.4). We also determined the dependence of the attraction and alignment interactions on the focal fish viewing angle ψ and the two fish relative heading angle $\Delta\phi$ (see Figure 2.6). Note that physical forces can induce a cognitive reaction and hence a change in the intrinsic velocity, so that $\vec{F}_{\text{Soc}}$ may also contain reaction term to the presence of physical forces $\vec{F}_{\text{Phys}}$ (this was not the case in our experiments, except maybe, when the fish would actually touch the wall). Conversely, social interaction may lead to a particle willingly applying a physical force (a human moving toward another one and then pushing her/him). As a consequence, the notion of a conserved energy and many other properties resulting from the conservative nature of standard physical forces are lost, leading to a much more difficult analytical analysis of the Fokker-Planck equation which can be derived from Eqs. (2.22,2.23).

It is obviously a huge challenge to characterize these social interactions in animal groups, in particular to better understand the collective phenomena emerging in various contexts (Camazine et al., 2001; Giardina, 2008; Sumpter, 2010). The system of equations Eqs. (2.22,2.23), for specific social interactions, also presents a formidable challenge, for instance to determine the stationary distribution $P(\vec{x},\vec{v},\vec{u})$. In the absence of physical forces, and in the limit of fast reaction (small τ), leading to a perfect matching between the velocity and the intrinsic velocity, we obtain

$$\vec{v} = \frac{d\vec{x}}{dt} = \vec{u}, \tag{2.24}$$

$$\frac{d\vec{u}}{dt} = -\frac{\vec{u}}{\tau'} + \vec{F}_{\text{Soc}} + \sqrt{\frac{2T'}{\tau}}\vec{\eta}'. \tag{2.25}$$

Interestingly, this system is formally equivalent to Equation (2.16) for a standard physical particle, although Equation (2.25) is formally an equation for the intrinsic velocity, equal to the actual velocity in the considered limit. Yet, the resulting stationary state $P(\vec{x},\vec{v}=\vec{u})$ is in general not known, because of the non-conservative nature of the social interactions discussed above and in the Introduction.

2.B Experimental procedures and data collection

Ethics statement Our experiments have been approved by the Ethics Committee for Animal Experimentation of the Toulouse Research Federation in Biology N°1 and comply with the European legislation for animal welfare. During the experiments, no mortality occurred.

Study species *Hemigrammus rhodostomus* (rummy-nose tetras, Figure 2.7) were purchased from Amazonie Labège (`http://www.amazonie.com`) in Toulouse, France. This species was chosen because it exhibits a strong schooling behaviour and it is very easy to handle in controlled conditions. Fish were kept in 150 L aquariums on a 12:12 hour, dark:light photoperiod, at 26.8°C (±1.6°C) and were fed *ad libitum* with fish flakes. Body lengths (BL) of the fish used in these experiments were on average 31mm (Table 2.1).

The experimental tank (120×120 cm) was made of glass and was set on top of a box to isolate fish from vibrations. The set-up, placed in a chamber made by four opaque white curtains, was surrounded by four LED light panels giving an isotropic lighting. Circular tanks (of radius $R = 176$, 250, and 353 mm) were set inside the experimental tank 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 26.69°C (±1.19°C) (details in Table 2.1). Reflections of light due to the bottom of the experimental tank are avoided thanks to a white PVC layer. Each trial started by setting one or two fish randomly sampled from their breeding tank into a circular tank.

Fish were let for 10 minutes to habituate before the start of the trial. A trial consisted in one or three hours of fish freely swimming (*i.e.* without any external perturbation) in a circular tank (Tables 2.1 and 2.2). Fish trajectories were recorded by a Sony HandyCam HD camera filming from above the set-up at 50 Hz (50 frames per second) in HDTV resolution (1920×1080p).

Two main sources of uncertainty in the measures from video recorded from above with only one camera occur:

1. by not knowing the water depth at which a fish swims, between 0 to 7 cm from the bottom of the tank;
2. because of parallax issues (the bigger the angle between a swimming fish and the camera axis, the bigger the error made estimating the position of the fish).

The contribution of each source in the uncertainty of our measures has been estimated by computing the lengths of the cells of a chessboard set at the bottom of the tank ($Z = 0$ cm) and at the top of the water level (here $Z = 6$ cm), coming from photographs shot at two zoom levels, the one used to record experiments in the tank of radius $R = 250$ mm and the one used to record experiments in the tank of radius $R = 353$ mm. As a result, the uncertainty due to the unknown position of the fish in the water column is higher than the uncertainty due to parallax (3.5% vs 0.5%).

2.C Data extraction and pre-processing

Positions of fish on each frame have been tracked with the tracking software idTracker 2.1 [5]. The idTracker output format gives fish identity and barycentre positions of individuals in the image (in pixels), where the latter needs to be converted into position in the experimental frame of reference (in millimetres). The intermediary Matlab files issued by the tracker store the background image, which is the information used to calculate wall positions and thus colliding distances more accurately). Also, the intermediary files give the area of the detected fish, which can be used to determine fish heading from shape detection, independently of the trajectories. The processing of the output and intermediary files is processed with a custom-built Matlab script, which is structured into several procedures:

1. Detection of tank walls;
2. Conversion to metric frame of reference;
3. Fish shape detection and body length/width measurements;
4. Fish activity selection and sampling;
5. Segmentation;
6. Segmented variables estimation.

This section aims to document each of these procedures.

2.C.1 Detection of tank walls

From the tracking software idTracker, several files associated with one video are produced. In particular, there is a matrix with as many elements as resolution of videos in pixels (1920×1080 pixels in our case) containing light intensity of each pixel, that is coded in a greyscale image of the video (Figure 2.8 A). The intersection of the bottom of the tank with the tank floor is shadowed (Figure 2.8 A). This shadow is manually enhanced to improve the detection of the tank walls.

Before running the tracking on idTracker, the user has to define a mask in order to exclude areas where individuals (here fish) cannot be tracked (e.g. outside the tank) (Figure 2.8 B). This mask gives a raw circular estimation of the contour of the tank with radius R (*outer circle*) and a second raw circular estimation with arbitrary radius $0.85 \times R$ (*inner circle*) is derived. For N_θ radii with angle $\theta_i \in [0, 2\pi)$, $i \in [1, 2, ..., N_\theta]$, light intensity is measured from greyscale image every pixel from inner circle to outer circle. To measure the light intensity of a pixel, the average over the focal pixel and its 8 nearest neighbours in the image matrix is considered. We take $N_\theta = 5 \times 360 = 1800$ to oversample these measures. Mean position (x_i, y_i) of the three smallest values of light intensity (*i.e.* darker values) associated with each θ_i is computed, yielding a noisy and discrete estimate of lower image positions of the tank walls. A first smoothing procedure is run on the positions (x_i, y_i) to exclude bad walls estimates (e.g. detection of rust in the upper edge of the wall). A raw centre (x_0, y_0) is defined as the mean position of all the (x_i, y_i) and estimate N_θ radii r_i (in pixels) for each θ_i. The following criterion is used: two consecutive radii cannot differ by more than 5 pixels. If it

is the case for the ith radius, it is replaced by the previous one and associated (x_{i-1}, y_{i-1}) are recomputed given θ_i and the new radius r_i (in pixels). This procedure gives a new series of radii r_i where the previous procedure excludes outliers (Figure2.8 D). Assuming that fish swim at constant height h from the bottom, we analytically calculate from the detected positions of tank walls at the bottom the positions of the tank walls at height h using the formula

$$\vec{r}_h = \left(1 + \frac{h}{h+D}\right)(\vec{r} - \vec{r}_{CCD}) + \vec{r}_{CCD}, \tag{2.26}$$

with D the metric distance between the optical centre of the camera and the bottom of the tank alongside the optical axis, and r_{CCD} the position of the centre of the image.

2.C.2 Conversion to metric frame of reference

A cubic spline estimation is computed to smooth the noisy signal of the radius of the detected tank (red line on Figure 2.8 D). The radius signal is repeated over 3 periods to avoid border effects, that is to make estimations at 0 and 2π connected. The Matlab package `immoptibox` (`http://www.imm.dtu.dk/~hbni/immoptibox/`) is used to estimate the cubic splines with function `splinefit`. Splines are piecewise-defined polynomial functions. Knots of the spline are chosen equally spaced. We take 30 knots on each period: the `splinefit` function estimates a local polynomial on each interval defined by two knots. The spline estimated over the second period (2π to 4π) is used for any subsequent calculations to ensure border continuity. The `splineval` function is used to find the radius of the tank wall corresponding to any angular position. Figure 2.8 C shows positions of estimated walls for 2000 θ (*i.e.* of the same magnitude as the number of pixels describing the tank contours) which allow to compute the centre of the tank (x_0, y_0) which will be used as the centre of the coordinate system of fish positions. Given the mean radius derived from these estimates and the *a priori* known radius in millimetres, the number of pixels per millimetres is computed (PixelsToMm ratio) (Figure 2.8 E). This value is the conversion ratio to translate image coordinates into the experimental metric frame of reference, which origin is taken to be the centroid of the estimated tank positions. Figure 2.8 F exemplifies the metric fish positions and velocities. The tank coordinates are also converted to metric coordinates through the same translation and a new metric spline is evaluated to obtain the tank metric coordinates for any angular position.

2.C.3 Fish shape detection and body length/width measurement

By removing the image background information from every frame, idTracker is able to detect an approximate fish shape, *i.e.* a set of pixels' coordinates with their corresponding light intensity, which is stored as intermediary files. These files are read to find the main axis of the fish through a principal component analysis, to estimate the typical body length (BL) and body width (BW) for every fish in all frames. BL and BW are calculated as the difference of the maximum and minimum value constituted by the projection of fish points along respectively main axis and secondary axis, then converted to metric values through the above-mentioned conversion ratio. From the main axis, a heading can be derived using the lower light intensity of fish image due to the black eye of the fish. Detection of the head direction along the main axis is done by evaluating the position relative to the barycentre of the blackest five percent of fish image points. To avoid inaccurate shape detection due to the identification of fish shade as fish shape when the fish is stopped against the wall, we approximate BL and BW as their mean value when the fish is moving faster than $15\,\mathrm{mm.s^{-1}}$.

2.C.4 Fish activity selection and sampling

Inter-fish variability in terms of activity is reduced by selecting the phase where a sustained swimming is observed. Considering that the observation period is much longer than the typical activity phase, these detected activity phases are sampled into sections of two minutes, allowing us to grasp the intra-fish variability occurring along an activity section.This procedure is based on the evaluation of fish speed relative to its mean body-length $u = \frac{v}{\mathrm{BL}}$, evaluated through a centred difference scheme with $0.16\,\mathrm{s}$ amplitude. First, the program detects whether the fish are **swimming**, **pausing** or **stopping**. **Swimming** is defined as swimming at a speed greater than a threshold velocity $u_{\min}$. **Pausing** is defined as the fastest fish of the group swimming at a speed smaller than $u_{\min}$ for a period of time smaller or equal to $\tau_s = 4\,\mathrm{s}$. **Stopping** is defined as the fastest fish of the group swimming at a speed smaller than $u_{\min}$ during more than $\tau_s = 4\,\mathrm{s}$. The program extracts sequences of frames where the fish is either **swimming** or **pausing**, removing **stopping** behaviour. From these sequences where the fish are active, *i.e.* not **stopping**, the program cuts series of the

same length τ_l. For each experiment, the program will give discontinued series where the fish are swimming. The number of series for each experiment can be used to estimate how much an experiment will participate in the statistics. The values $\tau_l = 120\,\mathrm{s}$ and $\tau_s = 4\,\mathrm{s}$ are chosen as a compromise between the amount of data available and the insensitivity of the results of activity selection to the mere parameters.The value of $u_{\min} = 0.5\,\mathrm{BL.s}^{-1}$ is a reasonably low threshold that allows to exclude the low activity phases where the fish uses pectoral fin swimming, of no interest for our study describing regular fish motion using body and caudal fin swimming. The proportion of time where individuals are detected active over the whole experiment is listed in Table 2.2 (column *Proportion of active swimming*).

2.C.5 Segmentation

H. rhodostomus swims in a burst-and-coast (or burst-and-glide) style. There is a succession of short acceleration phases during which the fish may also change its heading and each acceleration phase is followed by a gliding phase during which the velocity decreases and then the cycle starts again (Figure 1C, main text). The points of acceleration exhibited by fish when "bursting" is used to detect these decisions. Most heading changes occur at these decision points also called "kicks". Our assumption is that these kicks are sufficient to describe fish swimming behaviour, the passive phases containing only a consequence of the previous action, being entirely determined by physical forces. Thus we can minimize the amount of noise given by barycentre estimation and minor trajectory deviations by describing the trajectory as segments between kicks. In order to properly identify the acceleration events, we have to smooth the raw speed time series obtained by taking the modulus of the velocity vector through a centred difference scheme over a moving time window of bandwidth 0.08 s (4 frames). We use a Savitsky-Golay[1] filter of degree three over a 0.36 s time window (18 frames) to smooth the raw time series, allowing us to classify the time series into accelerating and decelerating state [38]. To limit remaining noise, we fuse any consecutive pair of accelerations separated by a deceleration lasting less than 0.08 s. We then discard any acceleration lasting less than 0.08 s as it is a too short period of time regarding the typical duration of a body motion. We assume that the times of the kicks coincide with the starting of the acceleration periods.

2.C.6 Segmented variable estimation

Assuming a fish instantaneously takes a new direction and velocity every time its body motion produces an acceleration, we reduce the full time-sampled trajectories of every two minutes activity sample b of a fish Id to a set of positions and times of interest $\{\vec{x}_i, t_i\}_b^{Id}$ corresponding to a set of decision events. From this point of view, the statistics of interest used in both data and simulation, and discussed below, are:

- The length and time intervals between two decision events;
- The absolute and relative to rotation direction change in orientation due to a decision event;
- The distance between the fish centroid and the closest point on the wall (*wall distance*);
- The top speed between decision events.

Length and time intervals

Length between two decision events is defined as the Euclidean distance between both decision points $l_i = \|\vec{x}_{i+1} - \vec{x}_i\|$. The duration $\tau_i = t_{i+1} - t_i$ of the kick initiated at t_i is calculated from decision times.

Heading

Heading of the fish ϕ_i during the kick initiated at t_i is identified to the direction of the vector between two decision points.

Computation of wall distances

θ_i, the angle in radians between the positive x-axis of the frame and (x_i, y_i) is computed from the current position of the fish (x_i, y_i). The radius r_{θ_i} for θ_i is computed based on the spline estimation described in the previous section. The wall distance is the Euclidean distance between fish position and estimated wall position as in $r_{\mathrm{w},i} = \|r_{\theta_i}\vec{e}_{\theta_i} - \vec{x}_i\|$.

[1] A. Savitzky, M. J. E. Golay (1964) Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* 36(8):1627–39.

Top speed between kicks

The top speed v_i between kicks is determined from the smoothed speed time series used by the segmentation procedure, taking the maximum value reached between kicks at time t_i and t_{i+1}.

2.C.7 Symmetrisation of the data

We did not observe any statistically relevant left/right asymmetry in the distribution of angles θ_w (1 and 2 fish; see Figure 1E), or ψ and $\Delta\phi$ (2 fish; see Figure 1F). Assuming perfect left/right symmetry amounts to saying that a trajectory as observed from the top of the tank (as we did) has exactly the same probability to occur as the very same trajectory but as seen from under the tank ("mirror trajectory"). For the mirror trajectory, all angles θ_w, ψ, and $\Delta\phi$ have the opposite sign compared to the original trajectory. Hence, the systematic angle change $\delta\phi$ of a fish due to the interaction with the wall (1 or 2 fish experiments) and with an other fish (2 fish experiments) must exactly satisfy the symmetry condition

$$\delta\phi(r_\mathrm{w}, -\theta_\mathrm{w}) = -\delta\phi(r_\mathrm{w}, \theta_\mathrm{w}), \tag{2.27}$$

for 1 fish experiments, and

$$\delta\phi(r_\mathrm{w}, -\theta_\mathrm{w}, d, -\psi, -\Delta\phi) = -\delta\phi(r_\mathrm{w}, \theta_\mathrm{w}, d, \psi, \Delta\phi), \tag{2.28}$$

for 2 fish experiments. In order to analyse and disentangle the interactions, notably the attraction and alignment interactions between fish, we have imposed general functional forms (see Eqs. (7,12,13) in the main text) obeying these conditions. Accordingly, exploiting this assumed but reasonable left/right symmetry, we have effectively doubled our data set by adding the mirror trajectory associated to each observed trajectory. This procedure not only reduces the statistical uncertainty on quantities depending on angles (by a factor $\sqrt{2}$, by the law of large numbers), but it also helps stabilizing the optimization procedure used to extract the various components of the interactions from $\delta\phi$, which is detailed in the next section.

2.D Analysis of the interactions

2.D.1 Interaction with the wall of a single fish

The position and orientation of a fish relative to the wall is fully determined by r_w and θ_w (see the main Figure 1E). As explained in the article (section 2.2.3), in addition to the random component of the angle change $\delta\phi$ between kicks, which accounts for the fish spontaneous motion, we look for a systematic angle change due to the presence of the wall of the form $\delta\phi_\mathrm{W}(r_\mathrm{w}, \theta_\mathrm{w}) = f_\mathrm{w}(r_\mathrm{w})O_\mathrm{w}(\theta_\mathrm{w})$, where $O_\mathrm{w}(\theta_\mathrm{w})$ is an odd function of θ_w. If the wall interaction tends to push back the fish toward the center of the tank and is isotropic, one has exactly that $O_\mathrm{w}(\theta_\mathrm{w}) \propto \sin(\theta_\mathrm{w})$ (the projection of a radial force on the angular acceleration, which is perpendicular to the velocity). Hence, the ansatz $\delta\phi_\mathrm{W}(r_\mathrm{w}, \theta_\mathrm{w}) = f_\mathrm{w}(r_\mathrm{w})O_\mathrm{w}(\theta_\mathrm{w})$ is the simplest generalization accounting for the fish anisotropic perception of its environment, while keeping a product form and still obeying left/right symmetry. Note that in the Gautrais *et al.* model (Gautrais et al., 2012), the sign function was phenomenologically used instead of the sin function. Despite having a qualitatively similar shape, and being both odd functions of θ_w as requested by symmetry (see above), the sign function has the unphysical/unbiological drawback of attributing a sharp discontinuous response to a fish when it approaches the wall from an arbitrary small angle from the left or the right (an angle sign that the fish could not measure with such a perfect precision).

In order to measure the actual $f_\mathrm{w}(r_\mathrm{w})$ and $O_\mathrm{w}(\theta_\mathrm{w})$, we first define a discrete mesh of the two-dimensional space $(r_\mathrm{w}, \theta_\mathrm{w})$, with each direction $r_\mathrm{w} \in [0, R]$ and $\theta_\mathrm{w} \in [-\pi, \pi]$ partitioned respectively in I and J boxes (typical values are $I = 40$ and $J = 30$). We tabulate the unknown functions $f_\mathrm{w}(r_\mathrm{w})$ and $O_\mathrm{w}(\theta_\mathrm{w})$ by defining f_i as the (mean) value of $f_\mathrm{w}(r_\mathrm{w})$ when r_w falls in box i, and O_j as the (mean) value of $O_\mathrm{w}(\theta_\mathrm{w})$ when θ_w falls in box j. We finally define ϵ_{ij} as the number of data points falling in the squared box of index i and j, and $\delta\phi_{ij}$ as the averaged experimental angle change for data points in this box.

f_i and O_j are then determined by minimizing the error

$$\Delta = \sum_{i=1}^{I}\sum_{j=1}^{J} \epsilon_{ij}(\delta\phi_{ij} - f_i O_j)^2. \tag{2.29}$$

This minimization is achieved by writing the equation $\partial\Delta/\partial f_i = 0$ and $\partial\Delta/\partial O_j = 0$ under the form

$$f_i = \frac{\sum_{j=1}^{J} \epsilon_{ij}\delta\phi_{ij}O_j}{\sum_{j=1}^{J} \epsilon_{ij}O_j^2}, \tag{2.30}$$

$$O_j = \frac{\sum_{i=1}^{I} \epsilon_{ij}\delta\phi_{ij}f_i}{\sum_{i=1}^{I} \epsilon_{ij}f_i^2}. \tag{2.31}$$

It is straightforward to realize that if the experimental $\delta\phi_{ij}$ were exactly of the form $f_i \times O_j$, the right-hand side of Eqs. (2.30,2.31) would indeed exactly recover f_i and O_j.

This system is solved iteratively starting from reasonable initial conditions, but most importantly with O_j being an odd function of $\theta_{\rm w}$. We checked that this procedure leading to the results of Figure 4 does not depend on the initial conditions.

In practice, knowing the f_i's and O_j's at a given iteration, we generate their values at the next iteration by computing

$$f'_i = (1-p)f_i + p\hat{f}_i, \tag{2.32}$$

$$O'_j = (1-p)O_j + p\hat{O}_j, \tag{2.33}$$

where $\hat{f}_i$ and $\hat{O}_j$ are given by the right-hand side of Eqs. (2.30,2.31), and p is a damping parameter that we took equal to 0.25. Obviously, the fixed point solution of Eqs. (2.32,2.33) ultimately coincides with the wanted solution of Eqs. (2.30,2.31).

Since multiplying O_j by an arbitrary constant and dividing f_i by the same constant leaves the product f_iO_j unchanged, we choose to normalize O_j (and all angular functions appearing in Figure 4 and 6) after each iteration such that its average square is unity:

$$\frac{1}{J}\sum_{j=1}^{J} O_j^2 = \frac{1}{2\pi}\int_{-\pi}^{+\pi} O_{\rm w}^2(\theta_{\rm w})\, d\theta_{\rm w} = 1 \tag{2.34}$$

The procedure described here converges to a relative accuracy of 10^{-6} in typically 100 iterations, leading to the result of Figure 4.

2.D.2 Attraction and alignment interaction between two fish

We define a procedure identical in spirit as above, but involving more unknown interaction functions now depending on the 3 parameters d, ψ, and $\Delta\phi$ defined in Figure 1F. We choose to restrict our analysis to data points for which the focal fish was at a distance greater than 2 BL$\sim$ 60 mm, for which the wall interaction is found to be negligible by the previous analysis.

Again, we partition the three-dimensional space $(d, \psi, \Delta\phi)$ in a mesh of $K\times L\times M$ boxes (with typically $K = 40$, $L = M = 30$). As previously, we define ϵ_{klm} as the number of data points falling in the cubic box of index k, l and m, and $\delta\phi_{klm}$ as the averaged experimental angle change of the focal fish for data points in this box.

As explained in the main text (section 2.2.5), the systematic angle change $\delta\phi = \delta\phi_{\rm Att}(d, \psi, \Delta\phi) + \delta\phi_{\rm Ali}(d, \psi, \Delta\phi)$ due to the attraction and alignment forces is parameterized by 6 unknown functions

$$\delta\phi_{\rm Att}(d, \psi, \Delta\phi) = F_{\rm Att}(d)O_{\rm Att}(\psi)E_{\rm Att}(\Delta\phi), \tag{2.35}$$

$$\delta\phi_{\rm Ali}(d, \psi, \Delta\phi) = F_{\rm Ali}(d)O_{\rm Ali}(\Delta\phi)E_{\rm Ali}(\psi), \tag{2.36}$$

with parity constraints (O functions are odd, E functions are even). These rather general functional forms translate into mathematical forms the notion of attraction and alignment which can in fact have cumulative or contrary effects depending on the relative position or orientation of the two fish (see Figure 6A). Intuitively, attraction means that if the other fish is on the right, the focal fish should turn to the right, and should turn by the same amount to the left in the "mirror situation" (see section 2.C.7).

Again we tabulate these 6 functions, each depending on only one variable, as $F_{{\rm Att},k}$, $O_{{\rm Att},l}$, $E_{{\rm Att},m}$, $F_{{\rm Ali},k}$, $O_{{\rm Ali},m}$, $E_{{\rm Ali},l}$. These functions are determined by minimizing the error

$$\Delta = \sum_{k=1}^{K}\sum_{l=1}^{L}\sum_{m=1}^{M} \epsilon_{klm}\times(\delta\phi_{klm} - F_{{\rm Att},k}O_{{\rm Att},l}E_{{\rm Att},m} - F_{{\rm Ali},k}O_{{\rm Ali},m}E_{{\rm Ali},l})^2. \tag{2.37}$$

Expressing that the derivative of Δ with respect to the 6 tabulated function is zero at the minimum, we obtain 6 equations similar to Eqs. (2.30,2.31), although a bit more complicated. For instance, $F_{\mathrm{Att},k}$ satisfies the fixed point equation

$$F_{\mathrm{Att},k} = \frac{\sum_{l,m} \epsilon_{klm} O_{\mathrm{Att},l} E_{\mathrm{Att},m} (\delta\phi_{klm} - F_{\mathrm{Ali},k} O_{\mathrm{Ali},m} E_{\mathrm{Ali},l})}{\sum_{l,m} \epsilon_{klm} O^2_{\mathrm{Att},l} E^2_{\mathrm{Att},m}}. \tag{2.38}$$

Note the presence of the counter term $F_{\mathrm{Ali},k} O_{\mathrm{Ali},m} E_{\mathrm{Ali},l}$ between the parentheses of Equation (2.38), meaning that the attraction force is evaluated by subtracting the estimated alignment interaction to the actual experimental angle change. Again, it is straightforward to check that if the experimental $\delta\phi_{klm}$ takes exactly the form $F_{\mathrm{Att,k}} O_{\mathrm{Att},l} E_{\mathrm{Att},m} + F_{\mathrm{Ali},k} O_{\mathrm{Ali},m} E_{\mathrm{Ali},l}$, Equation (2.38) exactly recovers the correct d dependence of the attractive force $F_{\mathrm{Att},k}$.

Finally, the resulting system of 6 equations is solved by the same iterative procedure as for the wall interaction (including the normalization of the average square of angular functions), leading to the results of Figure 6B, C, and D. In particular, the angular functions obtained are plotted with simple analytical forms given by

$$O_{\mathrm{Att}}(\psi) \propto \sin(\psi)[1 - 0.33\cos(\psi)], \tag{2.39}$$

$$E_{\mathrm{Att}}(\Delta\phi) \propto 1 + 0.48\cos(\Delta\phi) - 0.31\cos(2\Delta\phi), \tag{2.40}$$

$$O_{\mathrm{Ali}}(\Delta\phi) \propto \sin(\Delta\phi)[1 + 0.30\cos(2\Delta\phi)], \tag{2.41}$$

$$E_{\mathrm{Ali}}(\psi) \propto 1 + 0.60\cos(\psi) - 0.32\cos(2\psi), \tag{2.42}$$

the overall multiplicative constant being fixed by the normalization of the average square of these functions to unity. The analytical forms for $F_{\mathrm{Att}}(d)$ and $F_{\mathrm{Ali}}(d)$ are discussed in the main text, section 2.2.5.

One of the main interests of assuming these reasonable (see the discussion above and at the end of section 2.2.5 of the main text) product forms is to drastically limit the number of fitting parameters, but also to derive from these forms an *explicit* model. Indeed, if we were to produce a three-dimensional map of $\delta\phi$ on the $K\times L\times M$ mesh made of 36000 boxes ($K = 40$, $L = M = 30$), we would need 36000 fitting parameters at this resolution, barely smaller than the number of kicks available from experiment (~ 200000 for 2 fish). The present procedure only requires $2\times(K + L + M) = 200$ fitting parameters, which allows us to extract the main feature of the interactions with a high resolution and yet a rather small noise (see Figure 4 and 6, and in particular, the various angular functions).

2.E Parameter estimation and simulations

As explained in the main text (section 2.2.3) and illustrated in the insert of Figure 2D, the spontaneous angle change has a nearly Gaussian distribution of zero mean (left/righ symmetry) and variance $\gamma_{\mathrm{R}} \approx 0.35\,\mathrm{rad}$, as measured when the fish is far from the wall ($r_{\mathrm{w}} > 2\,\mathrm{BL}$). This value of γ_{R} was used in all simulations of the one fish dynamics, in the three circular tanks of radius $R = 176,\ 250,\ 353\,\mathrm{mm}$. For two fish, γ_{R} is found experimentally to be slightly larger, probably because each fish activity is stimulated by the presence of the other fish.

The total angle change PDF shown in the main graph of Figure 2D (one fish) and in Figure 5D (two fish) is dominated by the majority of kicks where the fish are very close to the wall, and has a width reduced by a factor nearly 3, so that a fair estimate of the spontaneous angle change intensity near the wall is (see Equation (6))

$$\gamma^0_{\mathrm{R}} = (1 - \alpha)\gamma_{\mathrm{R}} \sim \gamma_{\mathrm{R}}/3. \tag{2.43}$$

The exact value of γ^0_{R} (or α), of the comfort distance to the wall l_c introduced below Equation (7), which should naturally be of order of $1\,\mathrm{BL} \sim 30\,\mathrm{mm}$, and of the intensity of the interaction with the wall γ_{W} are tuned near their experimental expected value, resulting in the very satisfactory agreement with experiments obtained in Figure 2 and Figure 5. The parameter used in our simulations are summarized in Table 2.3.

For each of the 4 conditions (the three tank sizes for one fish and one tank size with two fish; see Table 2.2), the graphs of Figure 2 (one fish) and in Figure 5 (two fish) are obtained by typically simulating 100 runs with 10^6 kicks each, compared to the 100000-200000 kicks recorded experimentally for each condition/tank.

Table 2.1: Experimental conditions

	1 Fish			**2 Fish**
Arena radius (mm)	176	250	353	250
Number of experiments	11	12	8	16
Temperature (°C, mean ± se)	27.7 ± 0.2	26.8 ± 0.3	27.4 ± 0.3	26.1 ± 0.3
Body length (mm, mean ± se)	35.0 ± 1.1	28.3 ± 1.3	30.9 ± 1.5	30.3 ± 1.0

Table 2.2: List of experiments

Group size (Number of fish)	Tank Size (mm)	Date	Proportion of active swimming	Duration (min)	Number of kicks
1	176	2014-08-20	0.8	191	21012
		2014-08-20	0.1	193	2572
		2014-08-21	0.4	187	10122
		2014-08-21	0.7	197	19849
		2014-08-22	0.7	204	15691
		2014-08-22	0.1	202	2401
		2014-08-26	0.9	230	31603
		2014-08-26	0.1	208	1550
		2015-07-18	0.0	188	631
		2015-07-21	0.2	178	4459
		2015-07-25	0.5	185	11772
	250	2014-04-25	0.1	186	3199
		2014-04-25	0.3	191	6112
		2014-04-29	0.3	195	6363
		2014-04-29	0.2	222	5955
		2014-04-30	0.5	187	13632
		2014-05-06	0.3	181	6760
		2014-05-06	0.1	191	2206
		2014-05-07	0.5	187	12127
		2015-06-10	0.6	190	10584
		2015-07-19	0.5	185	11531
		2015-07-28	0.2	204	3833
		2015-07-31	0.3	198	7179
	353	2014-05-12	0.5	205	13180
		2014-05-13	0.7	207	16325
		2014-05-13	0.4	192	11362
		2014-05-28	0.2	197	4468
		2015-07-23	0.2	187	6501
		2015-07-26	0.3	189	6357
		2015-07-29	0.8	202	21944
		2015-07-30	0.9	209	23604
2	250	2013-01-31	0.9	65	13605
		2013-02-04	0.9	65	14473
		2013-02-13	1.0	41	9037
		2013-03-04	0.9	68	13663
		2013-03-14	0.0	86	1151
		2013-03-18	0.9	61	12401
		2013-03-27	0.9	63	14331
		2014-02-14	1.0	62	13365
		2014-03-11	1.0	61	15012
		2014-11-04	0.9	80	15287
		2014-11-05	0.3	77	6136
		2014-11-05	0.7	79	11910
		2014-11-05	0.5	75	8458
		2014-11-06	0.5	75	9262
		2014-11-06	0.8	78	14732
		2014-11-07	0.7	81	12904

Table 2.3: Parameters used in the simulations

# fish; R (mm)	l_c (mm)	γ_{R} (rad)	γ_{R}^0 (rad)	γ_{W} (rad)
1; 176	21	0.35	0.12	0.40
1; 250	30	0.35	0.14	0.12
1; 353	45	0.35	0.11	0.10
2; 250	30	0.45	0.15	0.15

Figure 2.7: A group of two fish from the study species *Hemigrammus rhodostomus*. Credits to David Villa ScienceImage/CBI/CNRS, Toulouse, 2015.

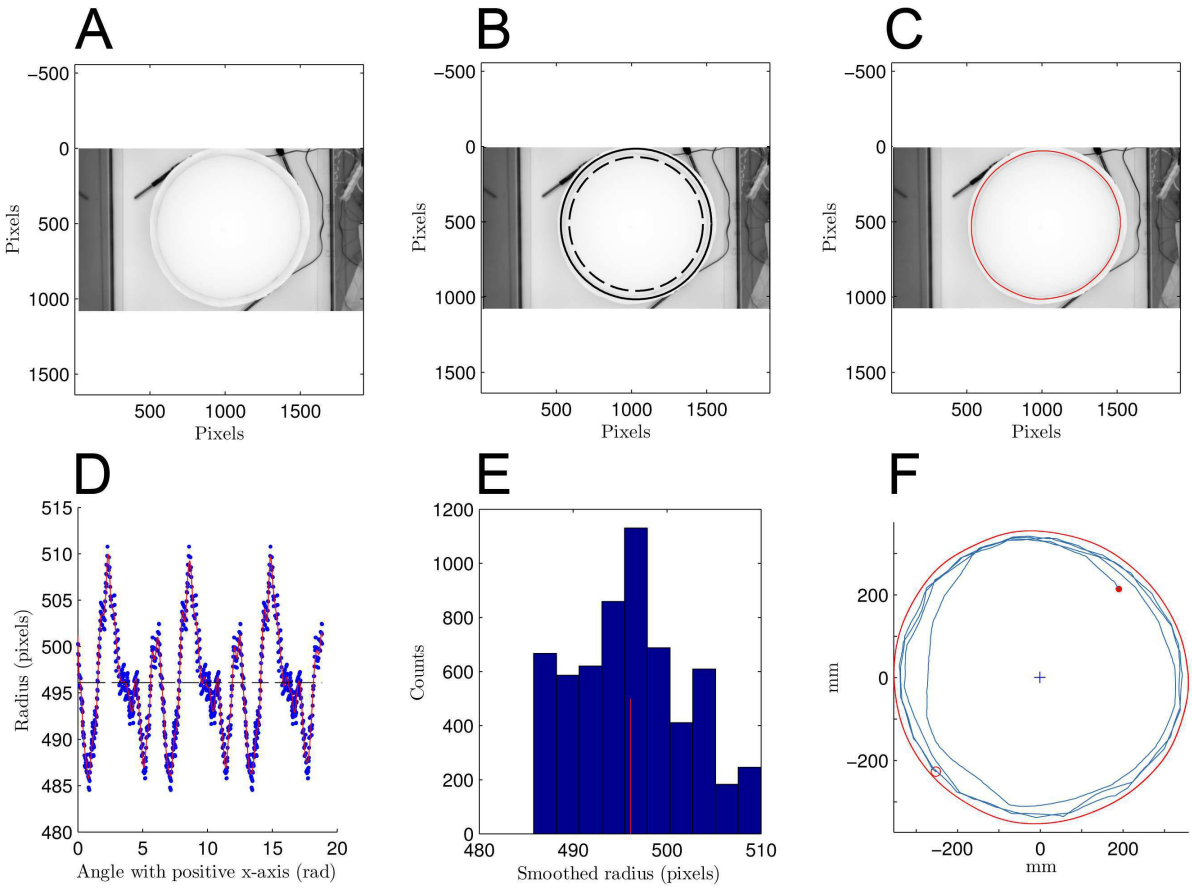


Figure 2.8: **A** Background image in greyscale extracted from a video file of an experiment with the biggest tank (radius $R = 353\,\text{mm}$). **B** Arena estimated from user-defined mask. The outer bold circle of radius R is derived from the mask drawn by the user of the tracking software and defining the area where tracking occurs. Inner dashed circle has arbitrary radius $0.85 \times R$. **C** Estimated walls of the tank. **D** Estimation of the radius along the circle. Red line stands for cubic spline smoothing and extrapolation over 2000 angular points. The signal is repeated 3 times to improve estimation on limits (at 0 and 2π). The second period is kept to compute wall distances. Estimation of local polynomials is done on 30 equally spaced ranges over one period. Dashed line shows the average radius. **E** Distribution of estimated radius in pixels. Red line stands for estimation of the average, used as radius approximation to compute the ratio of pixel to millimetres (PixelsToMm ratio is equal to 0.71 for this video). **F** Trajectory of a fish during 40 seconds (2000 points) reported inside the estimated walls. Filled and empty circle respectively stand for start and end points.

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

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.

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 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).

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 150 L 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 synchronised and polarised manner. Inside an experimental tank,

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

Group Size	Number of trials	Total duration	Body length (mm, mean ± se)
1	4	260 min	33.1 ± 1.8
2	10	652 min	33.3 ± 0.8
5	10	543 min	31.5 ± 0.3

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).

3.2.2 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 segmentation detects whether the fish is in its burst phase or not.

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 φ_i, the angle turned $\delta\phi$, its angle to the outer wall θ_w, its distance to the outer wall r_w, its distance to the neighbour d, the angular position of the neighbour ψ 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 $v_{\max}$ 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 the 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.

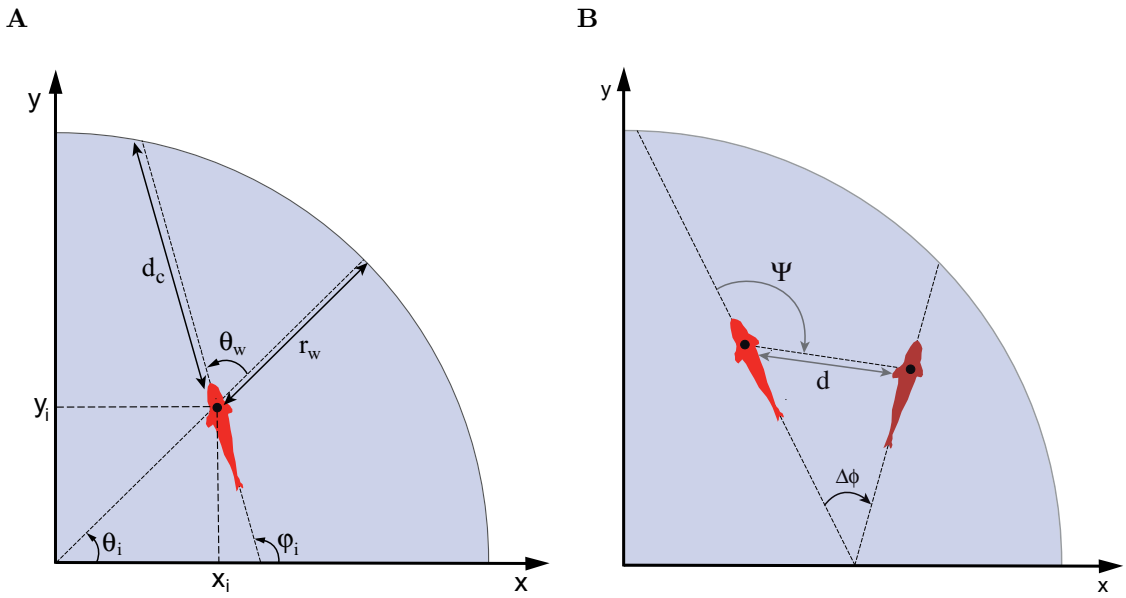


Figure 3.2: Symbols for analysis of fish motion in experiments with 1 fish (A) and 2 fish (B). Fish i has a position (x_i, y_i) at the onset of the burst. We measure its current heading φ_i, its angle to the outer wall θ_w, its distance to the outer wall r_w, its distance to neighbour d, the angular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

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:

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

$$\delta\phi(r_w, -\theta_w) = -\delta\phi(r_w, \theta_w), \tag{3.1}$$

for experiments with 1 fish and

$$\delta\phi(r_w, -\theta_w, d, -\psi, -\Delta\phi) = -\delta\phi(r_w, \theta_w, d, \psi, \Delta\phi), \tag{3.2}$$

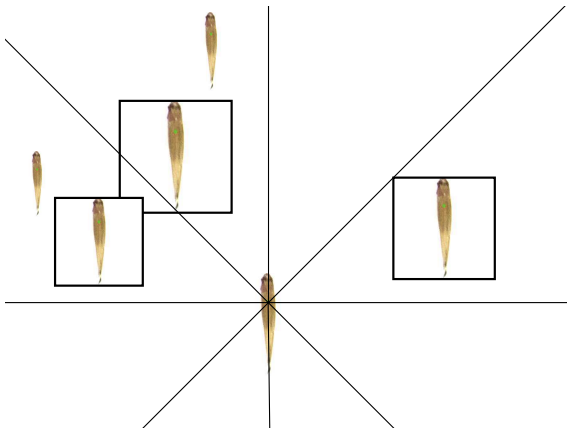


Figure 3.3: Discretisation of the local environment of a focal fish into 8 sectors. The individuals in the squares are labelled as influential because they are the nearest neighbours of the focal fish in each sector.

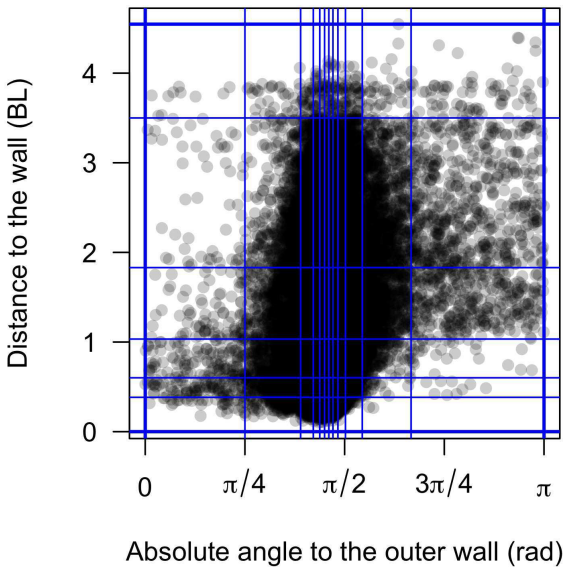


Figure 3.4: Discretisation of the space $\mathcal{S}_w(|\theta_w|, r_w)$ used to sample reactions of a single fish in response to the influence of the wall. Blue lines stand for the bin boundaries. Data points are grey and transparent – the darker a bin, the more points it contains.

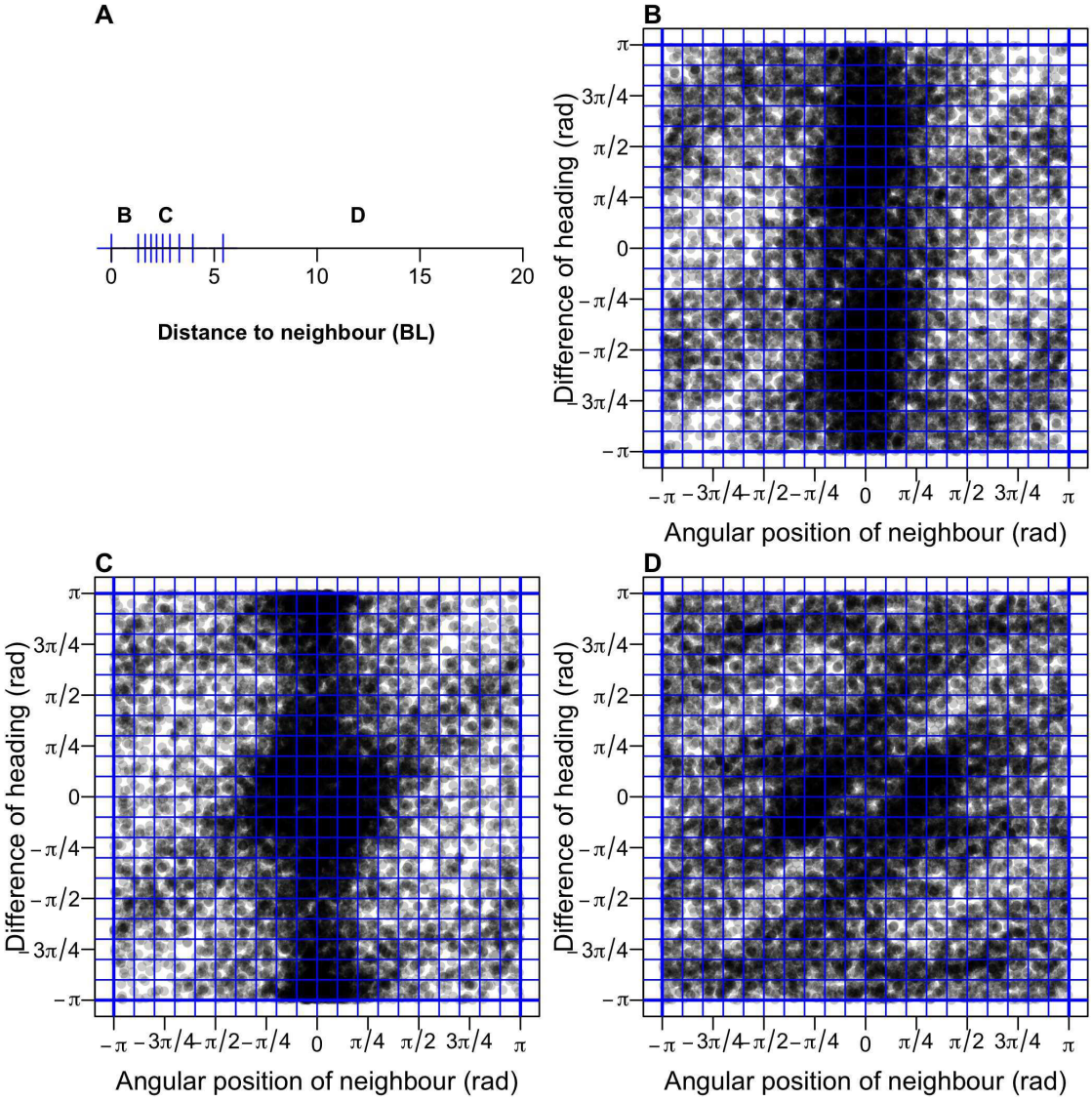


Figure 3.5: Discretisation of the space $\mathcal{S}_n(d, \psi, \Delta\phi)$ used to sample reactions of several fish in response to the neighbours, from experiments with 2 fish in the circular tank and away from the walls. Blue lines stand for the bin boundaries. A). Discretisation regarding distance 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.

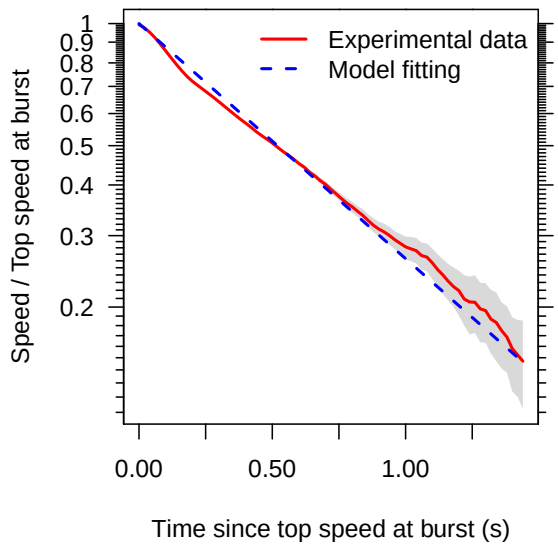


Figure 3.6: Decay of speed during coasting phases relative to top speed in experimental data (red curve) and as modelled by a drag due to water frictions with exponential decay of parameter $D \approx 0.71$ (blue). The grey shade shows the bootstrapped 95% confidence interval.

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 $\mathcal{S}_w(\theta_w, r_w)$ 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 $r_w > 2.5\,\mathrm{BL}$). This is done to remove the influence of the wall from the social interactions. We discretise the space $\mathcal{S}_n(d, \psi, \Delta\phi)$ 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 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:

$$\begin{cases} x_{k+1} = x_k + l_k \cos(\phi_k) \\ y_{k+1} = y_k + l_k \sin(\phi_k), \end{cases} \tag{3.3}$$

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 $l_b = \frac{1}{2}ab^2 + v_k b$, 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 $\mathcal{S}_w(\theta_w, r_w)$ and $\mathcal{S}_n(d, \psi, \Delta\phi)$ is set to avoid empty bins.

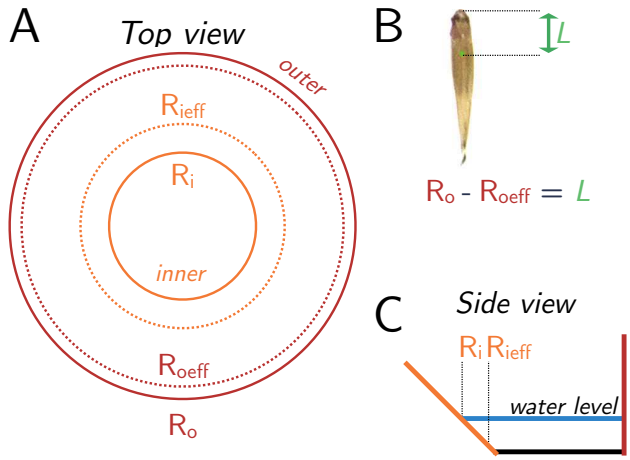


Figure 3.7: Parameters of the rejection algorithm (see section 3.2.4). The real radii of the outer wall (R_o) and of the inner wall (R_i) are not used to delimit the walls. Because of the length L of the front end of the fish to its barycentre used in simulations and of the conic shape of the inner wall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

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: $R_i = 6.27\,\mathrm{BL}$ and outer ring: $R_o = 10.40\,\mathrm{BL}$).

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 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).

Table 3.2: Default parameters used in simulations.

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

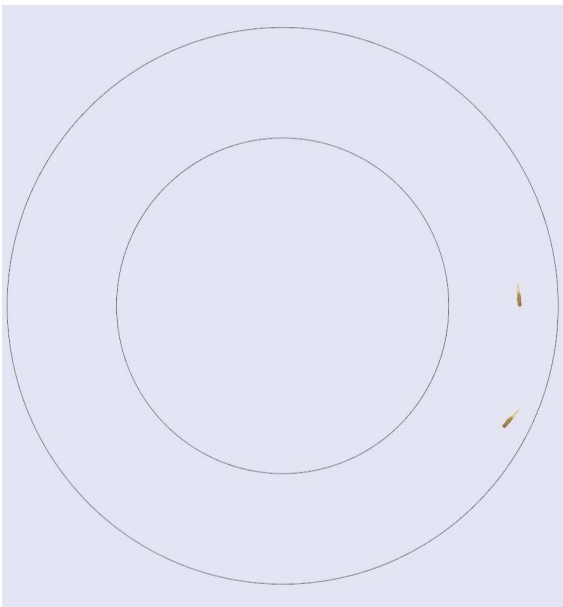


Figure 3.8: Frame of the real-time visualisation of a simulation with two fish. For parameters, see Table 3.2.

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.

	`nocor`	`reg`	`wei`	`inv`	`wei-inv`
Correlated data points	no	yes	yes	yes	yes
Weighted wall interaction	no	no	yes	no	yes
Invisible walls	no	no	no	yes	yes

- 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 at the opposite side of the tank. Otherwise (*yes*) the radius of the inner wall R_i was set to 0 and $R_{\text{ieff}} = 7.14$ (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 ($\approx 0.2\,\text{s}$) 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 (Figures 3.9A and B). The acceleration of fish during the bursting period has a large variance: all accelerations between $0.5\,\text{BL/s}^2$ and $12\,\text{BL/s}^2$ 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 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.

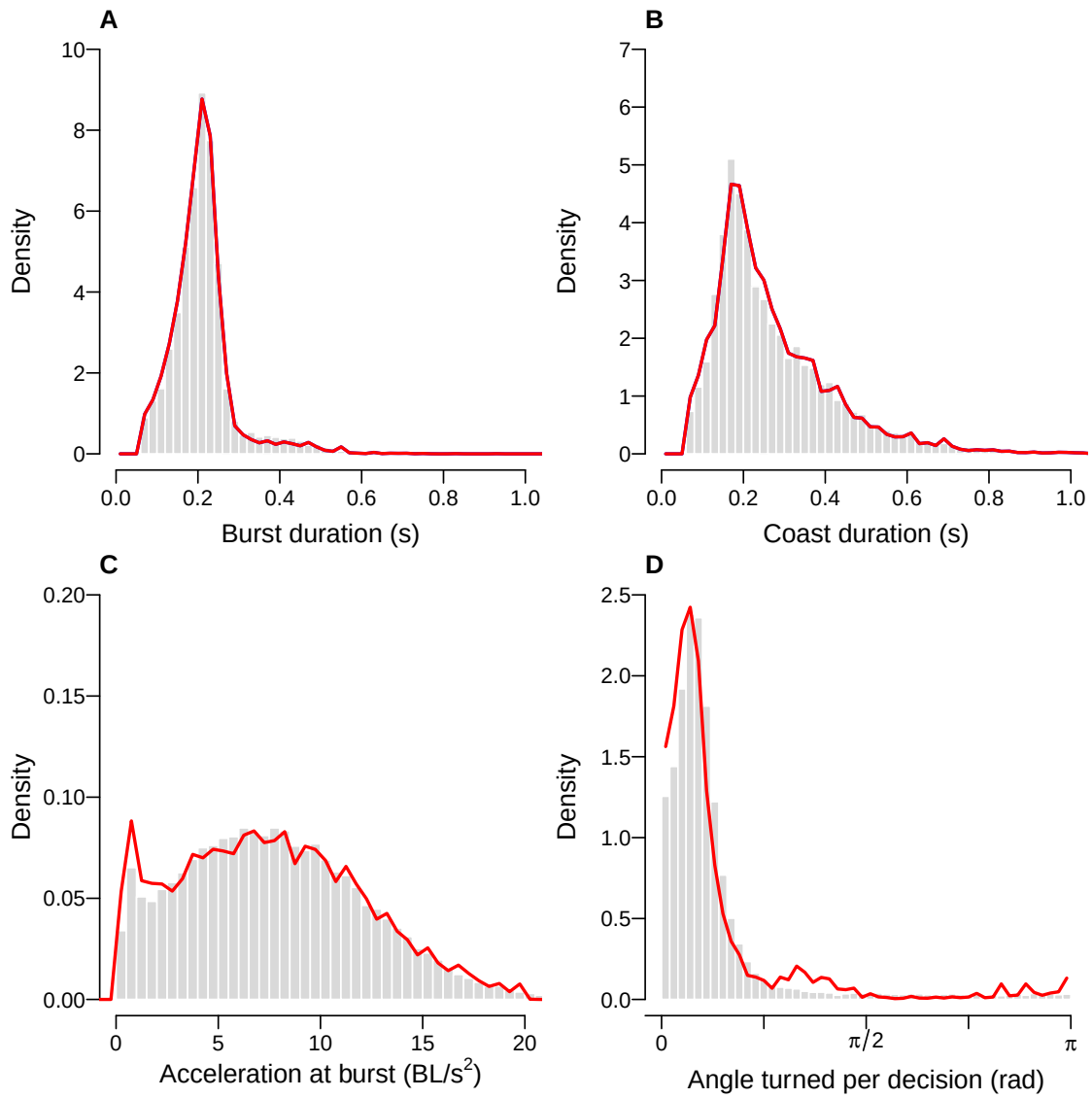


Figure 3.9: Distributions of the burst duration (A), coast duration (B), acceleration at burst (C) and absolute value of the turning angles (D) of 1 fish/agent in empirical data (bars) and simulations (red curve).

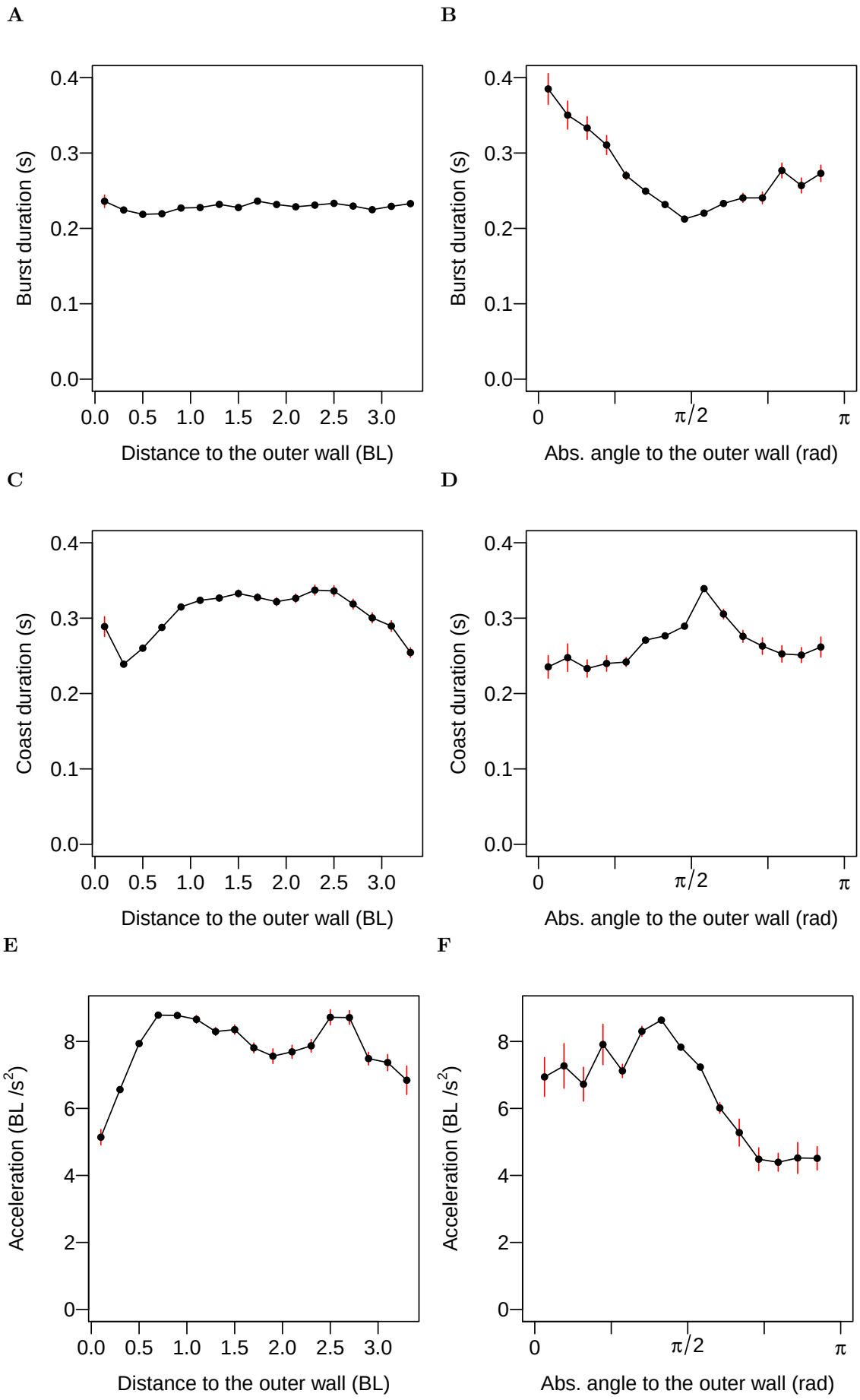

A
B
C
D
E
F
Burst duration (s)
Coast duration (s)
Acceleration (BL /s^2)
0.0
0.1
0.2
0.3
0.4
0
2
4
6
8
0.0 0.5 1.0 1.5 2.0 2.5 3.0
Distance to the outer wall (BL)
0 π/2 π
Abs. angle to the outer wall (rad)

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 48). 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 θ_w (0: facing the outer wall, π: facing the inner wall). Red vertical lines stand for $\pm$ standard error.

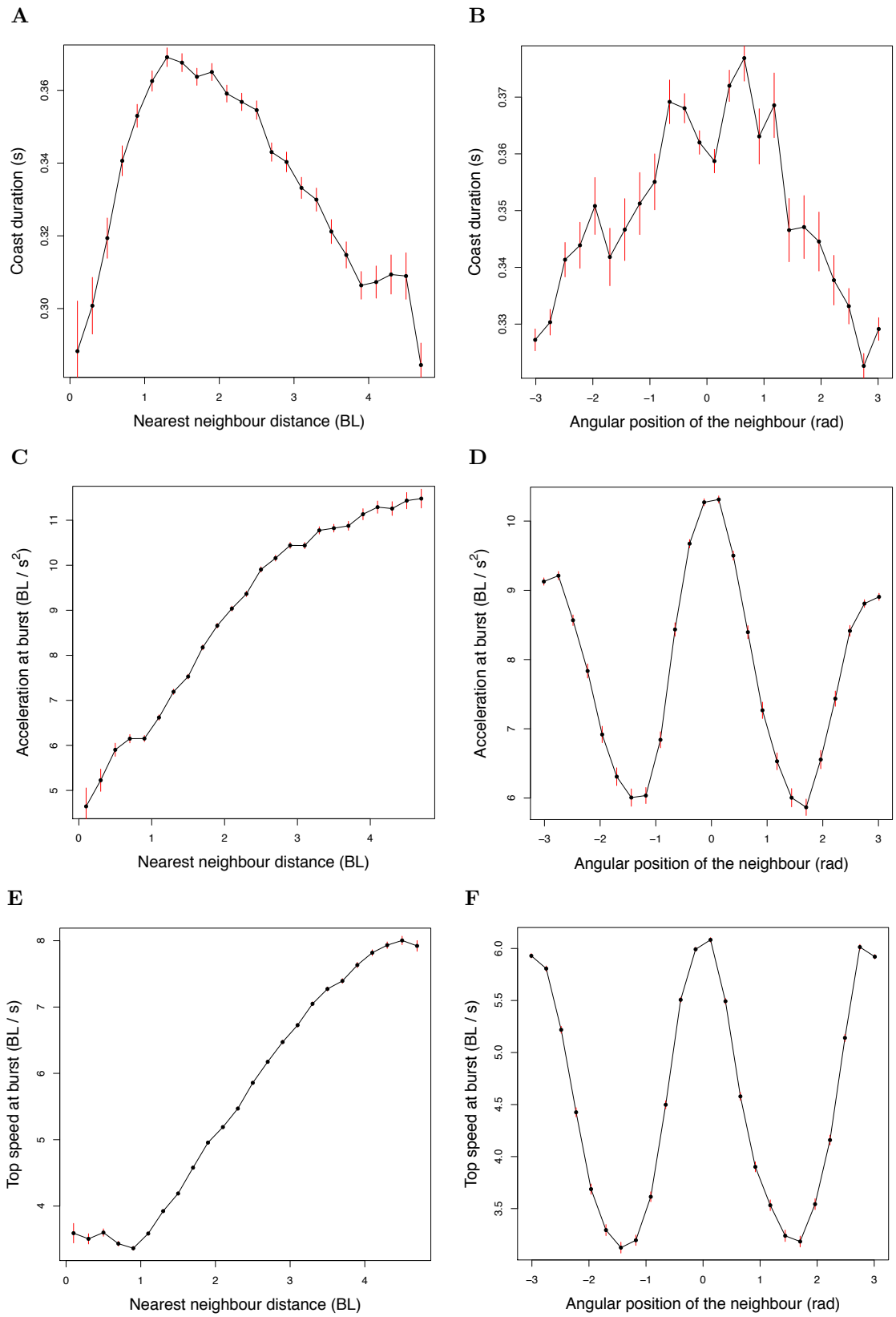


Figure 3.11: Effect of the angular position of the neighbour on the average coasting duration (A and B), the average acceleration of fish (C and D) and the average top speed at burst (E and F) in groups of 2 fish (page 49). Left figures show effect of the distance from the neighbour, right figures the effect of the angular position of the neighbour ψ (0: in front of the focal fish, $\{-\pi;\pi\}$: behind the focal fish). Red segments stand for $\pm$ standard error.

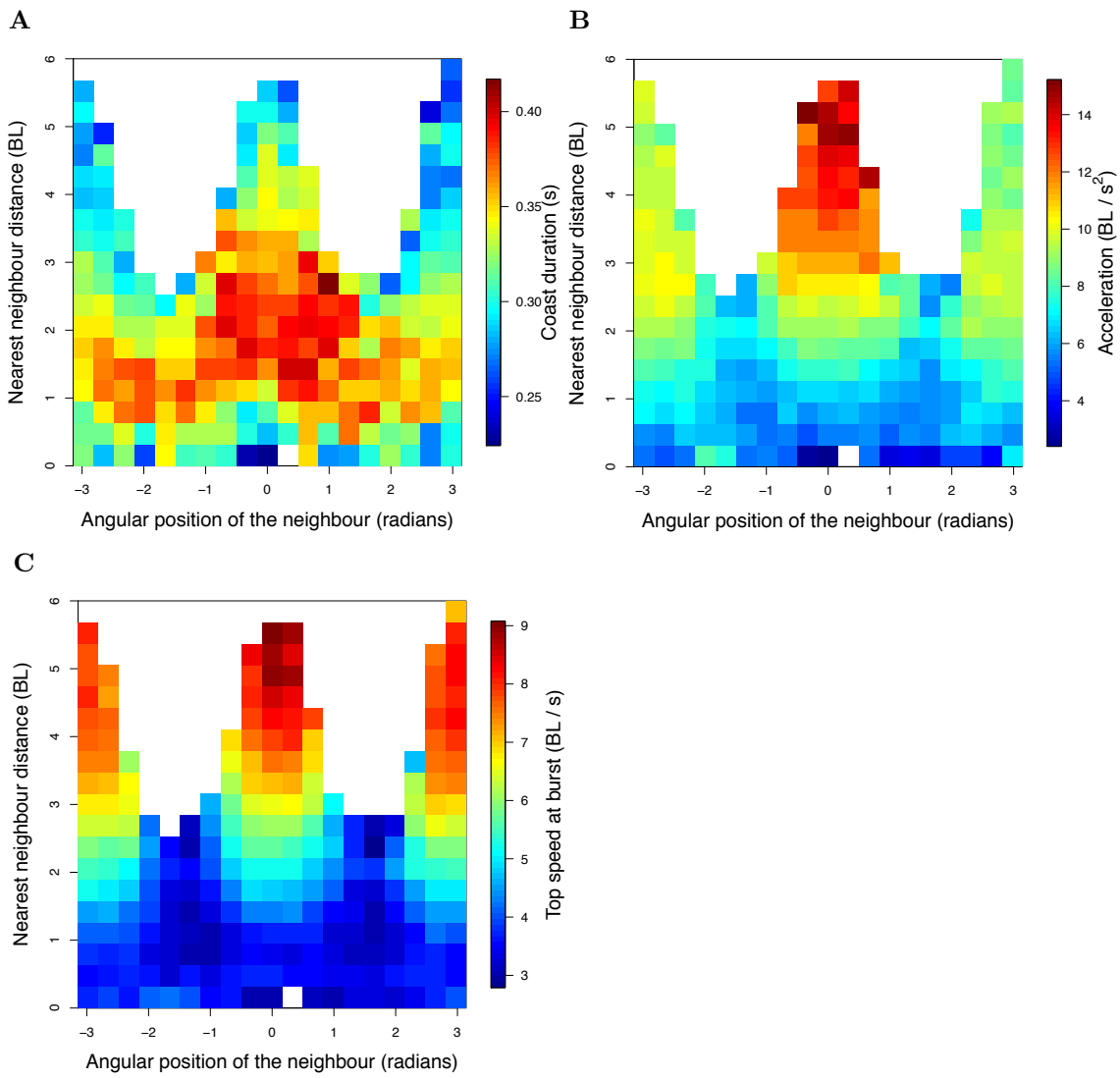


Figure 3.12: Effect of the relative position of the neighbour on the coasting duration (A), the acceleration of fish (B) and top speed at burst (C) in groups of 2 fish. White areas stand for absence of data or cells with less than 25 data points.

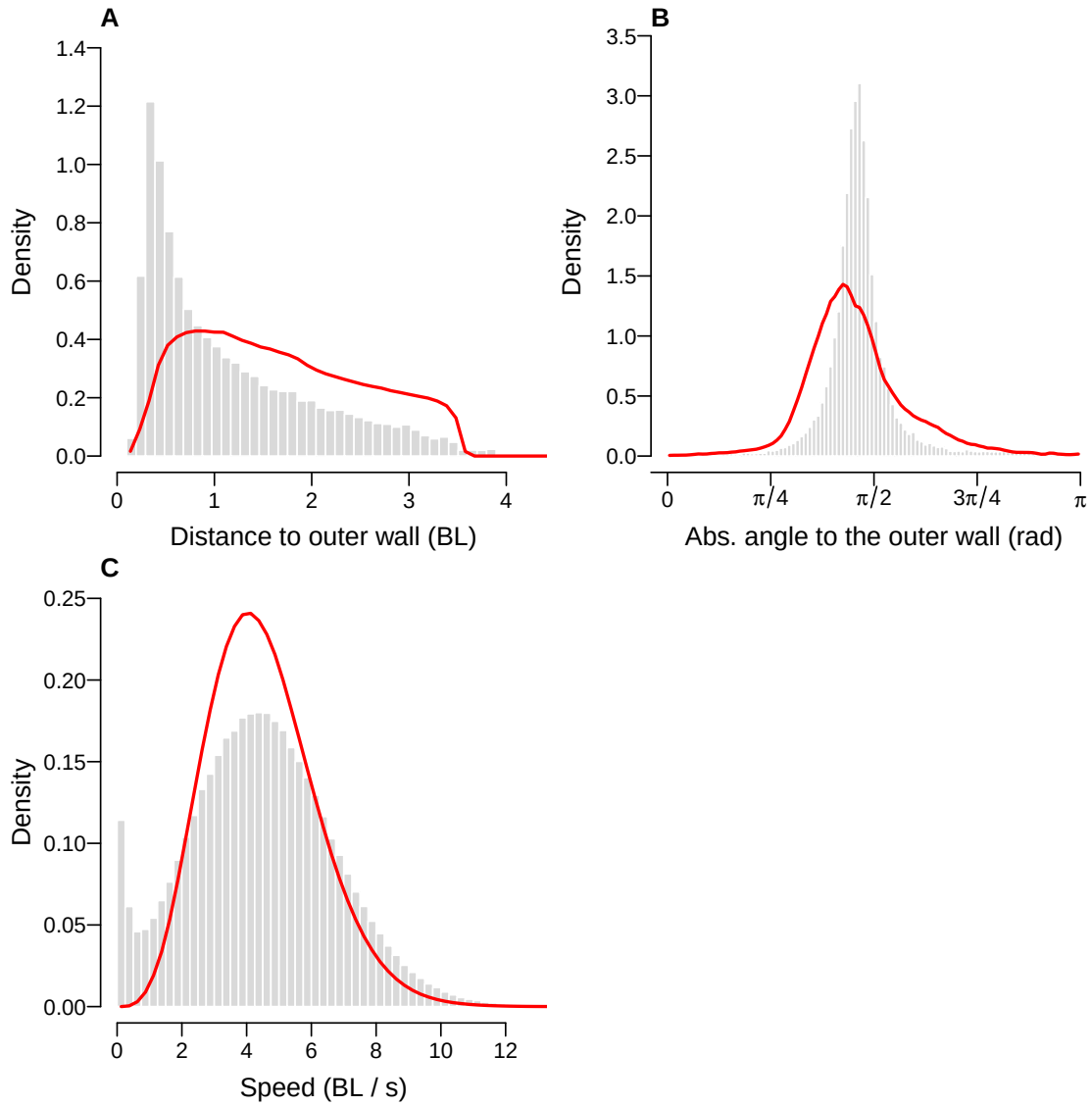


Figure 3.13: Distributions of the (A) distance to the outer wall, (B) absolute value of the angle to the outer wall and (C) speed of 1 fish/agent in experiments (bars) and simulations (red curve).

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

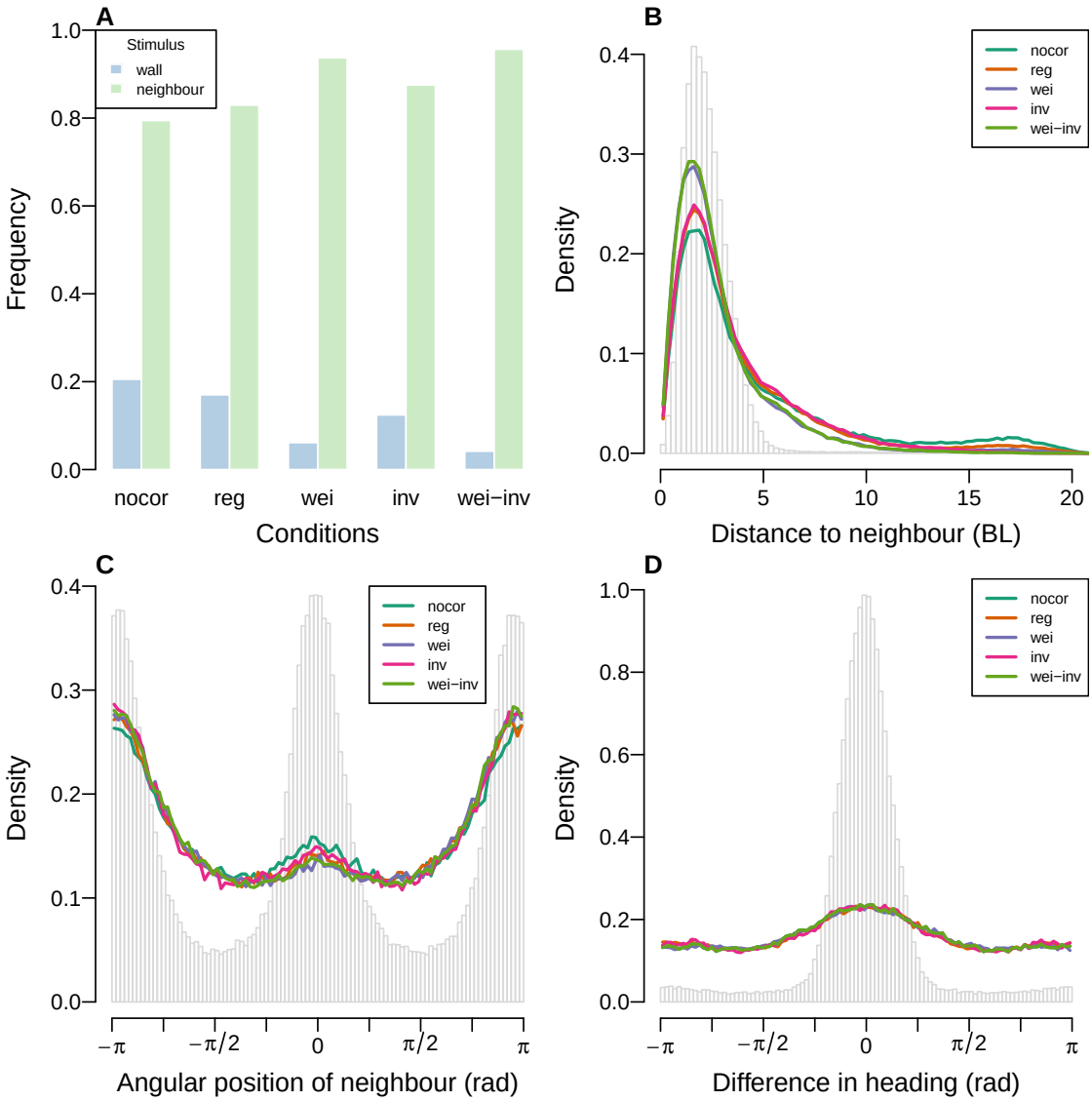


Figure 3.14: A). Relative frequency of the selected stimulus for each condition of simulation (see Table 3.3). Distributions of the distance to the neighbour (B), angular position of the neighbour ψ (C) and difference in heading between fish $\Delta\phi$ (D) measured at the onset of bursts in experiments (bars) and simulations (curves) with 2 fish.

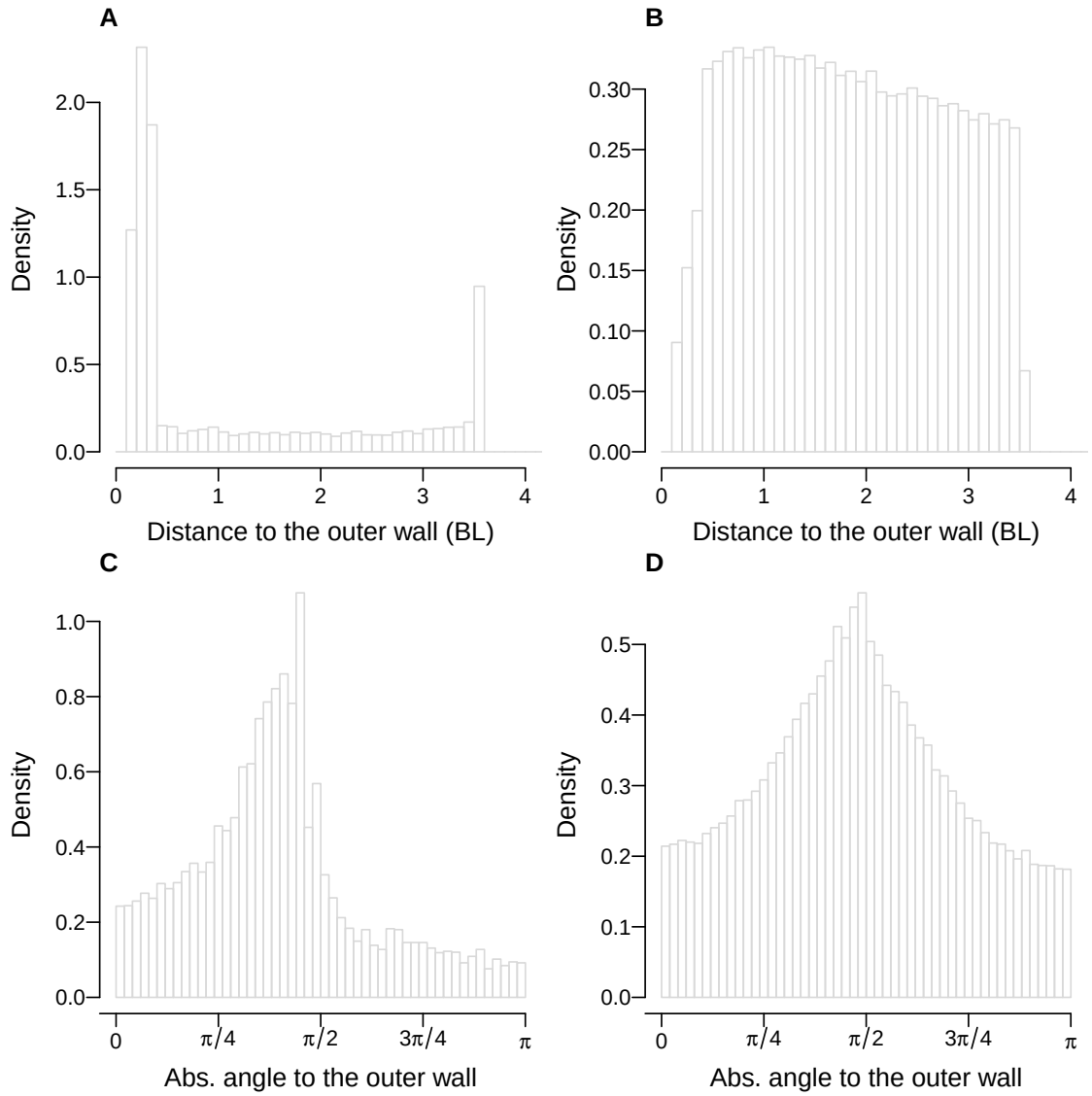


Figure 3.15: Model distributions of the distance to the outer wall and of the absolute value of the angle to the outer wall when the wall is selected as the strongest stimulus (respectively (A) and (C)) and when the neighbour is selected as the strongest stimulus (resp. (B) and (D)), for simulations with 2 fish in the `weighted wall interaction` (`wei`) condition.

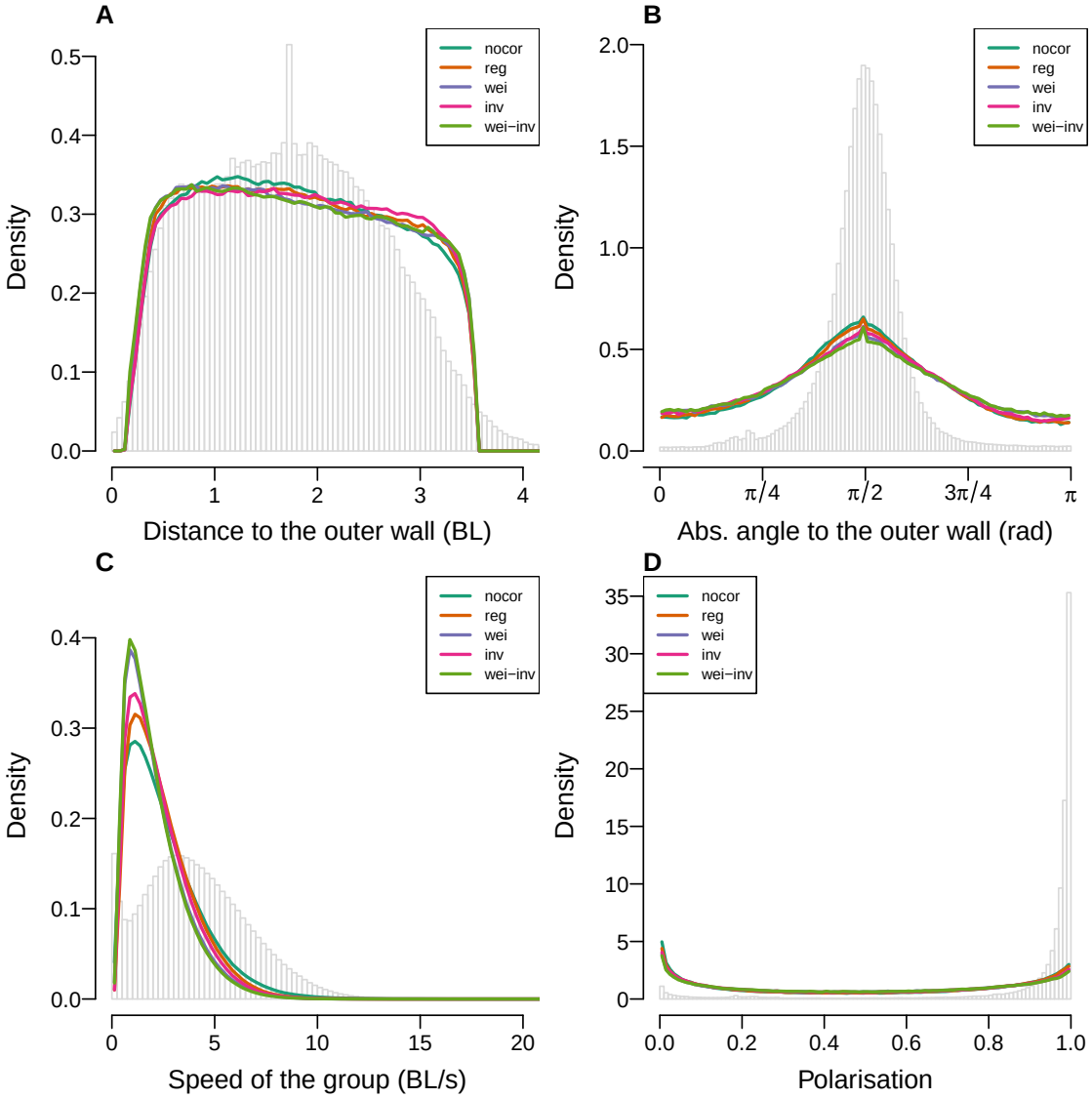


Figure 3.16: Distributions of the (A) distance to the outer wall, (B) absolute value of the angle to the outer wall, (C) speed of the group and (D) polarisation of the group measured on all frames in experiments (bars) and simulations (curves) with 2 fish.

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 is a slight tendency of agents to swim in the same direction as their neighbour (Figure 3.14D). This leads to two modes for the distribution of the polarisation (either well polarised when agents swim in the same direction or not polarised when they swim in opposite directions) (Figure 3.16D).

As for the location of fish or agents in the ring-shaped tank, both swim closer to the outer wall than to the inner one (Figure 3.16A) and parallel to the walls (Figure 3.16B). Agents are slower than real fish (Figure 3.16C).

The resemblance to empirical data is worse for groups of five fish or agents compared to previous group sizes (Figures 3.17 and 3.19). Namely, in simulations, the nearest neighbour is too often far away from the focal agent, when very close in experiments (Figure 3.17B), and there is not preference for frontal neighbours, in contrast to experiments (Figure 3.17C). There is no preference to be aligned with the nearest neighbour in simulations, in contrast to experiments (Figure 3.17D), resulting in a poor agreement of the distribution of the group polarisation – simulated groups are not as well polarised as the groups in experiments (Figure 3.19D). We find that the distribution of the turning angles by the agents does not fit the experimental distribution, especially for groups of 5 (Figure 3.18). Namely, with our assumption that fish would only react to the strongest stimulus, we find that agents make larger turns than fish in experiments, with a mode for values close to π (i.e. U-turns).

As for the location to the wall, agents are too close to the outer wall (Figure 3.19A) and not aligned enough with the walls (Figure 3.19B). They are slower than empirical groups (Figure 3.19C), with a distribution that is very close to the distributions with 2 agents or fish (Figure 3.17C) when groups of 5 fish in experiments are faster than groups of 2 fish. In short, the model fails to reproduce the features for real groups of 5 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 velocity 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

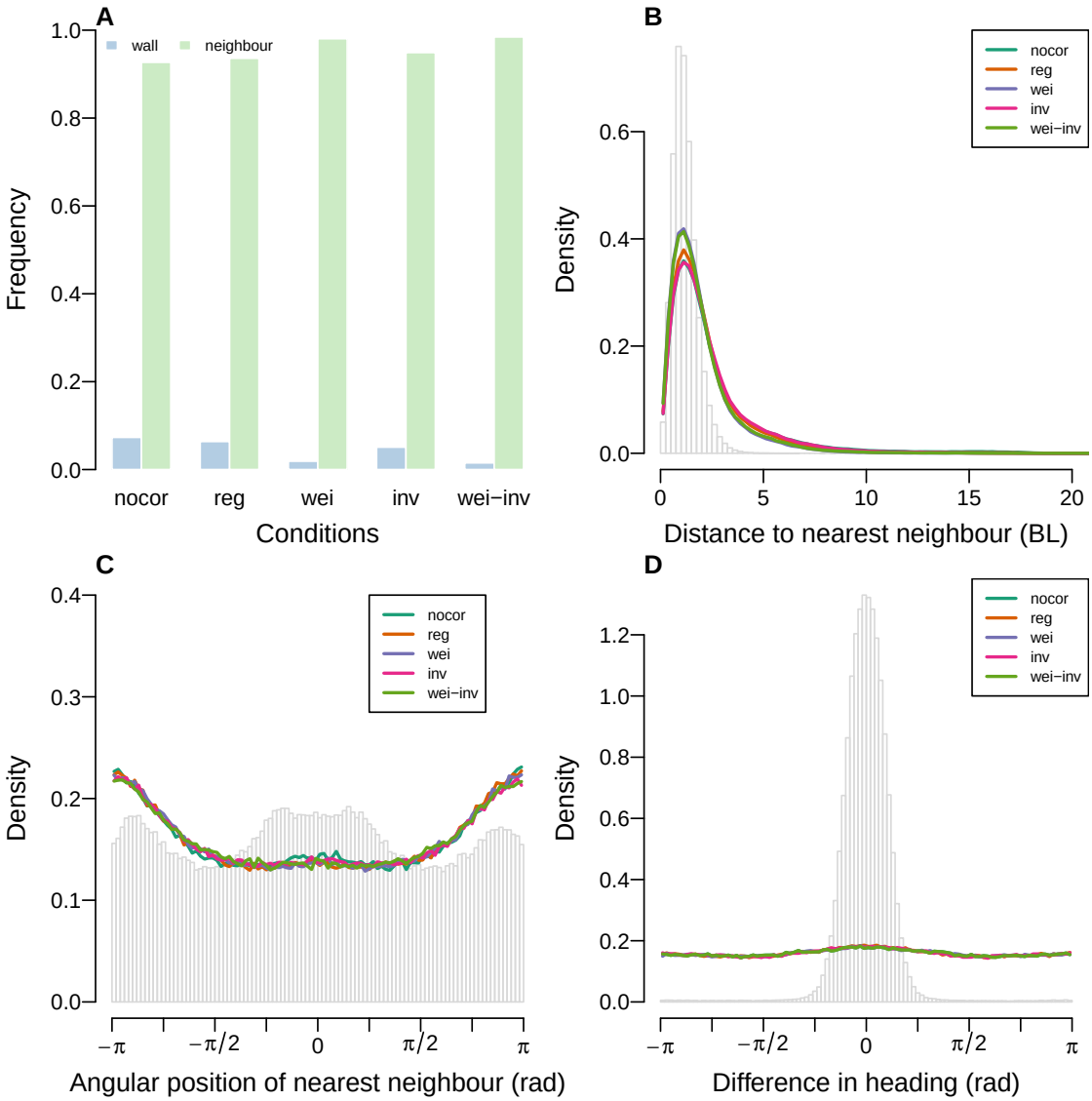


Figure 3.17: A). Relative frequency of the selected stimulus for each condition of simulation (see Table 3.3). Distributions of the (B) distance to the nearest neighbour, (C) angular position of the nearest neighbour ψ and (D) difference in heading between focal fish and nearest neighbour $\Delta\phi$ measured at the onset of bursts in experiments (bars) and simulations (curves) with 5 fish.

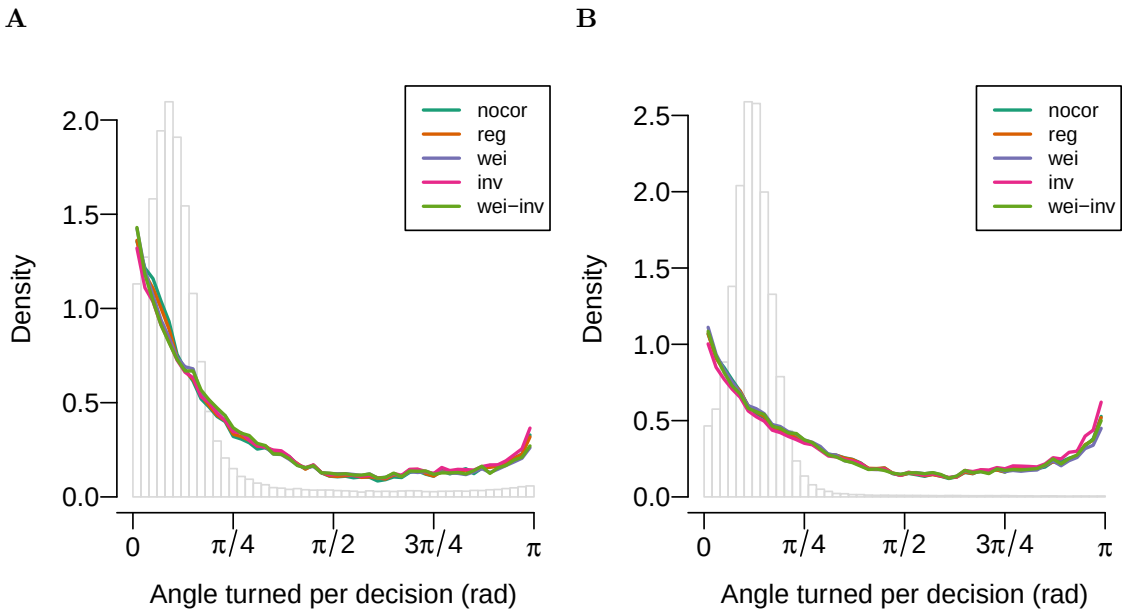


Figure 3.18: Distributions of the absolute value of the angle turned in experiments (bars) and simulations (curves) for groups of 2 (A) and 5 fish (B).

our results. Future research is required to find an efficient algorithm to make bins in 2 and 3-dimensional spaces with 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

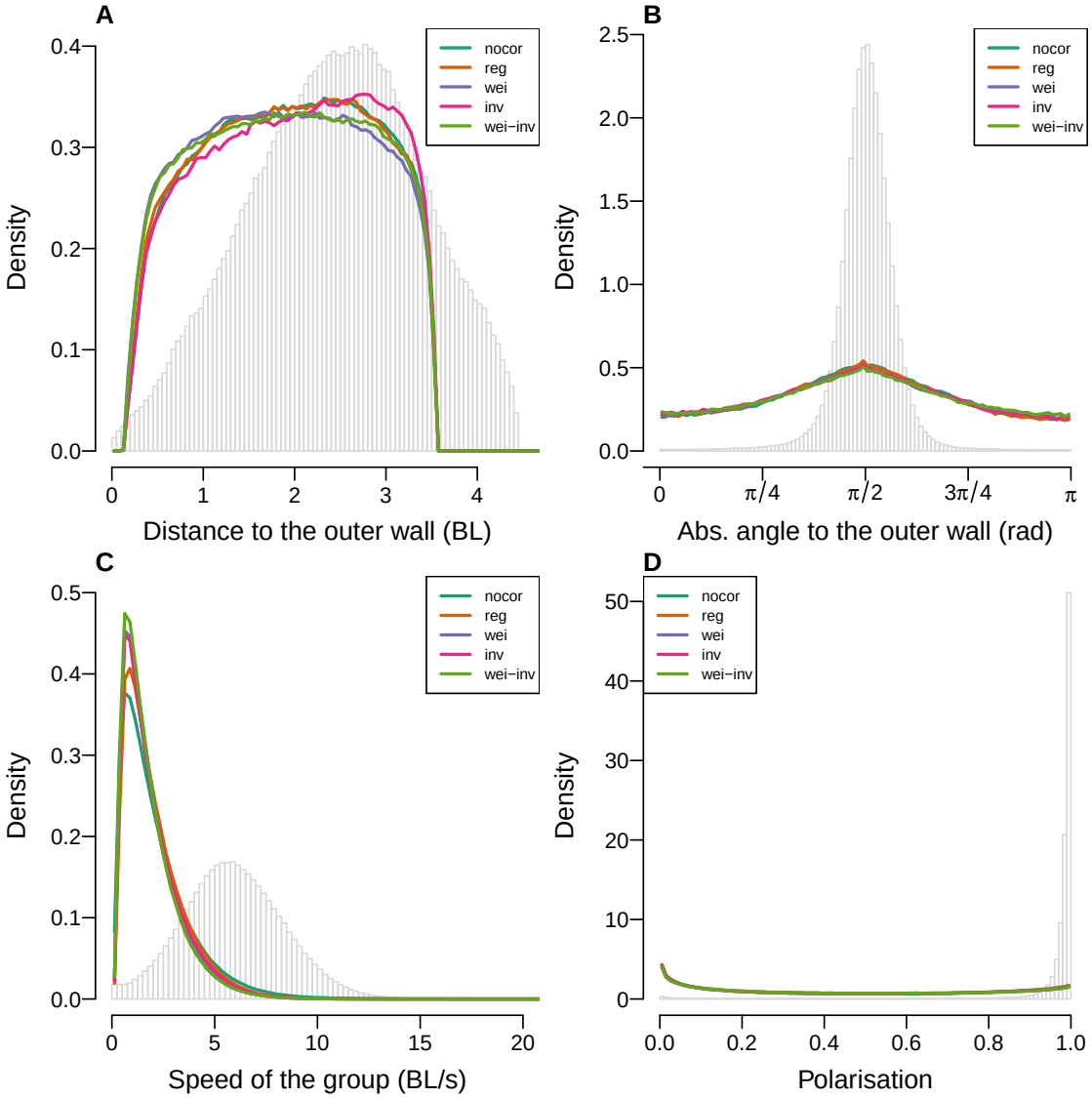


Figure 3.19: Distributions of the distance to the outer wall (A), absolute value of the angle to the outer wall (B), speed of the group (C) and polarisation of the group (D) measured on all frames in experiments (bars) and simulations (curves) with 5 fish.

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).

Part II

How does information propagate in groups of fish in response to perturbations?

Chapter 4

Domino-like propagation of collective U-turns in fish schools

Valentin Lecheval, Li Jiang, Pierre Tichit, Clément Sire, Charlotte K. Hemelrijk and Guy Theraulaz

Article published in Lecheval, V et al. 2018. "Social Conformity and Propagation of Information in Collective U-Turns of Fish Schools" Proceedings of the Royal Society of London B: Biological Sciences 285 (1877). https://doi.org/10.1098/rspb.2018.0251.

Abstract

Moving animal groups such as schools of fish or flocks of birds often undergo sudden collective changes of their travelling direction as a consequence of stochastic fluctuations in heading of the individuals. However, the mechanisms by which these behavioural fluctuations arise at the individual level and propagate within a group are still unclear. In the present study, we combine an experimental and theoretical approach to investigate spontaneous collective U-turns in groups of rummy-nose tetra (*Hemigrammus rhodostomus*) swimming in a ring-shaped tank. U-turns imply that fish switch their heading between the clockwise and anticlockwise direction. We reconstruct trajectories of individuals moving alone and in groups of different sizes. We show that the group decreases its swimming speed before a collective U-turn. This is in agreement with previous theoretical predictions showing that speed decrease facilitates an amplification of fluctuations in heading in the group, which can trigger U-turns. These collective U-turns are mostly initiated by individuals at the front of the group. Once an individual has initiated a U-turn, the new direction propagates through the group from front to back without amplification or dampening, resembling the dynamics of falling dominoes. The mean time between collective U-turns sharply increases as the size of the group increases. We develop an Ising spin model integrating anisotropic and asymmetrical interactions between fish and their tendency to follow the majority of their neighbours nonlinearly (social conformity). The model quantitatively reproduces key features of the dynamics and the frequency of collective U-turns observed in experiments.

Contribution of authors

C.K.H. and G.T. conceived and designed the study; V.L. and P.T. performed experiments; V.L. and C.S. developed the model; V.L., L.J., C.S. and P.T. analysed data; V.L., C.K.H., C.S. and G.T. wrote the paper.

4.1 Introduction

The flexible coordination of fish in schools brings important benefits (Radakov, 1973; Pitcher and Magurran, 1983; Krause and Ruxton, 2002). A striking consequence of this flexibility is the performance of rapid and coherent changes in direction of travel of schools, for instance as a reaction to a predator in the neighbourhood (Pitcher and Wyche, 1983; Pitcher and Parrish, 1993). In many species, it is only a small number of individuals that detects the danger and changes direction and speed, initiating an escape wave that propagates across the entire school (Gerlotto et al., 2006; Herbert-Read et al., 2015). Besides, sudden collective changes of the state of a school may also happen without external cause as a consequence of stochastic

effects (Tunstrøm et al., 2013). In these cases, local behavioural changes of a single individual can lead to large transitions between collective states of the school, such as between the schooling and milling states. Determining under what conditions fluctuations in individual behaviour, for instance in heading direction, emerge and propagate within a group is a key to understanding transitions between collective states in fish schools and in animal groups in general.

Only few theoretical and experimental studies have addressed these questions (Kolpas et al., 2007; Calovi et al., 2015). Calovi et al. (Calovi et al., 2015) used a data-driven model incorporating fluctuations of individual behaviour and attraction and alignment interactions among fish to investigate the response of a school to local perturbations (i.e. by an individual whose attraction and alignment behaviour differs from that of the rest of the group). They found that the responsiveness of a school is maximum near the transition region between the milling and schooling states, where the fluctuations of the polarisation are also maximal. This is entirely consistent with what happens in inert physical systems near a continuous phase transition. For instance, in magnetic systems, the polarisation of the atomic spins of a magnet near the transition point has diverging fluctuations and response to a perturbation by a magnetic field. The fluctuations of school polarisation are also expected to be strongly amplified at the transition from schooling to swarming observed when the swimming speed of individuals decreases (Gautrais et al., 2012; Calovi et al., 2014). During such a transition, the behavioural changes of a single individual are more likely to affect the collective dynamics of the school. However, the tendency of fish to conform to the speed and direction of motion of the group can also decrease the fluctuations at the level of the group with increasing group size (Herbert-Read et al., 2012). Social conformity refers to the nonlinear response of individuals to adjust their behaviour to that of the majority (Latane, 1981; Efferson et al., 2008; Morgan and Laland, 2012).

In the present work, we analyse in groups of different size under which conditions individual U-turns occur, propagate through the group, and lead to collective U-turns. We let groups of rummy-nose tetra (*Hemigrammus rhodostomus*) swim freely in a ring-shaped tank. In this set-up, fish schools can only head in two directions, clockwise or anticlockwise, and they regularly switch from one to the other. In a detailed analysis of empirical data, we reconstruct individual trajectories of fish and investigate the effect of group size on both the tendency of individuals to initiate U-turns and the collective dynamics of the U-turns. We develop an Ising-type spin model, a simple model for magnets in the physical context, to investigate the consequences on the dynamics and the propagation of information during U-turns, of the local conformity in heading, of the fish anisotropic perception of their environment, and of the asymmetric interactions between fish. We use tools and quantitative indicators from statistical physics to analyse the model. In particular, we introduce the notion of local (respectively, global/total) pseudo-energy which, in the context of a fish school, becomes a quantitative measure of the "discomfort" of an individual (respectively, of the group) with respect to the swimming direction of the other fish.

4.2 Material and Methods

4.2.1 Experimental procedures and data collection

70 rummy-nose tetras (*Hemigrammus rhodostomus*) were used in our experiments. This tropical freshwater species swims in a highly 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 (Supplementary Information (SI), Figure 4.5A). For each trial, n fish were randomly sampled from their breeding tank ($n \in \{1, 2, 4, 5, 8, 10, 20\}$). Each fish only participated in a single experiment per day. For each group size, we performed between 9 and 14 replications (see SI, Table 4.1). 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). Finally, we tracked the positions of each individual using idTracker 2.1 (Pérez-Escudero et al., 2014), except for groups of 20 fish. Details about experimental set-up, data extraction, and pre-processing are given in SI.

4.2.2 Detection and quantification of individual and collective U-turns

Since fish swim in a horizontal ring-shaped tank, their heading can be converted into a binary value: clockwise or anticlockwise. Before a collective U-turn, the fish are all moving in the same direction, clockwise or anticlockwise. When one fish changes its heading to the opposite direction, it can trigger a collective U-turn.

From the heading angle $\varphi_i(t)$ and angular position $\theta_i(t)$ of an individual i at time t (SI, Figure 4.6), the angle of the fish relative to the wall is computed as

$$\theta_{wi}(t) = \varphi_i(t) - \theta_i(t), \tag{4.1}$$

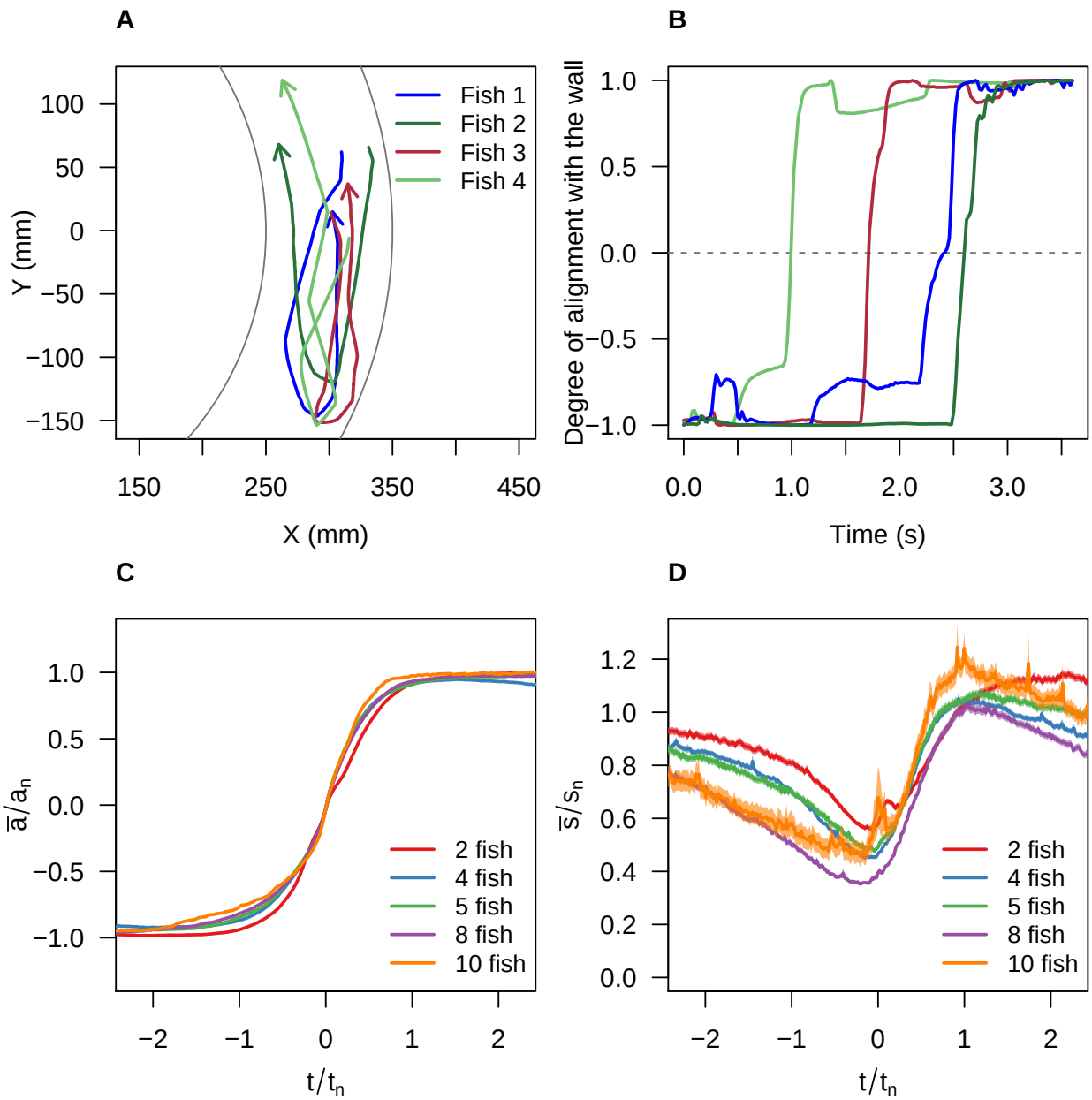


Figure 4.1: Individual trajectories (A) and degree of alignment $a_i(t)$ of fish with the wall (B) during a U-turn in a group of 4 fish. C) Normalised degree of alignment with the wall, averaged over all fish and U-turns, against the rescaled time t/t_n for groups of size n, where t_n is a measure of the mean duration of a U-turn. $t = 0$ is set when $\bar{a}/a_n = 0$. D) Average individual speed $\bar{s}$ normalised by the average speed s_n of the group, against the rescaled time t/t_n.

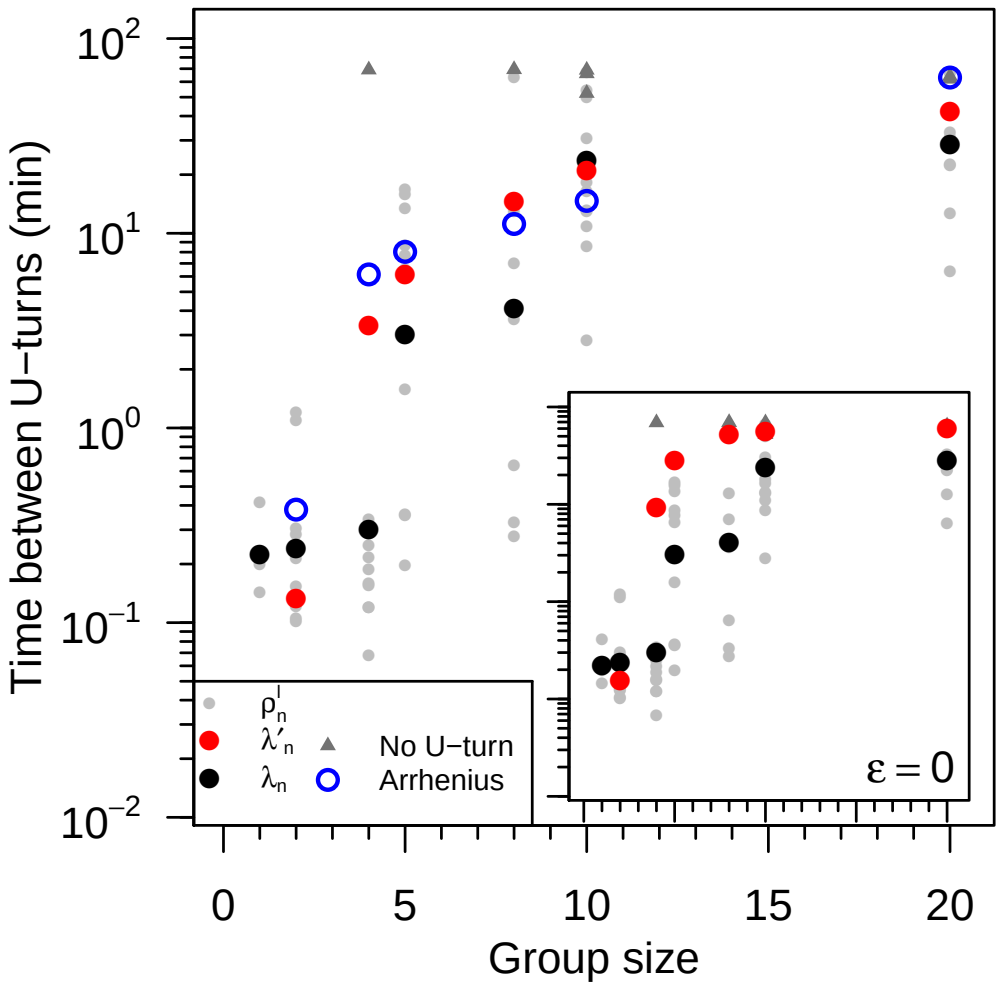


Figure 4.2: Average time between two consecutive collective U-turns as a function of group size. Average time between collective U-turns ρ_n^l in each experiment l with n fish defined as the duration of an experiment T_n^l divided by the number of collective U-turns performed during this experiment (grey dots). Experiments without any collective U-turn are indicated by grey triangles, with $\rho_n^l = T_n^l/1$. Average of the log of the time between collective U-turns over all experiments ($\lambda_n = \exp(\langle \log \rho_n^l \rangle)$; black dots) and over 1000 simulations (λ'_n; $J = 1.23$ and $\epsilon = 0.31$; red dots). Prediction of the Arrhenius law (open blue circles). Inset: results of the model without asymmetric influence ($J = 1.23$ and $\epsilon = 0$).

and thus the degree of alignment to the circular wall can be defined as

$$a_i(t) = \sin(\theta_{wi}(t)). \tag{4.2}$$

The degree of alignment $a_i(t)$ between a fish i and the outer wall is 1 when it is moving anticlockwise, -1 when moving clockwise and 0 when it is perpendicular to the wall. When a group of fish makes a collective U-turn, the degree of alignment to the wall averaged over all individuals of the group $\bar{a}(t)$ changes sign. We used this as the criterion for detecting collective U-turns automatically from the smoothed time-series of $\bar{a}(t)$ using a centred moving average over 9 consecutive frames. Figure 4.1A shows individual trajectories during a typical collective U-turn in a group of 4 fish and Figure 4.1B reports the corresponding evolution of the degrees of alignment $a_i(t)$. Further details about U-turns detection and the calculation of the quantities of interest are detailed in SI, Methods.

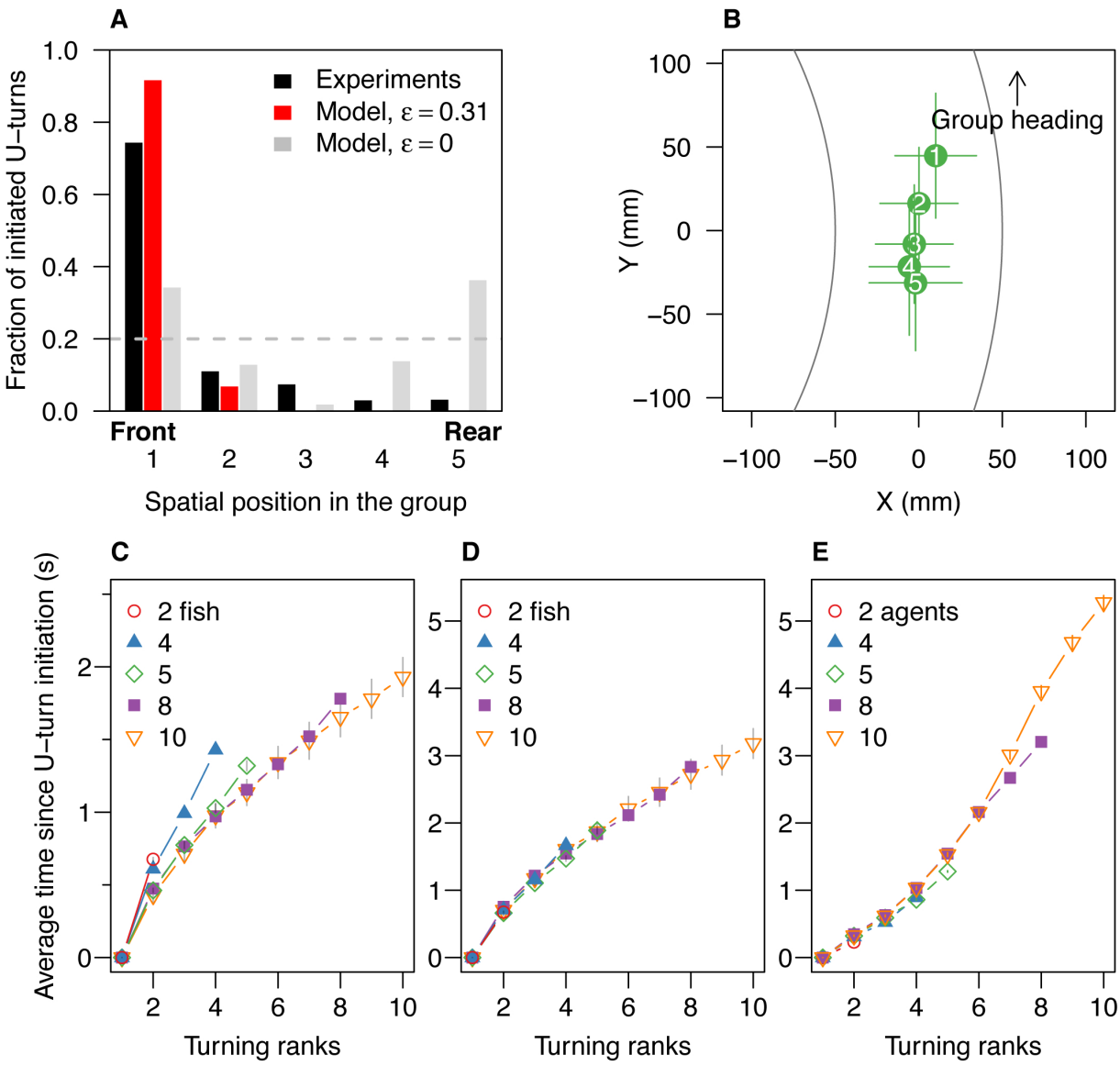


Figure 4.3: Spatiotemporal propagation of collective U-turns. A) Spatial position distribution of the initiator in groups of 5 fish in experiments (black) and in simulations with asymmetric influence ($J = 1.23$ and $\epsilon = 0.31$; red) and without asymmetric influence ($J = 1.23$ and $\epsilon = 0$; grey). Spatial positions go from 1 (position at the front) to 5 (position at rear). The dashed line shows the uniform distribution $1/5 = 0.2$, when spatial position has no effect on the initiation of collective U-turns. B) Average relative positions ($\pm$ sd) of all individuals at initiation of collective U-turns, ranked by order of turning (i.e. rank 1 is initiator) in groups of 5. Positions have been corrected so that all groups move in the same direction, with the outer wall at their right hand-side. The origin of the coordinate system is set to the centroid of the average positions of individuals. Average time interval since the beginning of a collective U-turn as a function of turning rank and group size in experiments (C and D) and in simulations (E). In D, the time is scaled by the factor $r_n = s_n/s_2$, where s_n is the average speed of groups of size n, revealing a behaviour almost independent of n.

4.3 Results

4.3.1 Spatiotemporal dynamics of collective U-turn

Hemigrammus rhodostomus fish form highly cohesive schools during our experiments (SI, Figure 4.7A) and adjust their speed and heading to that of their group members. In a former study (Calovi et al., 2017), we have shown that this is achieved through attraction and alignment interactions that have been precisely measured. Figure 4.2 indicates that the average time interval between two U-turns in groups of 10 fish (one U-turn every 20 min) is two orders of magnitude larger than in groups of 2 fish (one U-turn every 0.2 min). In experiments in which no collective U-turn was observed (grey triangles on Figure 4.2), we took the total period of observation as the interval until the next U-turn. Therefore, the average time λ_n between U-turns measured in groups of 4, 8, 10, and 20 fish are slightly underestimated. Thus, as group size increases, the number of collective U-turns decreases, because the propensities of a fish to initiate and propagate a U-turn decrease (SI, Figure 4.8). Like in many other species, individual fish tend to adopt the behaviour of the majority of the group members and thus inhibit the initiation of U-turns (Herbert-Read et al., 2012).

As shown in Figure 4.1C, the dynamics of collective U-turns, and in particular the evolution of the mean alignment $\bar{a}(t)$, is similar for all group sizes, once time is rescaled by the mean U-turn duration (see SI for the method used to compute the scaling parameter t_n, which is an effective measure of the U-turn duration). In SI, Figure 4.9 shows that t_n increases approximately linearly with group size n. In groups of all sizes, fish progressively decrease their speed before turning collectively and accelerating sharply (Figure 4.1D). The duration of this deceleration (and then acceleration) phase is much longer than the time for the group to complete a U-turn (compare Figure 4.1C and Figure 4.1D). Moreover, the speed minimum of the group in Figure 4.1D is reached near the midpoint of the U-turn, when $t = 0$ and the mean alignment is $\bar{a} = 0$ in Figure 4.1C.

Our results show that the propagation of information is on average sequential, both in space and time. This resembles a chain of falling dominoes, for which the time interval between successive falls is constant, without any positive feedback.

Collective U-turns are usually initiated at the front of the school and the change of swimming direction propagates towards the rear (Figures 4.3A and B and SI, Figures 4.10 and 4.11) and Table 4.2 for statistical tests). Note that Figure 4.3B does not show the actual shape of groups but only the average and relative positions of fish. In particular, the x-coordinates are not perfectly centred on 0 (the centroid of the average positions) for all turning ranks because the foremost fish tends to swim significantly closer to the outer wall than the fish swimming at the rear, in line with previous results in groups of two fish in the same species (Calovi et al., 2017) (SI, Table 4.3 for statistical tests). At the time of the turn of each individual, fish almost turn at the same location as the previous ranks, respectively to the y-coordinates (SI, Figure 4.12 and Tables 4.4 and 4.5).

Although the time interval between the turning of the first and the second fish is longer than it is for others, the time interval between the successive turns of individuals is almost constant in a given group size (Figure 4.3C). This can be derived from the fact that the time since the initiation of the collective U-turn increases linearly with the turning rank. The linear propagation of information in all group sizes suggests that there is no amplification of the individual tendency to perform a U-turn: the time between two successive individuals performing U-turns does not decrease with the number of fish that already performed a U-turn. This suggests that individuals only pay attention to a small number of neighbours at a given time as was shown in (Jiang et al., 2017, see Appendix B).

The mean time interval between two successive individual U-turns decreases with group size (see Figure 4.3C where the slopes decrease with n, or SI, Figure 4.13). However, when these time intervals are multiplied by a factor r_n proportional to the average speed s_n of groups of size n ($r_n = s_n/s_2$), they collapse on the same curve (Figure 4.3D). This suggests that the shorter reaction time of fish in larger groups is mostly due to their faster swimming speed. Larger groups swim faster (SI, Figure 4.7B), presumably because fish are interacting with a greater number of neighbours and are closer to each other (SI, Figure 4.7C). Indeed, fish have to avoid collisions with obstacles and other fish and the faster they swim, the shorter their reaction time, a well-known psycho-physiological principle (Smeets and Brenner, 1994).

In summary, our results show that U-turns are mostly initiated by fish located at the front of the school. U-turns are preceded by a decrease in the speed of the group. Once the U-turn has been initiated, the wave of turning propagates in a sequential way, suggesting that fish mainly copy the behaviour of a small number of individuals.

4.4 Modelling collective U-turns

4.4.1 Model description

We now introduce an Ising-type spin model (Castellano et al., 2009; Brendel et al., 2003) to better understand the impact of social conformity, anisotropy and asymmetry of interactions, and group size on the propagation of information during U-turns. Each agent i has a direction of motion $d_i \in \{-1, 1\}$ with $d_i = -1$ representing swimming clockwise and $d_i = 1$ swimming anticlockwise. A U-turn performed by an agent i corresponds to a transition from d_i to $-d_i$. In the model, the relative positions of individuals and the interaction network (i.e. the influential neighbours η_i of an agent i) are kept fixed in time (SI, Figure 4.14). Agents are positioned in staggered rows (SI, Figure 4.7D for experimental data supporting an oblong shape that becomes longer when the group size increases, as previously found by others, e.g. (Hemelrijk et al., 2010)) and only interact with their direct neighbours. The strength of interactions between an agent i and its neighbour j is weighted by a parameter α_{ij} that depends on the spatial position of j relatively to i. α_{ij} controls the anisotropy and asymmetry of the interactions between individuals, assuming that fish react stronger to frontal stimuli, in agreement with previous experimental results on *H. rhodostomus* (Calovi et al., 2017). We define $\alpha_{ij} = 1+\epsilon$ when agent j is in front of agent i, $\alpha_{ij} = 1$ if j is at the side of i, and $\alpha_{ij} = 1-\epsilon$ if j is behind i, where the asymmetry coefficient $\epsilon \in [0, 1]$ is kept constant for all group sizes.

The propensity of an individual i to make a U-turn depends on the state of its neighbours η_i and on the interaction matrix α_{ij}. The "discomfort" E_i of an agent i in a state d_i is given by

$$E_i = -d_i \sum_{j\in\eta_i} J_{ij} d_j, \quad J_{ij} = \alpha_{ij} J, \tag{4.3}$$

with J_{ij} the coupling constant between two neighbours i and j, set by the two positive parameters of the model, ϵ and $J > 0$. When the anisotropy of perception and asymmetry of interactions are ignored ($\epsilon = 0$), $\alpha_{ij} = 1$ for all neighbouring pairs (i, j). E_i is minimal (and negative) when the focal fish i and its neighbours point in the same direction, and maximal (and positive) if the focal fish i points in the opposite direction of its aligned neighbours. A small value of $|E_i|$ corresponds to its neighbours pointing in directions nearly averaging to zero.

If an individual flips ($d'_i = -d_i$), the new discomfort is $E'_i = d_i \sum_{j=\{\eta_i\}} J_{ij} d_j$, and we have

$$\Delta E_i = E'_i - E_i, = 2Jd_i \sum_{j\in\eta_i} \alpha_{ij} d_j. \tag{4.4}$$

$\Delta E_i < 0$ when an agent flips to the most common state of its neighbours, whereas $\Delta E_i > 0$ when it flips to the state opposite to this most common state. In this $\epsilon = 0$ case,

$$\mathrm{E} = \frac{1}{2} \sum_{i=1}^{n} E_i \tag{4.5}$$

corresponds to the total actual *energy* of the magnetic system. In this context, the fully polarised state where all fish are aligned corresponds to the so-called ground-state energy, the lowest possible energy of the system. For $\epsilon \neq 0$, the asymmetry between the perception of i by j and that of j by i breaks this interpretation in terms of energy (Calovi et al., 2017). Yet, for $\epsilon > 0$, it is still useful to define E as a pseudo-energy, as will be discussed later, since it remains a good indicator of the collective discomfort of the group, i.e. the lack of heading alignment within the group.

The dynamics of the model is investigated using Monte Carlo numerical simulations inspired from the Glauber dynamics (Glauber, 1963). Within this algorithm, at each time step $t_{k+1} = t_k + 1/n$ (n is the number of agents), an agent is drawn randomly and turns (update d_i to $d'_i = -d_i$) with the acceptance probability

$$P = \frac{1}{2} - \frac{1}{2} \tanh\left(\frac{\Delta E_i}{2T}\right), \tag{4.6}$$

which is a sigmoid, going from $P \to 1$ for $\Delta E_i \to -\infty$ (maximal acceptance if the discomfort decreases sharply), to $P \to 0$ for $\Delta E_i \to +\infty$ (no direction switch if the discomfort would increase dramatically). In equation 4.6, T plays the role of the temperature and we chose $T = 1$. Indeed, since ΔE_i is proportional to J, the probability P only depends on the parameter $J' = J/T$, and T can then be absorbed in the constant J.

The acceptance probability P represents the social conformity in our model and its strength (i.e. the nonlinearity of P) is mainly controlled by the parameter J (SI, Figure 4.15B). For large $J > 0$, this dynamics

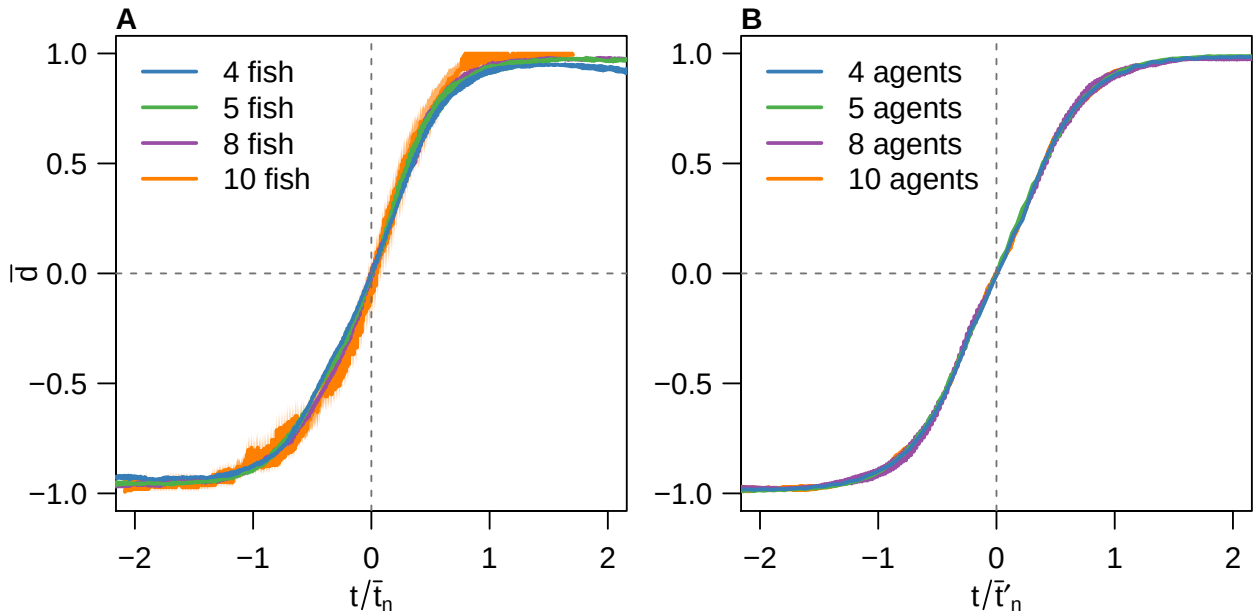


Figure 4.4: Mean swimming direction $\bar{d}$ averaged over all collective U-turns as a function of scaled time $t/\bar{t}_n$ and $t/\bar{t}'_n$ for all group sizes in (A) experiments and (B) model. t_n and $\bar{t}'_n$ are obtained by data scaling (see SI, Methods). The shadows stand for the standard error. In contrast to Figure 4.1, $t = 0$ is set to the time $(t_E - t_S)/2$ (experiments) or $(t'_E - t'_S)/2$ (model), where t_S stands for the start of the collective U-turn (first frame where at least one direction $-d_i \times d_0$ is positive) and t_E for the end of the collective U-turn (first frame where all directions $-d_i \times d_0$ are positive). In A, time has also been shifted so that $\bar{d}(t = 0) = 0$.

will favour the emergence of strongly polarised states, while for $J = 0$, all fish directions will appear with the same probability during the dynamics. In physics, such a model favouring alignment between close spins is known as the Ising model, which crudely describes ferromagnetic materials, i.e. magnets.

In summary, J controls the directional stiffness of the fish group, while ϵ describes the fish anisotropic perception of their environment and the asymmetric interactions between fish. After inspecting the (J, ϵ) parameter space (see SI, section 4.A.5), we find that the parameter values $J = 1.23$ and $\epsilon = 0.31$ lead to a fair agreement between the model and experimental data, as will be shown in the next section.

4.4.2 Simulation results versus experimental data

Our model quantitatively reproduces the effect of group size on the dynamics of collective U-turns (Figures 4.2 and SI, 4.8). This suggests that the tendency of individuals to initiate U-turns and move in the opposite direction of the whole decreases with group size. However, note the poor agreement between simulations and experimental data in groups of 4. One explanation for this may be the age and body size of the fish, since body size influences the strength of interactions between fish (Romenskyy et al., 2017) (SI, Table 4.1). It is possible to set a different coupling constant J for each group size to account for this effect, with the risk of overfitting (SI, Figures 4.16A and B).

Even though there is no strict notion of energy in our model when $\epsilon > 0$, we can still compute the mean pseudo-energy barrier $\Delta\mathrm{E}_n$ as a function of group size n. It is defined as the mean difference between the maximum value of the pseudo-energy E during the U-turn and the reference energy computed when all the agents have the same direction (i.e. before and after a U-turn). With the interpretation of E (respectively, E_i) as a quantitative indicator of the discomfort of the group (respectively, of the fish i), the (pseudo) energy barrier $\Delta\mathrm{E}_n$ is hence a measure of the collective effort of the group to switch direction. We find that the energy barrier $\Delta\mathrm{E}_n$ increases sublinearly with group size n (SI, Figure 4.17). We then expect that the higher/larger is the (pseudo) energy barrier $\Delta\mathrm{E}_n$, the more difficult it will be for the group to perform a U-turn, as it must necessarily pass through an intermediate state of greater discomfort as the group size n increases. As a consequence, the average time between U-turns is also expected to increase as n and the (pseudo) energy barrier $\Delta\mathrm{E}_n$ both increase. In fact, for $\epsilon = 0$, for which E represents a true energy, this mean time interval between direction changes is exactly given by the Arrhenius law, which can be analytically proven for our spin model evolving according to the Glauber Monte Carlo dynamics.

In physical chemistry, the Arrhenius law describes for instance the switching time between two states A and B of a molecule, separated by an energy barrier associated to an intermediate state through which the molecules must necessarily pass to go from state A to state B. The Arrhenius law (Atkins and De Paula, 2011) stipulates that the mean transition time τ between two states separated by an energy barrier $\Delta \mathrm{E}_n$ grows like

$$\tau = \tau_0 \exp\left(\frac{\Delta \mathrm{E}_n}{T}\right), \tag{4.7}$$

where τ_0 is a time scale independent of n, and T is the same temperature as the one appearing in equation 4.6 (here, $T = 1$). Despite the fact that $\epsilon > 0$, for which E is not anymore a true energy, we still find in Figure 4.2 that the (pseudo) Arrhenius law reproduces fairly well the experimental mean interval between U-turns as a function of group size n, explaining the wide range of observed time intervals, but with a modified constant effective temperature $T \approx 1.9$ (and $\tau_0 \approx 0.09\,\mathrm{min}$). It is remarkable that the mean time between U-turns (a purely dynamical quantity) grows exponentially fast with $\Delta \mathrm{E}_n$ (the pseudo-energy difference between two static configurations), considering that both quantities are measured in two completely independent ways.

The sequential propagation of information is also reproduced well by the simulations of the model, both in space (Figure 4.3A and SI, Figures 4.10 and time (Figure 4.3C and SI, Figure 4.18). When the perception of agents is isotropic (i.e. $\epsilon = 0$), the location of the U-turn initiation is no longer mainly at the front of the group but depends on the number of influencing neighbours (Figure 4.3A and SI, Figure 4.14). The lower the number of influential neighbours, the higher the number of collective U-turns. For groups of 5 and $\epsilon = 0$, the agents triggering most of the U-turns are the first and last agents because they only have two influencing neighbours.

Regarding the propagation in time, simulations reproduce the linear propagation of information at the individual scale, except for the largest group sizes. This can be improved by changing the topology of the interaction network of these group sizes (SI, Figures 4.16C and D). Figure 4.4A and B show that once rescaled by the U-turn duration, the average direction profile is nearly independent of the group size, and that the model prediction is in good agreement with experimental data. It takes about the same amount of time to turn the first and second half of the fish, both in experiments and in the model, although the first half of the fish is slightly slower to turn than the second half in experiments. This is consistent with the results reported on Figure 4.3C, where the interval between the turning of the first and the second fish was longer than between the turns of the following fish. The durations of collective U-turns are Log-normally distributed, both in experiments and in the model (SI, Figure 4.19).

Despite its simplicity and having only two free parameters (J and ϵ), our model reproduces quantitatively the experimental findings, both at the collective scale (the frequency of collective U-turns, average direction profile, duration of U-turns...) and at the individual scale (the spatiotemporal features of the propagation of information). Note that a linear response of the agents to their neighbours cannot reproduce the order of magnitude of the U-turn durations measured in the experiments (SI, Figure 4.15). Social conformity is thus a good candidate as an individual mechanism underlying the observed patterns including the time intervals between successive collective U-turns for different group sizes, the distribution of the U-turn duration, and the spatial propagation of information.

4.5 Discussion

How information propagates among individuals and determines behavioural cascades is crucial to understand the emergence of collective decisions and transitions between collective states in animal groups (Giardina, 2008; Sumpter et al., 2008; Wang et al., 2012; Attanasi, Cavagna, Del Castello, Giardina, Grigera, Jelic, Melillo, Parisi, Pohl, Shen and Viale, 2014). Here, we addressed these questions by analysing the spontaneous collective U-turns in fish schools.

We find that collective U-turns are preceded by a slowing down period. It has been shown in other fish species that speed controls alignment between individuals (Gautrais et al., 2012), leading slow groups to be less polarised than fast groups (Steven V. Viscido et al., 2004; Hemelrijk and Hildenbrandt, 2008; Tunstrøm et al., 2013; Calovi et al., 2014). In general, at slower speed, there is less inertia to turn, resulting in weaker polarisation (Kunz and Hemelrijk, 2003; Hemelrijk et al., 2010) and thus an increase of the fluctuations in the swimming direction of the fish (Marconi et al., 2008). Moreover, as the fish speed decreases, the fish school is in a state closer to the transition between the schooling (strong alignment) and swarming (weak alignment) states, where (Calovi et al., 2015) have shown that both fluctuations in fish orientation and the sensitivity of the school to a perturbation increase. It is therefore not surprising that U-turns occur after the group has slowed down.

U-turns are mostly initiated by the fish located at the front of the group. At the front, individuals experience a lesser influence from the other fish. This is due to the perception anisotropy which results in

individuals interacting more strongly with a neighbour ahead than behind. Therefore, frontal individuals are more subject to heading fluctuations and less inhibited to initiate U-turns. Similarly, in starling flocks, the birds that initiate changes in collective travelling direction are found at the edges of the flock (Attanasi et al., 2015).

We found no evidence for dampening or amplification of information as fish adopt a new direction of motion. Moreover, on average, turning information propagates faster in larger groups: 0.19 s per individual in groups of 10 fish, and 0.26 s per individual in groups of 5 fish (SI, Figure 4.13). This appears to be the consequence of the increase of the swimming speed with group size, which requires that individuals react faster. Indeed, our results show that the interval between successive turns of individuals during a collective U-turn decreases with swimming speed, although distance between individuals may also play a role (Jiang et al., 2017). However, the mean time interval between successive individual U-turns is almost constant and independent of the group size, once time has been rescaled by the group velocity. This points to a domino-like propagation of the new direction of motion across the group. This sequential spatiotemporal propagation of information also suggests that each fish interacts with a small number of neighbours.

We found that the level of homogeneity in the direction of motion of the schools increases with group size resulting in a lower number of collective U-turns. This phenomenon has been previously described in other fish species (Day et al., 2001; Herbert-Read et al., 2012) as well as in locusts in a similar set-up (Buhl et al., 2006).

We developed an Ising-type spin model in which fish adopt probabilistically the direction of the majority of their neighbours, in a nonlinear way (social conformity) influenced by the anisotropic and asymmetrical interactions between fish. Since the probability that a fish chooses a direction is a nonlinear function of the number of other fish having already chosen this direction, as previously shown (Sumpter and Pratt, 2009; Ward et al., 2008), it is thus more difficult for a fish to initiate or propagate a U-turn the larger the number of fish swimming in the opposite direction (Efferson et al., 2008). The model also introduces quantitative indicators of the individual and collective discomfort (lack of alignment of heading among group members), the latter being represented by a measure of global pseudo-energy of the group. Larger groups have to overcome a larger pseudo-energy barrier to switch between the clockwise and anticlockwise fully polarised states. In physics and chemistry, the fast exponential increase of the switching time between two states as a function of this energy barrier is described by the Arrhenius law, which can be proven using the tools of statistical physics. We find that direct numerical simulations of the model and an effective Arrhenius law both quantitatively reproduce the sharp increase of the mean time between U-turns as the group size increases. The model also shows that asymmetric interactions and the anisotropic perception of fish are not essential to explain the decrease of collective fluctuations and hence the U-turns frequency as the group size increases. Social conformity (Latane, 1981; Morgan and Laland, 2012) (controlled by the magnitude of our parameter J) suffices to cause fewer fluctuations with increasing group size, leading to an increased robustness of the polarised state ("protected" by increasing pseudo-energy barriers).

Moreover, our model reveals that the front to back propagation of information results from the perception anisotropy and asymmetry of the fish (the ϵ parameter). Without perception anisotropy and asymmetry, U-turns are initiated by the fish that have fewer influential neighbours (in our simulations, those are the fish at the boundary of the group – all individuals would have the same probability to initiate a U-turn with periodic boundary conditions) and propagated to their neighbours without favouring any direction. Finally, the duration of a U-turn as a function of group size is quantitatively reproduced by the model, while the simulated mean direction temporal profiles during U-turns are very similar to the experimental ones, and are independent of the group size, once time is properly rescaled by the mean U-turn duration for the corresponding group size.

In summary, our work supports that social conformity, asymmetric interactions, and the anisotropic perception of fish are key to the sequential propagation of information without dampening in fish schools, at least in the small group sizes considered. Future work will be needed to disentangle the respective roles of the network topology and the actual functional forms of social interactions between fish on the propagation of information.

Supplementary Information

4.A Experimental procedures and data collection

Fish were purchased from Amazonie Labège (`http://www.amazonie.com`) in Toulouse, France. They were kept in 150 L aquariums on a 12:12 hour, dark:light photoperiod, at 27.5° C (±0.8° C) and were fed *ad libitum* with fish flakes. The body length of the fish in the experiments was on average 3.4 cm (± 0.44 cm).

The experimental tank (120×120 cm) was made of glass and was set on top of a box to reduce vibrations. It was surrounded by four opaque white curtains and illuminated homogeneously by four LED light panels. Inside an experimental tank, a ring-shaped corridor 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 4.5A). The corridor was 10 cm wide with a circular outer wall of radius 35 cm. 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. Fish were introduced in and acclimatised to the experimental tank during a period of 10 minutes before the trial starts. During each trial of one hour, individuals were swimming freely without external perturbation. Note that six experiments with a single fish have been discarded because of the inactivity of the individuals.

Table 4.1: **Number of trials, total duration of trials, number of collective U-turns and average body length of individuals for each group size.**

Group Size	Number of trials	Total duration	Total number of collective U-turns	Body length (mm, mean ± se)
1	4	260 min	1058	33.1 ± 1.8
2	10	652 min	1135	33.3 ± 0.8
4	10	684 min	1868	36.1 ± 0.6
5	10	543 min	500	31.5 ± 0.3
8	9	602 min	459	35.9 ± 0.6
10	14	832 min	49	33.4 ± 0.4
20	11	703 min	30	Not available

4.A.1 Data extraction and pre-processing

Sometimes, fish were misidentified by the tracking software, 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. For groups of 20 fish, only the number of collective U-turns and the time interval between two consecutive U-turns have been recorded.

Time series of positions were converted from pixels to meters and the origin of the coordinate system was 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. Table 4.1 summarises the data collected in our study.

4.A.2 Detection and quantification of individual and collective U-turns

A collective U-turn in a group of n fish starts when the degree of alignment to the wall $a_i(t)$ of the fish i that initiates the U-turn is 0 and it ends when the degree of alignment to the wall $a_j(t)$ of the last fish j that turns is 0. For each collective U-turn, we ranked the order with which each individual turned r_i (where $r_i = 1$ refers to the individual i initiating it) and the spatial positions of each individual at the initiation of the U-turn. In order to compare the spatial positions of individuals swimming in groups of various shapes, we compute at the beginning of the U-turns $\Phi_i = -(\theta_i - \theta_f)/(\theta_f - \theta_l)$, where the angle $\theta_i - \theta_f$ between each individual and the fish in front of the group, normalised by the angle $\theta_f - \theta_l$ between the first and last fish. We discretised $\Phi \in [0, 1]$ in n cells with increasing indices and the spatial position π_i is given by the index of the cell that contains Φ_i. π_i is 1 if an individual is very close to the front of the group when the first individual turns and n if it is close to the back of the group at this time.

To compute the ranks of turning and the spatial positions of individuals at the initiation of the U-turns, we needed to make sure that fish were responding to the initiation of a specific U-turn (and not to a previous

Table 4.2: Results of the linear mixed effects models fitted on each group size to test the effect of the rank of turning on the position regarding the y-coordinates at initiation of collective U-turns (see Figure 4.11). Collective U-turns with missing positions at initiation have been discarded.

Group size	Number of collective U-turns considered	Estimated slope (±se)	χ^2	p-value
2	1114	-34.83 ±4.89	18.72	< 0.001
4	1655	-25.09 ±1.78	29.19	< 0.001
5	472	-17.42 ±2.59	18.03	< 0.001
8	272	-11.34 ±0.76	45.25	< 0.001
10	33	-11.52 ±2.25	11.85	< 0.001

U-turn very close in time). Therefore, we only considered situations where fish were swimming for at least 2 seconds in the same direction before and after the U-turns.

Failed collective U-turns (i.e. U-turns initiated by one or more individuals that are not fully propagated) are also detected. A failed U-turn is detected when the average of the sign of the degree of alignment is not $|1|$ and when the sign of the average degree of alignment does not switch. To address possible noise in experimental data, the average of the sign of the degree of alignment has to be different from $|1|$ during at least 25 frames (half a second).

For a given group size, we compute the average rate of U-turns (failed or not) initiated per individual as

$$\frac{u_n + f_n}{nT_n}, \tag{4.8}$$

with n the group size, u_n the number of collective U-turns (fully propagated), f_n the number of failed collective U-turns and T_n the duration of the experiments. The probability that a collective U-turn is fully propagated is computed by

$$\frac{u_n}{u_n + f_n}. \tag{4.9}$$

4.A.3 Data scaling

Data scaling shown in Figure 1 is obtained by finding the value of the time parameter t_n that minimises the least-square error between the normalised degree of alignment with the wall averaged over the U-turns at a given group size n and that averaged over the U-turns of a group size of reference (namely, groups of 5 fish). To compute error bars, t_n has been bootstrapped by applying the least square method randomising the collective U-turns considered in the averaged normalised degree of alignment for each group size. For each group size, $N = 1000$ bootstrapped samples have been obtained. The same method has been used in Figure 4.

4.A.4 Statistical tests

We used R (R Core Team, 2016) and the package `lme4` (Bates et al., 2015) to perform a linear mixed effects analysis (with restricted maximum likelihood) of the relationship between x and y-coordinates (respectively) and ranks of turning (fixed effect). As random effect, we have intercept for the experiment as well as by-experiment random slopes to account for the non-independence of the U-turns within a group size. The examinations of residuals did not reveal any obvious deviations from homoscedasticity or normality. P-values were obtained by likelihood ratio tests of the full model with the fixed effect against the null model with intercept and random effect only. The slope estimated with restricted maximum likelihood and the result of the likelihood ratio tests are reported in Tables 4.3 and 4.2.

4.A.5 Model implementation

For given J and ϵ, we compute numerically the prediction for the number of collective U-turns u'_n for a group of size n made during T' Monte-Carlo time steps. We define the error function

$$\Delta = \sum_n (\frac{\tau'_n}{\tau_n} - 1)^2 \tag{4.10}$$

Table 4.3: Results of the linear mixed effects models fitted on each group size to test the effect of the rank of turning on the position regarding the x-coordinates at initiation of collective U-turns (see Figure 4.11). Collective U-turns with missing positions at initiation have been discarded.

Group size	Number of collective U-turns considered	Estimated slope (±se)	χ^2	p-value
2	1114	-12.04 ±4.89	5.13	0.02
4	1655	-4.04 ±0.41	56.97	< 0.001
5	472	-2.04 ±1.28	2.27	0.13
8	272	-0.95 ±0.42	19.44	< 0.001
10	33	-0.19 ±0.50	0.14	0.71

Table 4.4: Results of the linear mixed models fitted on each group size to test the effect of the rank of turning on the position regarding the y-coordinates when the individual turns (see Figure 4.12). Collective U-turns with missing positions at fish turns have been discarded.

Group size	Number of collective U-turns considered	Estimated slope (±se)	χ^2	p-value
2	1114	-5.86 ± 2.43	4.87	0.03
4	1655	-5.54 ± 1.51	8.85	< 0.001
5	472	-2.13 ± 1.71	1.29	0.26
8	272	1.40 ± 2.47	0.27	0.60
10	33	1.32 ± 1.59	0.54	0.463

Table 4.5: Results of the linear mixed models fitted on each group size to test the effect of the rank of turning on the position regarding the x-coordinates when the individual turns (see Figure 4.12). Collective U-turns with missing positions at fish turns have been discarded.

Group size	Number of collective U-turns considered	Estimated slope (±se)	χ^2	p-value
2	1114	-9.68 ± 4.00	5.00	0.03
4	1655	-3.91 ± 0.28	36.24	< 0.001
5	472	-3.51 ± 2.19	2.42	0.12
8	272	-2.98 ± 0.75	9.22	< 0.001
10	33	-5.36 ± 1.26	16.78	< 0.001

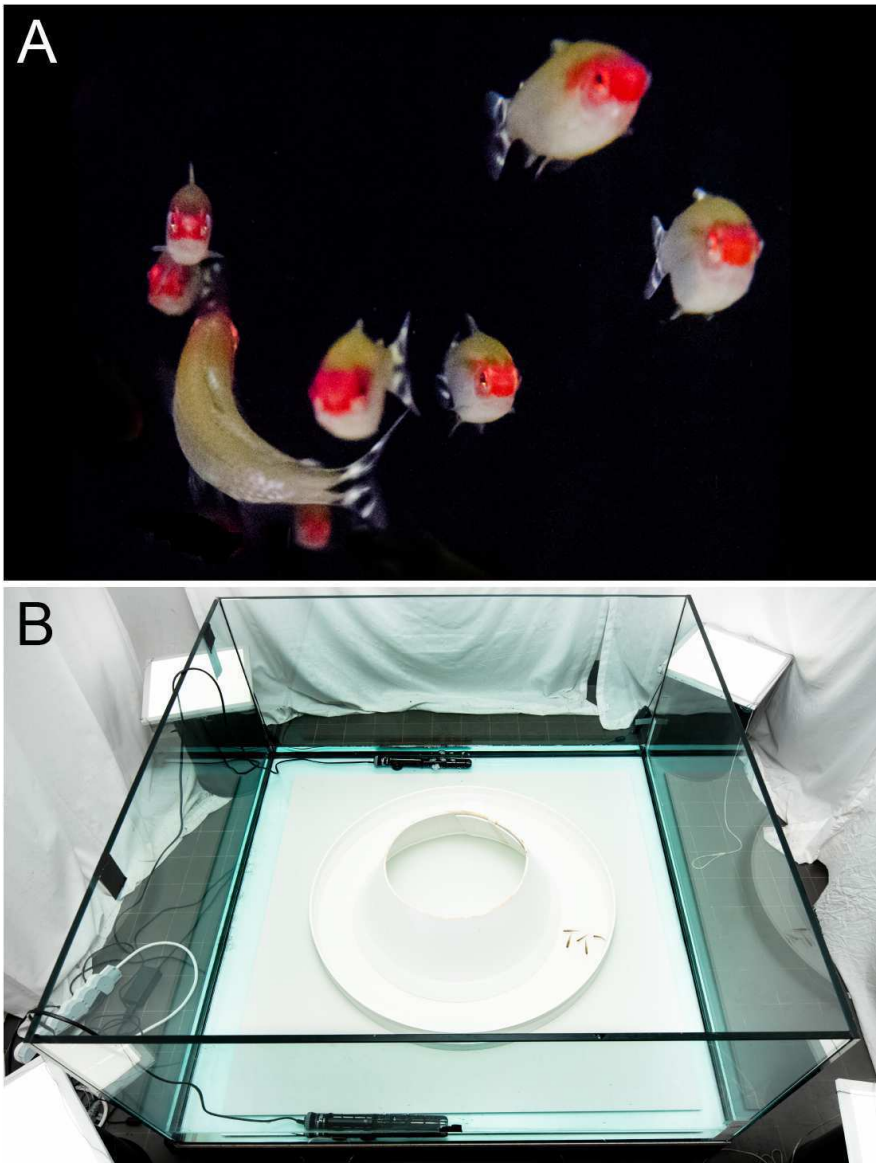


Figure 4.5: Experimental set-up. (A) A photo of a spontaneous U-turn initiated by a single fish in a group of eight *Hemigrammus rhodostomus* fish, (B) Experimental ring-shaped tank, credits to David Villa ScienceImage/CBI/CNRS, Toulouse.

with $\tau_n = \frac{u_n}{T_n}$ the experimental rate of collective U-turns (with T_n the total duration of all the experiments of the group size n, in minutes), $\tau'_n = \frac{u'_n}{T' t_0}$ the rate of collective U-turns in simulations. t_0 has the dimension of a time and translates Monte-Carlo time into actual experimental minutes, and is determined by minimising the error Δ, i.e. by solving the equation $\frac{\partial \Delta}{\partial t_0} = 0$.

The model has been implemented in R (and run with R 3.3.1) with a C++ subroutine using the package `Rcpp` (Eddelbuettel and Francois, 2011; Eddelbuettel, 2013). The sensitivity analysis has been conducted with parallel computing using the R package `parallel` (R Core Team, 2016).

4.B Supplemental figures

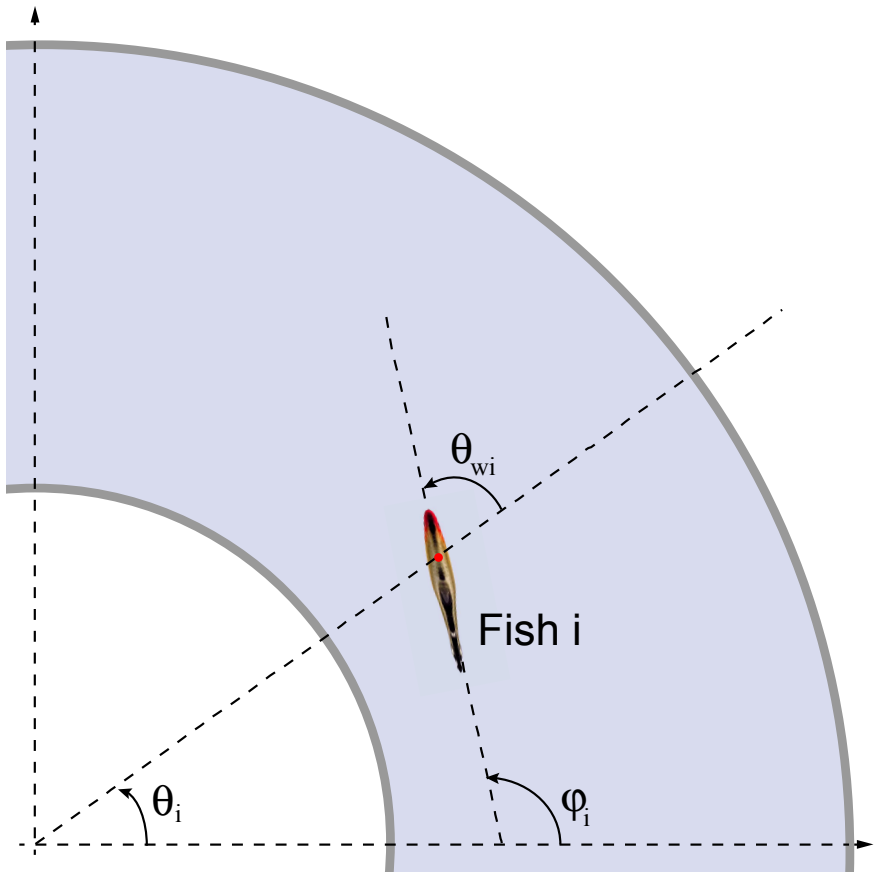


Figure 4.6: Variables used to describe the position, heading and relative orientation of fish relative to the wall in the experimental set-up: θ_i is the angle formed by the position vector of fish i and the horizontal line, φ_i is the heading of fish i, and θ_{wi} is the angle of incidence of fish i relative to the wall w.

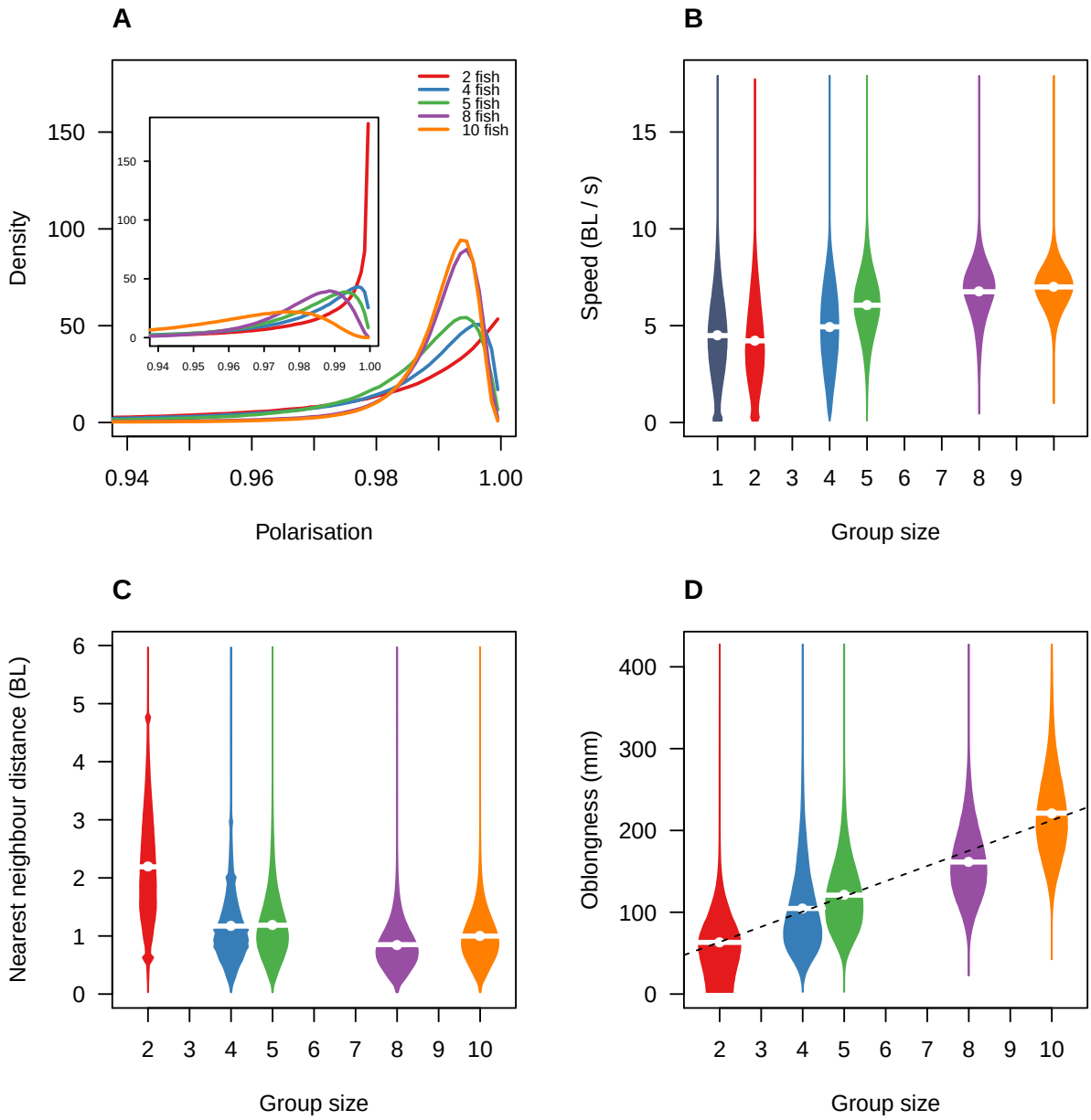


Figure 4.7: Influence of group size on internal structure, speed and shape of the schools. A) Distribution of polarisation, measured as the absolute value of the degree of alignment with the wall a_i and (inset) as the order parameter $\Phi = \sqrt{(\sum_N (\cos\phi_i)^2 + \sum_N (\sin\phi_i)^2)/N}$. Both parameters tends to 0 when the group is disordered and to 1 when the group is perfectly ordered. B) Distribution of the speed of the group, averaged over the speed of each individual, at each time, as a function of group size. C) Distribution of the nearest neighbour distance, measured on each individual, at each time, as a function of group size. D) Distribution of the oblongness of the group, measured on each frame as the maximum distance between positions of fish projected on the axis tangent to the swimming direction of the centre of mass of the group, as a function of group size. Dashed line stands for fitted linear model, $R^2 = 0.98$. B, C and D are violin plots, showing the rotated and mirrored histograms of the respective random variable. White belts stand for the mean.

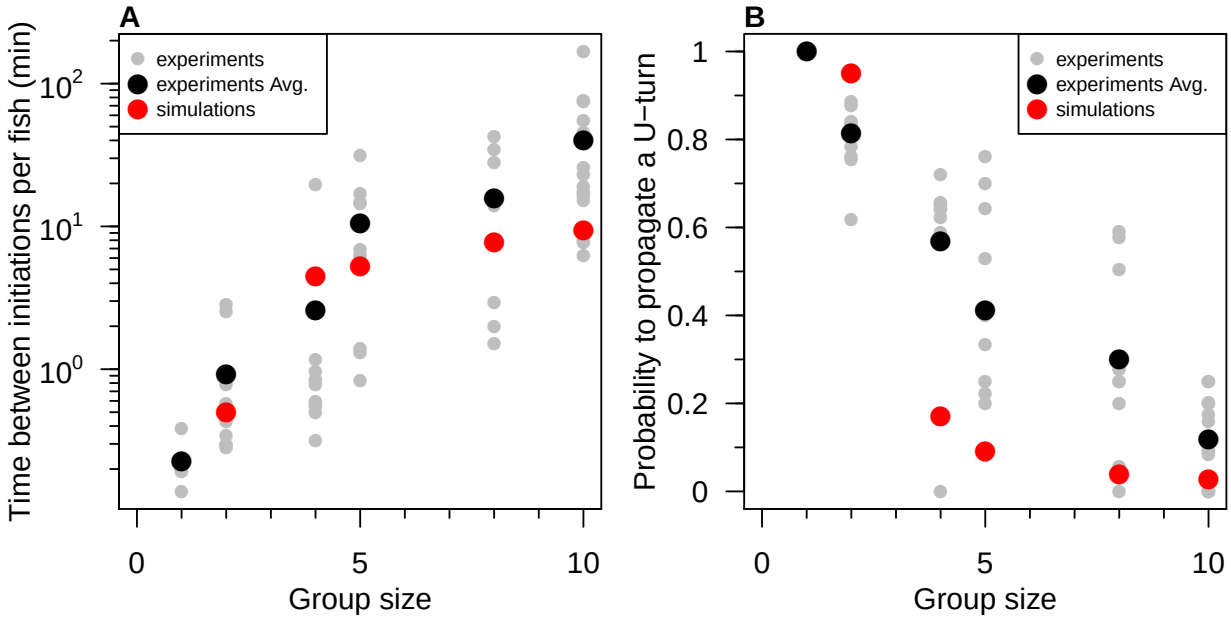


Figure 4.8: A). Time between U-turn initiation (failed or fully propagated) per fish as a function of group size. B). Probability that an initiated U-turn is fully propagated as a function of group size (see equations 4.8 and 4.9).

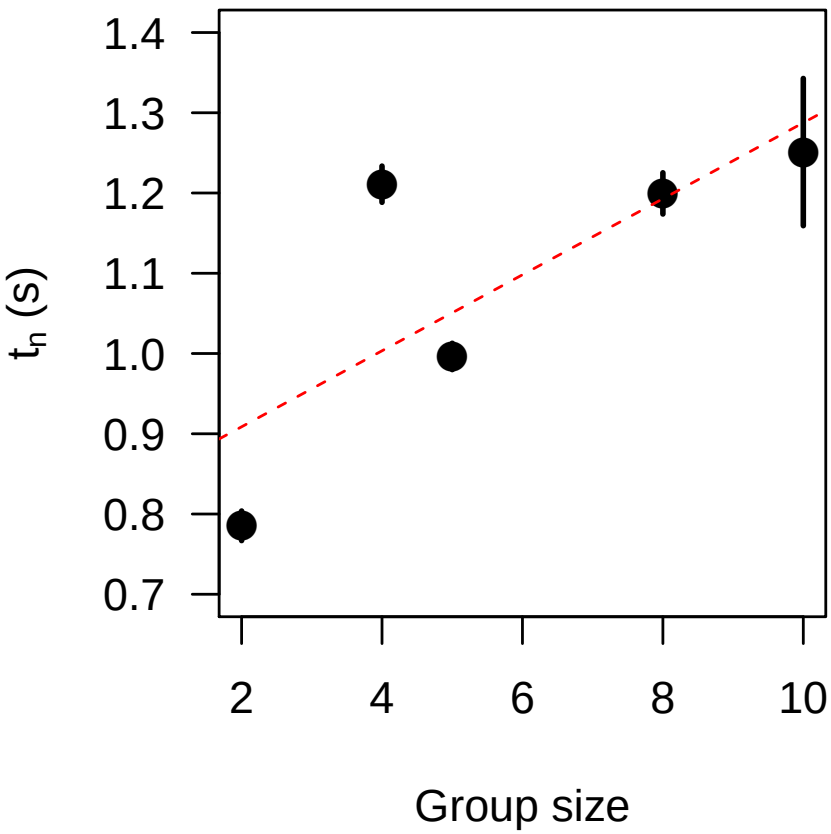


Figure 4.9: Mean ± sd of the bootstrapped sample of the scaling coefficient t_n used in Figures 1C and 1D; red dashed line stands for a fitted linear model, $R^2 = 0.59$.

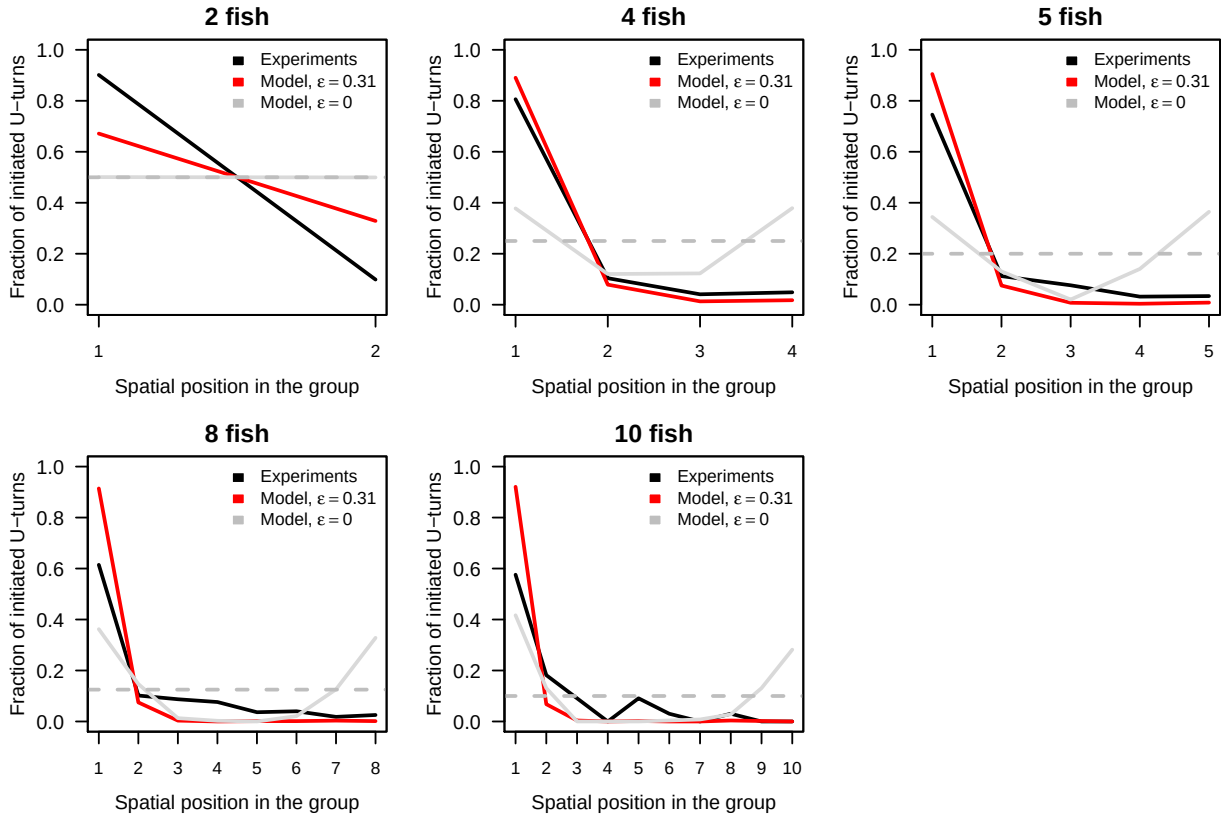


Figure 4.10: Spatial position of the U-turn initiator in groups of 2, 4, 5, 8 and 10 fish, in data of experiments and simulations.

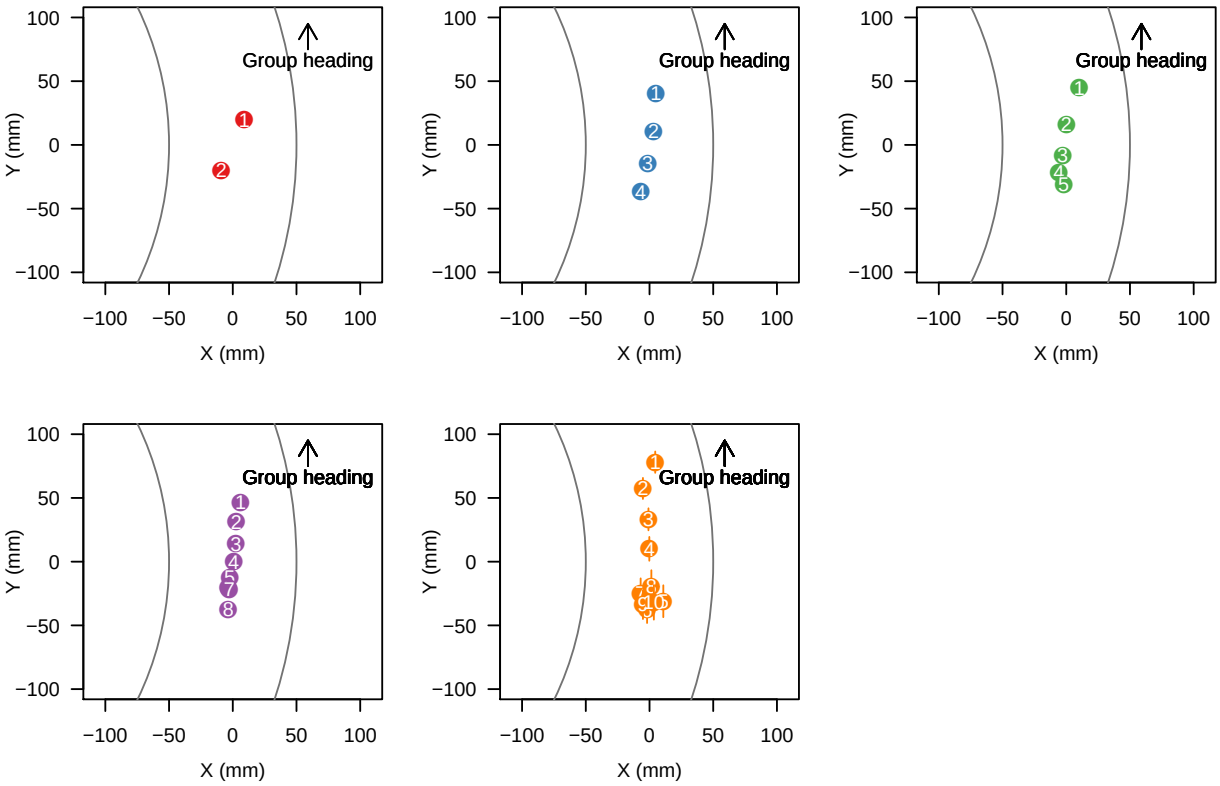


Figure 4.11: Average positions at U-turn initiation of individuals that turn subsequently, indicated by their ranks of turning (where rank 1 is the initiator of the U-turns) in experiments for groups of 2, 4, 5, 8 and 10 fish. Positions have been corrected so that all groups move in the same direction, with the outer wall at their right-hand side. Error bars indicate the standard error of the x and y-coordinates (smaller than the circles if not visible). The origin of the coordinate system is set to the centroid of the average positions of individuals. Statistical tests regarding the effect of the ranks of turning on the x and y-positions are reported in Tables 4.3 and 4.2.

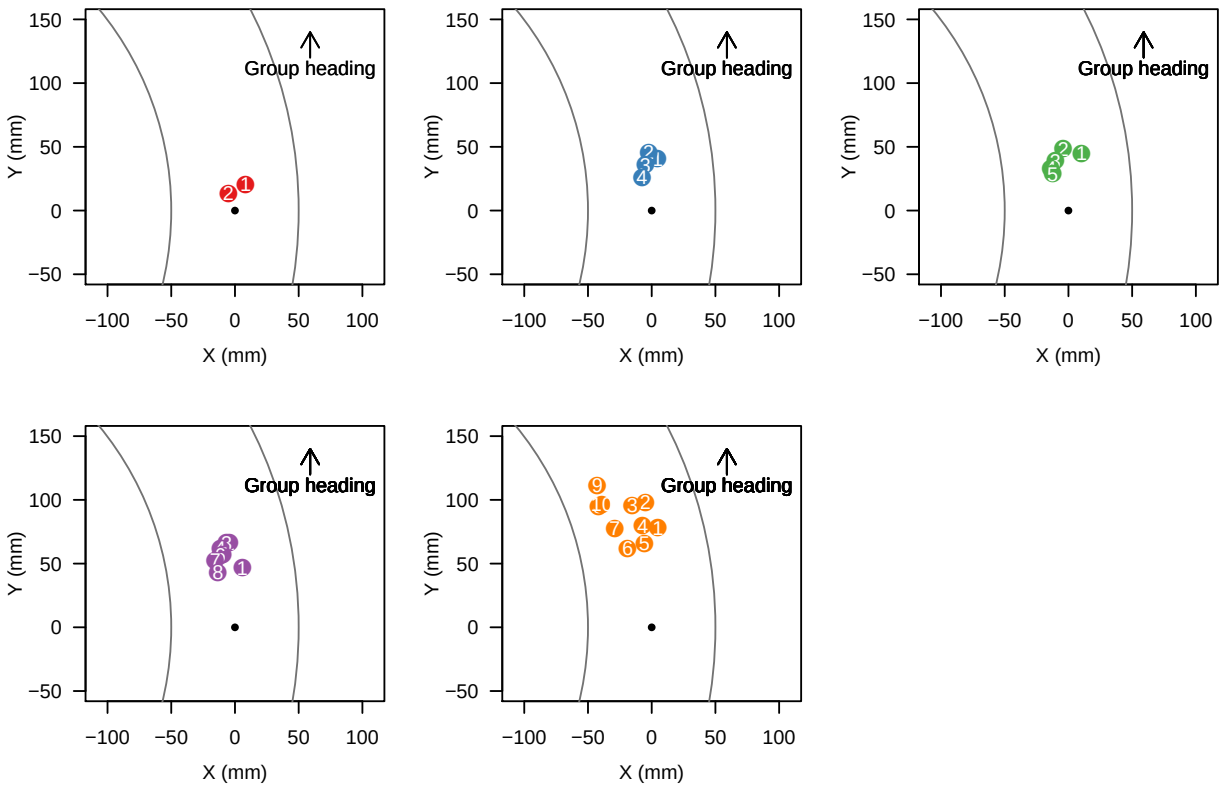


Figure 4.12: Average positions of individuals that turn subsequently, indicated by their ranks of turning (where rank 1 is the initiator of the U-turns) in experiments for groups of 2, 4, 5, 8 and 10 fish. Positions have been corrected so that all groups move in the same direction, with the outer wall at their right hand-side. Error bars indicate the standard error of the x and y-coordinates (smaller than the circles if not visible). The origin of the coordinate system (black dot) is set to the centroid of the average positions of individuals at the initiation of the collective U-turns. Statistical tests regarding the effect of the ranks of turning on the x and y-positions are reported in Tables 4.5 and 4.4.

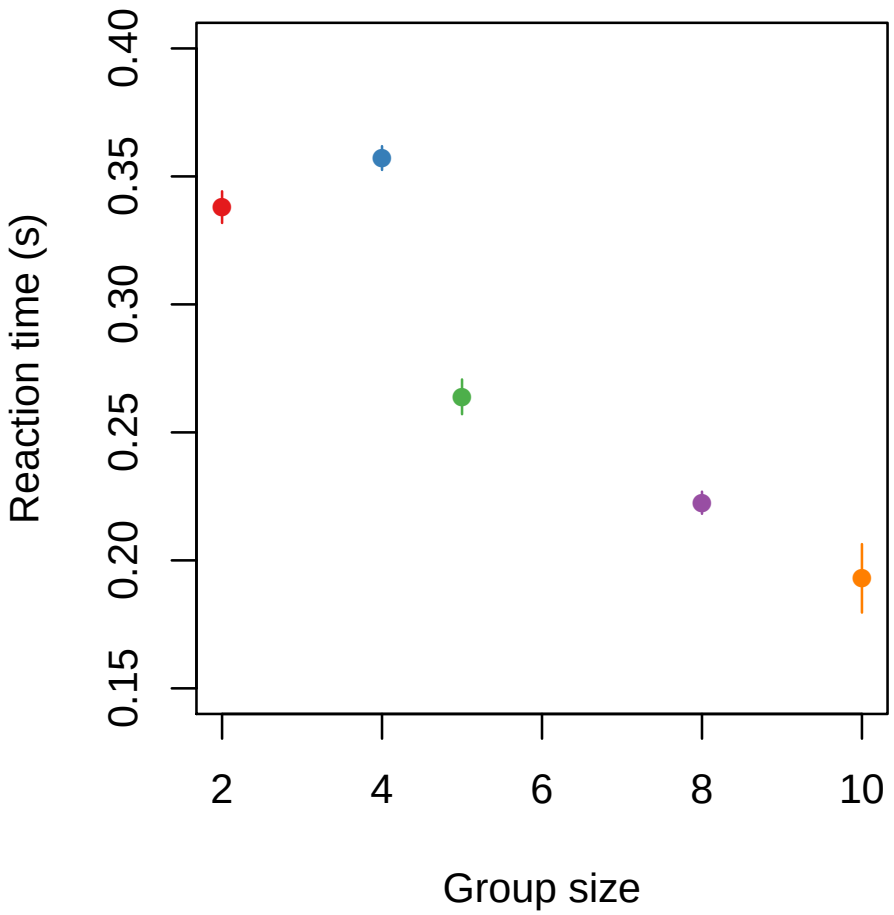


Figure 4.13: Reaction time measured as the average time interval between subsequent individuals making a U-turn (±s.e.) as a function of group size.

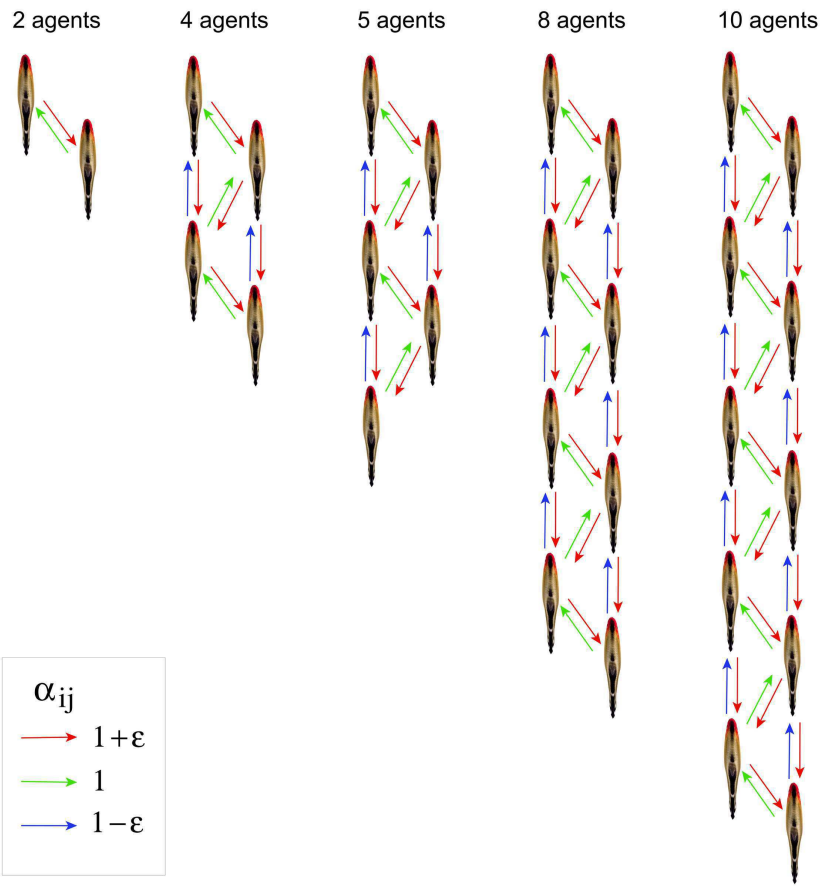


Figure 4.14: Topology of the interaction network in the simulations for different group sizes. Arrows indicate interactions going from the influencing agent to the influenced one. The colour of the arrow refers to the weight of the interaction, namely $\alpha_{ij} = 1 + \epsilon$ (red arrow), $\alpha_{ij} = 1$ (green arrow) and $\alpha_{ij} = 1 - \epsilon$ (blue arrow). The number of influencing neighbours of a focal agent can be derived from the number of pairs of arrows connected to the agent. For instance, in groups of 5 agents, each agent has, respectively (from front to back), 2, 3, 4, 3 and 2 influencing neighbours.

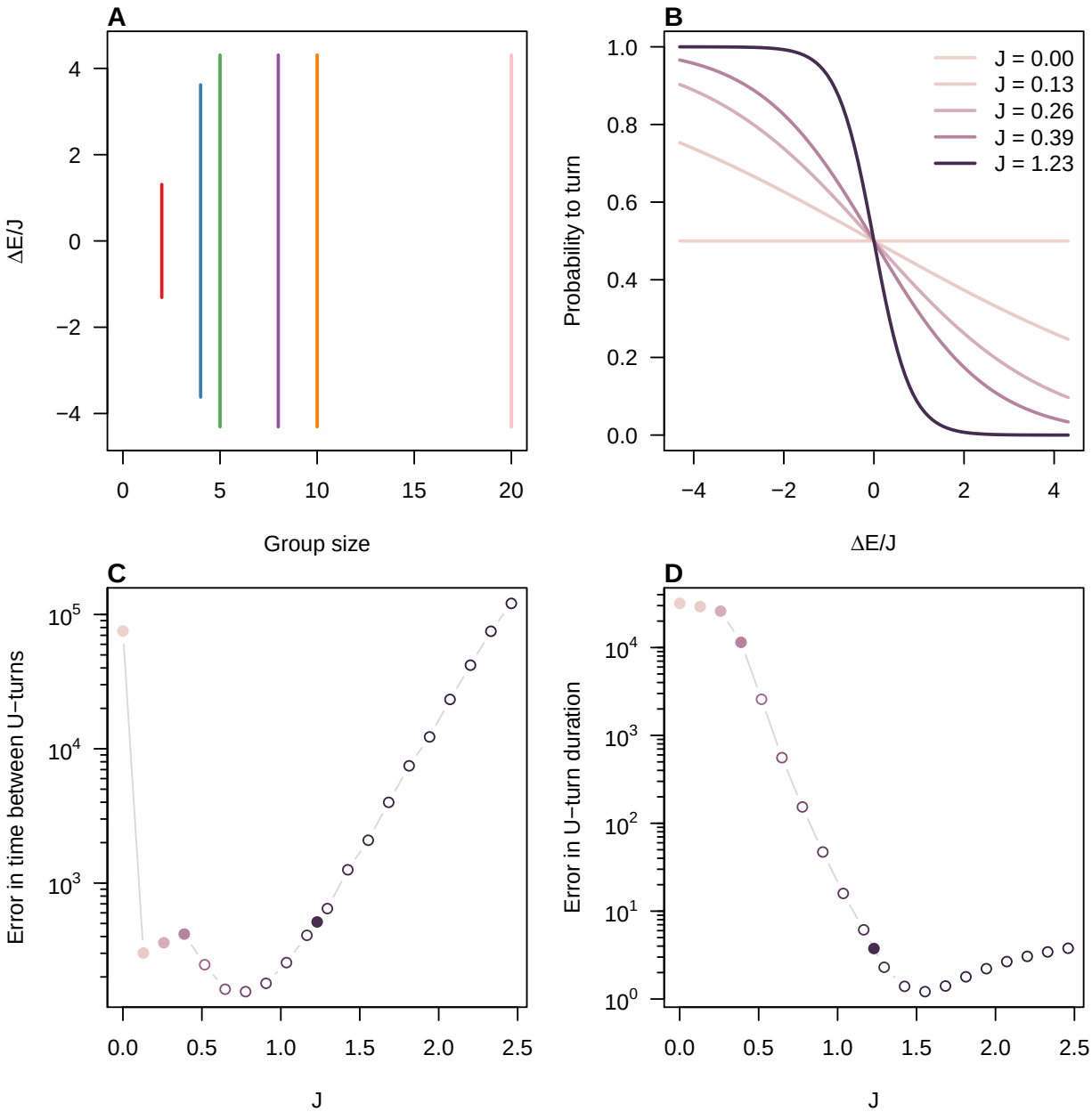


Figure 4.15: Influence of nonlinearity in the model, with the anisotropy parameter $\alpha = 0.31$. A) Range of $\Delta E/J$ values given $\alpha = 0.31$ for each group size. B) Probability P to accept an individual U-turn (see main text, equation 3.5) for different values of J (including $J = 1.23$, used in the fitted model). C) Error between simulations and experimental data regarding the average time between U-turns defined as $\Delta = \sum_n (\frac{\tau'_n}{\tau_n} - 1)^2$. D) Error between simulations and experimental data regarding the U-turn durations, defined as $\Delta' = \sum_n (\frac{t'_{e,n}}{t_{e,n}} - 1)^2$. In B, C and D, colours depend on the nonlinearity of the acceptance probability function (the darker, the more the response is nonlinear). Errors in C and D are computed from simulations with $T' = 10^6$ Monte-Carlo time steps. In C and D, filled dots stand for the values of J considered in B.

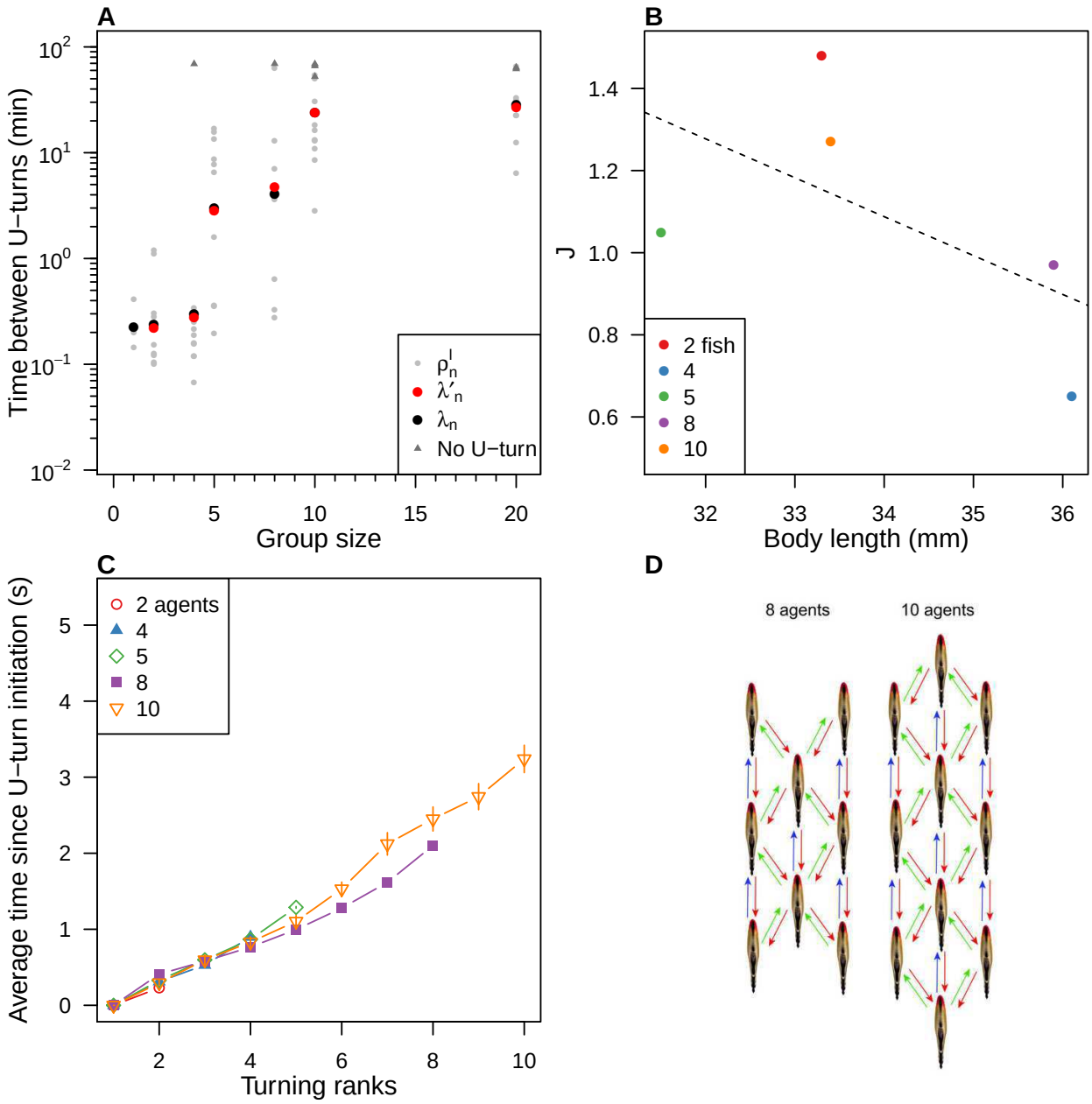


Figure 4.16: A). Average time between two consecutive collective U-turns as a function of group size with simulations performed with a value of the strength of interaction J per group size. All other parameters are set to the values of the main text. B). Values of J used in (A) as a function of the average body length of each group size. C). Average time interval since the beginning of a collective U-turn as a function of turning rank and group size in simulations, with same parameters as in main text. For 8 and 10 agents, the topologies of the interaction network have been changed to those shown in (D).

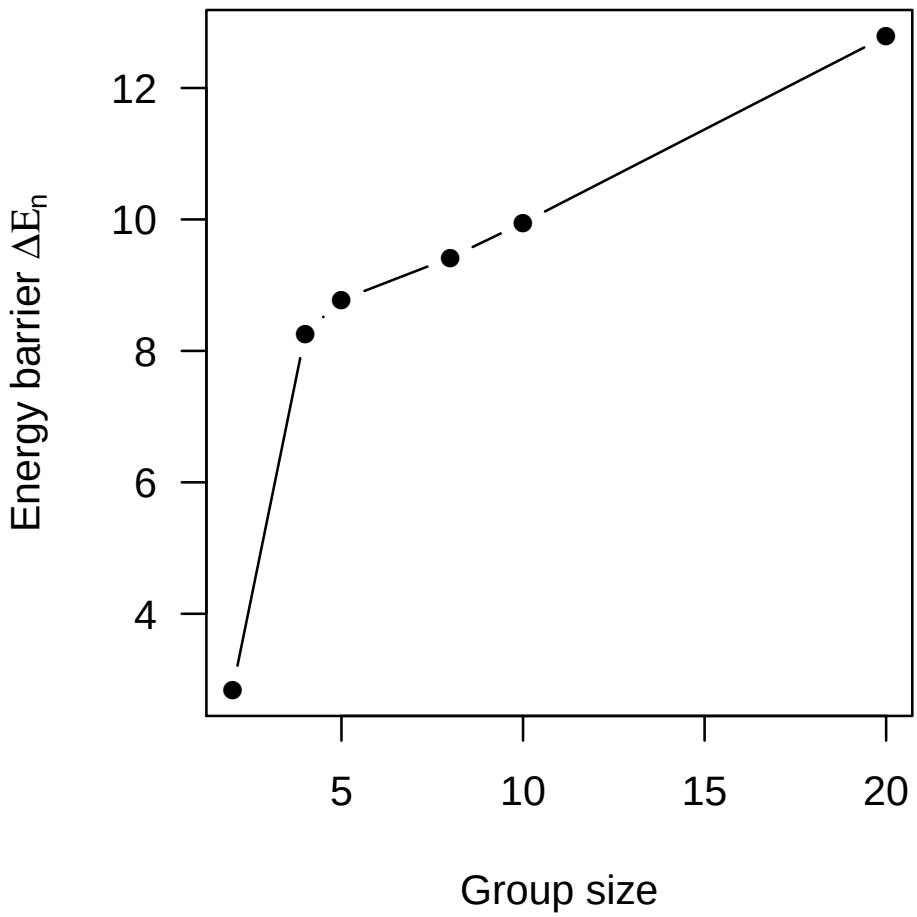


Figure 4.17: Energy barrier as a function of group size. The energy barrier ΔE_n for a group of size n is calculated in simulations as the difference between the average maximum of energy reached during U-turns and the reference energy when all agents are heading in the same direction.

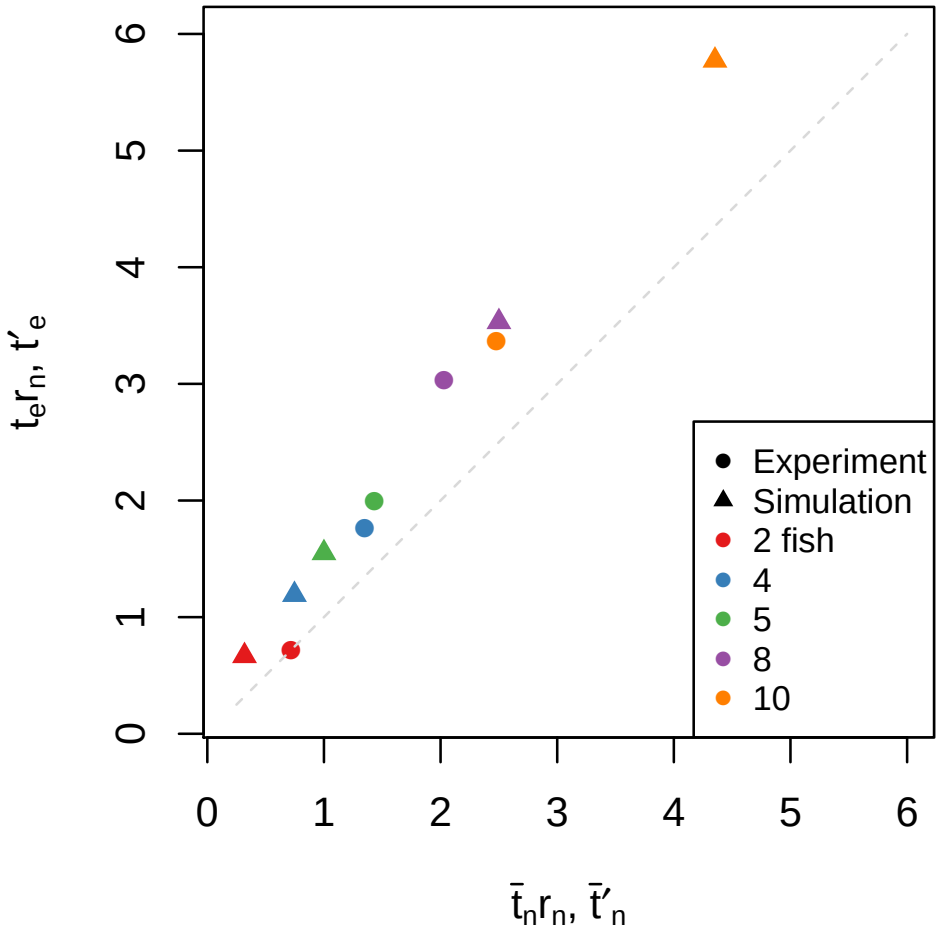


Figure 4.18: Duration of the collective U-turns, both in experiments $t_e r_n$ and the model t'_e as a function of the scaling coefficients ($\bar{t}_n r_n$ in experiments and $\bar{t}'_n$ in the model), computed by data scaling.

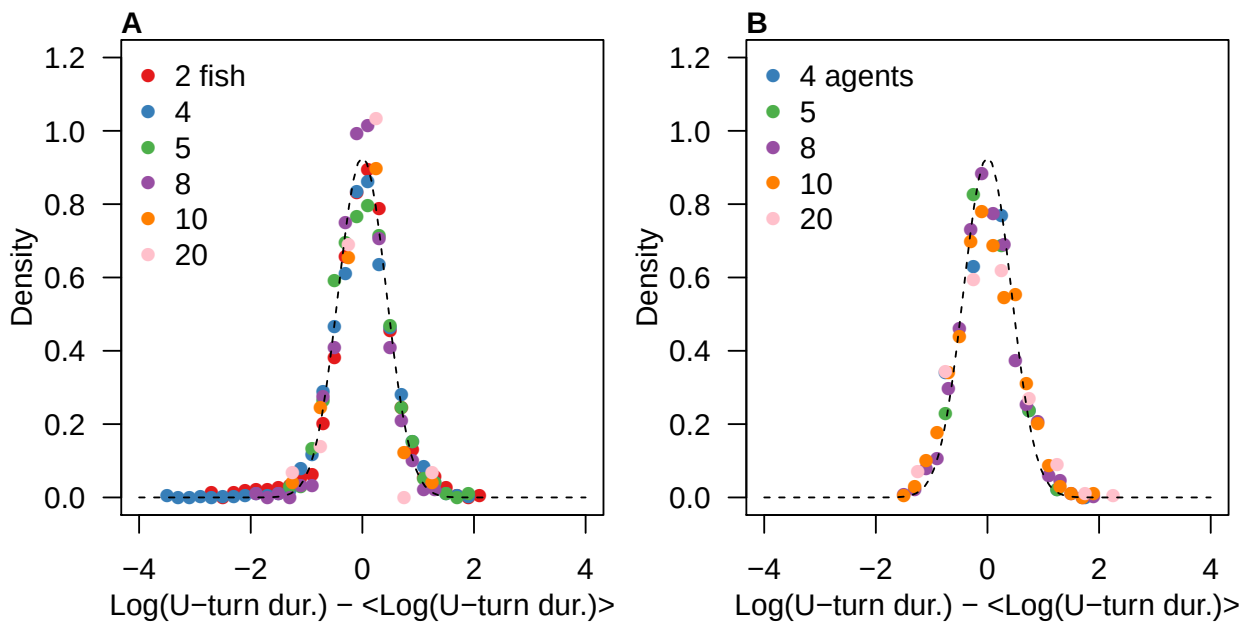


Figure 4.19: Probability distribution of the durations of U-turns normalised by the average duration, for each group size in experiments (A) and numerical simulations of the model (B). Dashed line is the probability density function of the Normal distribution $\mathcal{N}(\mu = 0, \sigma = 0.43)$.

Chapter 5

Conditioning an avoidance response in groups of rummy-nose tetra (*Hemigrammus rhodostomus*)

Valentin Lecheval, Patrick Arrufat, Stéphane Ferrere, Charlotte K. Hemelrijk and Guy Theraulaz

Abstract

We develop an experimental method to induce controlled perturbations in a group of fish and investigate the propagation of information. We use the paradigm of the shuttle box to condition aversive escape reactions in groups of rummy-nose tetra (*Hemigrammus rhodostomus*) in response to a green light. We show that aversive conditioning can (i) be used in this species, (ii) trigger collective escape reactions and (iii) be transferred from the training set-up to a new setting. We characterise and quantify the aversive conditioning and discuss long-term habituation and forgetting. We discuss these preliminary results in the context of propagation of information in reaction to external stimuli. Our findings suggest that the proportion of conditioned individuals in a group is critical to trigger collective escape reactions in response to external stimuli. Our conditioning experiments open promising possibilities for investigating the collective responses and propagation of information within groups of fish in response to perturbations mimicking sudden changes in the environment, such as predator attack.

Contribution of authors

V.L., C.K.H. and G.T. conceived and designed the study; P.A. and S.F. developed the set-ups. V.L. performed experiments; V.L. analysed data; V.L., C.K.H. and G.T. wrote the paper.

5.1 Introduction

Rummy-nose tetra (*Hemigrammus rhodostomus*) is a species of fish that swims in schools (i.e. the collective motion is highly coordinated) with high levels of polarisation (see previous chapter). Fish schools are of particular interest to investigate collective behaviour of animal societies in various ecological contexts, such as, for instance, under predator threat. When a school reacts collectively to an external perturbation, individuals react either directly to the perturbation itself or to startle response of its neighbours (Domenici and Batty, 1994). Therefore, the information responded to differs among group members. Conditioning experiments are of interest in the field of collective behaviour to investigate such phenomena experimentally. This can be done by manipulating the behaviour of a few individuals only by conditioning them to startle in reaction to a stimulus (Pillot et al., 2010, 2011; Miller et al., 2013; Toulet et al., 2015). Here, we use aversive conditioning. This implies that an initially neutral stimulus becomes aversive after repeated pairing with an unconditioned aversive stimulus. For animals in motion, the elicited (trained) escape response is of interest to discover how information propagates in fish schools when a single or a few individuals spot a predator in their neighbourhood. Unfortunately, all previous experiments of aversive conditioning conducted in fish were done with species that do not form fish schools (i.e. groups where individuals are highly coordinated and

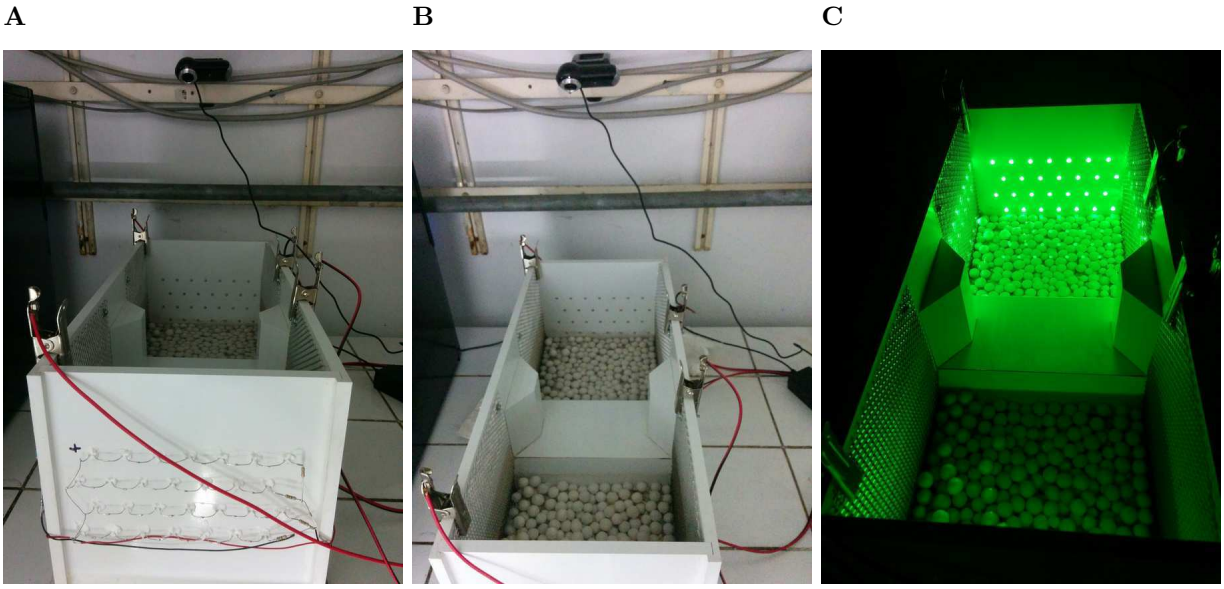


Figure 5.1: Pictures of the shuttle box ($46 \times 23 \times 21$ cm) used for the aversive conditioning. In (C), the conditioned stimulus is turned on in one of the two compartments.

aligned) (e.g. in zebrafish (*Danio rerio*) (Agetsuma et al., 2012) or goldfish (*Carassius auratus*) (Portavella et al., 2004)). In this paper, we developed experiments with *H. rhodostomus* to condition its escape response in order to investigate the propagation of information within a group of fish in a controlled way. We use aversive conditioning (electric shocks associated with changes in lighting) similar to the classical behavioural paradigm of the shuttle box, already tested in other fish species (Horner et al., 1961; Woodard and Bitterman, 1973; Pironi and Schmidt, 1988; Portavella et al., 2004; Pradel et al., 1999; Xu et al., 2007; Agetsuma et al., 2012). The aversive conditioning involves training fish to escape when a green light is turned on. We subsequently investigate how the proportion of conditioned fish in a group affects its propensity to perform a collective escape. We show in *H. rhodostomus* (i) that individuals learn to escape in response to a green light after aversive conditioning, (ii) that the conditioned escape can be transferred to a new experimental setting and (iii) that a minimum proportion of conditioned individuals in the group is needed to propagate the learnt behaviour.

5.2 Material and methods

5.2.1 Animals

A group of 70 rummy-nose tetras (*Hemigrammus rhodostomus*) were purchased from Amazonie Labège (`http://www.amazonie.com`) in Toulouse, France. The rummy-nose tetra is a tropical freshwater species. Fish were kept in 150 L aquariums on a 12:12 hour, dark:light photoperiod, at 27.5° C (±0.8° C) and were fed *ad libitum* with fish flakes.

5.2.2 Experiment 1: avoidance conditioning in a shuttle box

Subjects

Similar experiments (see following sections) with zebrafish have shown that groups of 5 fish learn faster than single individuals or groups of two (Gleason et al., 1977). For this reason and to avoid inactivity of solitary fish, we conducted the conditioning of 6 fish at the same time. 6 fish – thereafter *conditioned fish* – were randomly sampled from the breeding tank in July 2016 and kept in a different breeding tank, under the same conditions as the other fish. All other individuals are labelled as *naïve fish*.

Conditioning apparatus

Fish were trained using a shuttle box modified from Horner et al. (1961). The shuttle box ($46 \times 23 \times 21$ cm) is made of white plastic (polyvinyl chloride, PVC) (Figure 5.1A). The two compartments ($20 \times 15 \times 20$ cm) have 84 green light-emitting diodes (LEDs) set on 4 rows of 7 LEDs for each panel (Figures 5.1B-C). The LEDs of

each compartment can be turned on independently from those of the other compartment. Compartments are separated by a trapezoidal hurdle 7.5 cm high. Two metallic plates conduct electricity in each compartment. Compartments are filled with 4.5 cm of white gravels and white ceramic balls to enrich the environment of fish, in line with policies for animal welfare. The water comprised 50% of water purified by reverse osmosis and 50% of water treated by activated carbon, heated at 27° C. Its level is set at 3 cm above the hurdle (i.e. 10.5 cm in each compartment). Training is monitored by a webcam (Logitech QuickCam Pro 9000). Light and electric stimuli are both delivered manually.

Conditioning procedure

Before and after the training session, the group of 6 conditioned fish is let in the set-up without any stimulus for respectively 15 min (habituation) and 5 min. Water is changed after each experiment.

Acquisition

Training sessions were performed on a regular basis (77 sessions in total) from July 2016 to July 2017 with the same 6 fish [1]. Each training session consisted of 20 trials (intertrial time of 2 min). In each trial, at least one of the 6 fish had to cross the hurdle to avoid mild electric shocks (7 V, 2.7 mA, measured on the electrodes) (unconditioned stimulus, US) administered via electrodes 3 s[2] after onset of a light signal (green light, conditioned stimulus, CS). The light signal was always turned on in that compartment where the majority of the individuals were located. A trial is labelled as a Success if at least one fish crossed the hurdle before the onset of electric shocks (i.e. within 3 s after onset of the CS) and Failed in all other cases. If there is no fish crossing the hurdle within 15 s after the onset of electric shocks, all stimuli (CS + US) are turned off and the trial is labelled as Failed. The proportion of trials labelled as Success over the 20 trials will be referred to as the *proportion of escapes*.

10 training sessions without US have been conducted on groups of 6 naïve fish randomly sampled from the breeding tank to measure the proportion of escapes due to the spontaneous exploration of the apparatus (Control).

Our dataset consists in the output (Failed or Success) of 200 trials for naïve fish and 1540 trials for conditioned fish.

Ethical use of animals

To minimise the stress induced by the experiments with aversive conditioning, we carefully followed the Three Rs principle (Replacement, Reduction and Refinement) (Russel and Burch, 1959). The Three Rs are a guiding principle for ethical use of animal in testing, that recommends to Replace experiments that kill or harm animals with alternative techniques, to Reduce the number of animals used, and to Refine experiments to reduce suffering. We use aversive conditioning to Replace empirical experiments that would involve predators hunting their preys. Regarding Reduction, only 6 individuals have been used during the whole conditioning, conducted over almost a year. As for Refinement, the power of the electric shocks has been set to the *minimum* power (7 V, 2.7 mA, measured on the electrodes) triggering a reaction in fish visible by eye. Individuals were carefully observed every day for wounds or abnormal behaviour. Visible injuries are endpoints, i.e. an animal with wounds would be removed from the experiment. Experiments have been conducted in a different room than that of the breeding tanks. Thus, experiments were not visible by fish in the breeding tank, in agreement with policies for animal welfare.

All experiments comply with the European legislation for animal welfare. Our experiments have been approved by the Ethics Committee for Animal Experimentation of the Toulouse Research Federation in Biology (CEEA N°1).

5.2.3 Experiment 2: test in a new environment

Experimental set-up

The square experimental tank (120×120 cm) was made of glass and was set on top of a box to reduce vibrations. The room was heated at the same temperature as the water in the tank to avoid gradients of temperature. The tank was surrounded by four opaque white curtains and illuminated homogeneously by

[1] One of the 6 fish died, 10 months after the beginning of the experiment. The death did not occur during an experiment and no wounds were noticed. Experiments were thus performed with the other 5 fish thereafter.

[2] The CS-US interval had first been set to 5 s and then to 3 s after the 7 first experiments, to elicit faster and stronger avoidance reactions.

Table 5.1: Experimental conditions tested in Experiment 2 with the new environment. *1vs4* and *5vs0* refer to the number of conditioned vs naïve fish.

Condition	Replicates	Trials per replicate
Control	11	5
1vs4	9	5
5vs0	2	10

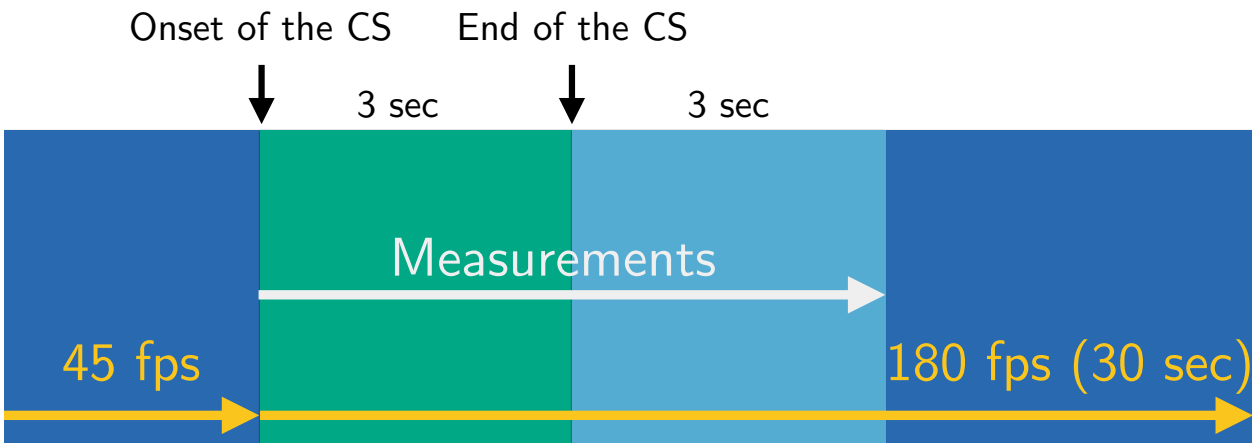


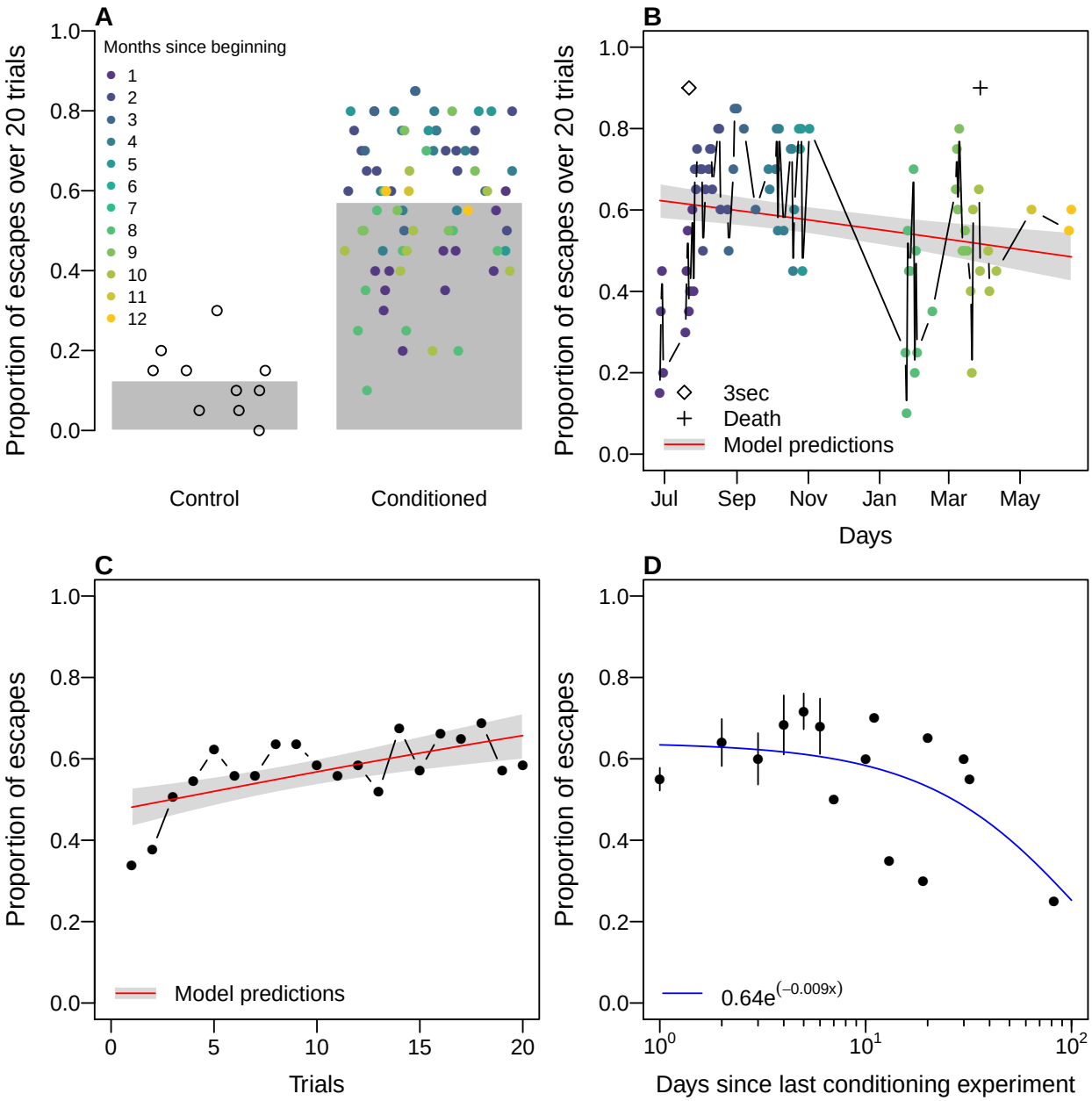
Figure 5.2: Schematic view of the trials of Experiment 2. Measurements are detailed in the text.

four LED light panels. Inside the experimental tank, a ring-shaped corridor 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 28.8° C (±0.7° C) . The corridor was 10 cm wide with a circular outer wall of radius 35 cm. 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. The outer wall has 2 rows of the same LEDs as used in the shuttle box, equally spaced by 1 cm. Only 35 cm of the entire diameter (thus 70 LEDs) can be turned on (CS).

Experimental procedure for testing

These experimental tests were conducted after the conditioning experiments presented in the previous section. Conditioned fish were trained for 10 trials of the conditioning experiment a few hours before an experiment was done in the new set-up. Per day, fish participated only in a single experiment. Experiments concerned groups of 5 fish and were done in the ring-shaped tank. Trajectories were recorded by a high-speed camera (R&D Vision) recording from above the set-up at 45 Hz or 180 Hz, both in high resolution (2000×2048 pixels). Groups swim spontaneously for 20 min, with trajectories recorded at 45 Hz, until the CS is turned on for the first time. At the onset of the CS, the camera automatically switches to record fish positions at 180 Hz. The CS is turned on for 3 s and trajectories are recorded subsequently for 30 s, still at 180 Hz. Several trials are performed on each group of fish, with an intertrial duration of 1 min. After the first trial, fish swim spontaneously for 1 min, recorded at 45 Hz before the CS. Three conditions are tested, the *Control* condition, with 5 naïve fish, a condition with 4 naïve fish and 1 conditioned (*1vs4*) and a condition with 5 conditioned fish (*0vs5*). Number of replicates and trials are reported in Table 5.1.

For each trial, we measure the number of fish that turn at least once in the direction opposite to the CS in the first 3 s following the onset of the CS (i.e while the light is on) and during the 3 s seconds after (i.e. when the light turns off). These two quantities estimate the strength of the collective reaction: if all individuals react fast and strongly to the CS, we expect them to turn while the light is on. We also measure whether all individuals are swimming in the direction opposite to the CS 6 s after its onset or not. If yes, we count one collective escape for the respective trial. This quantity is used to disentangle non-coordinated behaviours (e.g. where all individuals make U-turns several times and eventually the group does not swim away from the CS) to collective escape behaviours (i.e all group members make one clear U-turn and the group swims away from the CS). In the condition with one conditioned fish and 4 naïve ones (*1vs4*), we monitor if the conditioned fish turned first. The conditioned fish was recognised by eye by the operator. No tagging was used, because of the difficulty to tag small fish without altering their swimming behaviour.



We removed trials where not all fish were aligned in the same direction just before the CS is turned on (4 trials out of 55 in the condition *Control*, 8 trials out of 45 in the condition *1vs4* and 0 out of 20 in the condition *5vs0*).

5.3 Results

5.3.1 Experiment 1

We use logistic regressions (with R (R Core Team, 2016)) to investigate the effect of the conditioning experiments on the escape response of fish. We model the probability p that the binary result r of each trial (Failed ($r = 0$) or Success ($r = 1$)) is 1 as a function of the index of the trial i, the number of days d since the beginning of the conditioning for the considered group and of c that indicates whether the group is a control ($c = 0$) or a conditioned one ($c = 1$). We start by evaluating the full model:

$$\text{logit}(p) = \beta_0 + \beta_1 i + \beta_2 d + \beta_3 c + \beta_4 id + \beta_5 ic + \beta_6 cd, \tag{5.1}$$

where $\beta_j, j \in [0, 6]$ are the regression coefficients. We search the model that minimises the Akaike information criterion (AIC) from the full model ($\text{AIC}_{\text{full}} = 2231$). The AIC addressed the trade-off between

Figure 5.3: Results of the conditioning experiments (page 93). A). Proportion of escapes in control and conditioned conditions. Open circles and dots show the proportion of escapes of each experiment. The colour of the dots stand for the time since the beginning of the experiments. B). Time series of the proportion of escapes in 20 trials. An empty diamond stands for a change of protocol, from a duration of 5 s of CS to 3 s of CS before onset of electric shocks (to trigger stronger aversive responses). The cross indicates the death of one of the 6 individuals. Subsequent training sessions were conducted with the remaining 5 individuals. Colour of the dots are the same as in (A). The first point of the time series is averaged over the first conditioning replicate and all control replicates. Red line and grey shade represent the model (Eq. 5.2) predictions and its 95% confidence interval. C). Performance of the group averaged over all training sessions as a function of the trial. Red line and grey shade represent the model (Eq. 5.2) predictions and its 95% confidence interval. D). Proportion of escapes against time since last training session, with exponential fit. Points stand for the proportion of escapes averaged over all experiments with the same amount of days since the previous conditioning experiments and vertical bars stand for the standard error. There is no bar when there is only one data point.

Table 5.2: Exponential of the estimated logistic regression coefficients (odd ratios and confidence intervals) of the model shown in Equation 5.2. We also report the p-values of the Wald statistic that tests for each β_j the null hypothesis $\beta_j = 0$ (no significant effect of the jth variable). Hierarchical stepwise likelihood ratio tests give the same significances. For β_1 to β_3, values greater than 1 indicate positive correlation while values less than 1 indicate negative correlation between p and the respective explanatory variable.

	Estimate	2.5%	97.5%	p-value
$\exp(\hat{\beta}_0)$	0.073	0.044	0.116	< 0.001
$\exp(\hat{\beta}_1)$	1.062	1.040	1.084	< 0.001
$\exp(\hat{\beta}_2)$	12.281	7.918	19.802	< 0.001
$\exp(\hat{\beta}_3)$	0.999	0.9998	0.9999	< 0.001

Table 5.3: P-values of the likelihood ratio tests performed to assess the significance of the effects of the condition and of the trial to predict the occurrence of collective escapes from binomial generalised linear mixed effect models.

	Only Condition	Only Trial	Null Model
Full Model	0.4	0.03*	0.03*
Only Condition			0.02*
Only Trial			0.2

the goodness of fit of the model and the number $j-1$ of explanatory variables. We find that

$$\text{logit}(p) = \beta_0 + \beta_1 i + \beta_2 c + \beta_3 id, \tag{5.2}$$

(hereafter called *the model*) minimises the AIC ($\text{AIC}_{\text{mod}} = 2228$). We report the exponential of the estimated regression coefficients (odd ratios) in Table 5.2. This model predicts correctly the output of 63% of all trials and all coefficients $\beta_j, j \in [0,3]$ are significantly different from 0 (Table 5.2).

Our results show that *H. rhodostomus* can learn the task of the shuttle box, in a social context. The proportion of escape responses (avoidance of the CS) is statistically larger in the conditioned group (0.57 ± 0.02, mean±standard error) than in the control groups (0.125 ± 0.03) (Figure 5.3A and $\exp(\hat{\beta}_3) = 12.281$, Table 5.2). Control groups reach a non-null proportion of escapes because they randomly swim in the shuttle box and may cross the barrier during the CS, either by accident or because fish are afraid of the CS. We qualitatively assess an effect of the CS on the behaviour of the naïve fish which is different from the behaviour of conditioned fish. In general, groups of naïve fish become excited during CS but without any avoidance behaviour – they are even sometimes attracted towards the LEDs that emit the light. The time series of the proportion of successful responses of the conditioned group shows that there is a decrease in reaction to the stimulus (Figure 5.3B with the model fit and $\exp(\hat{\beta}_3) = 0.999$). The learning of the conditioning occurs within a month (purple dots) and the best performances are achieved until November (blue dots). From January to July (green and yellow dots), despite 27 conditioning experiments, it was not possible to reach the performances obtained in the other months. It seems that fish experience long-term habituation, even though our protocol cannot properly test it (i.e. we should test the specificity of habituation to the stimulus to exclude a fatigue effect). Although fish react less to the CS in the last sections of the time series, the achieved performance is, on average, still higher than the controls. Within a training session, fish perform significantly better after several trials, typically 3 (Figure 5.3C and $\exp(\hat{\beta}_1) = 1.062$). This effect decreases with the number of days since the beginning of the experiments, possibly because of the long-term habituation previously discussed (interaction term $\beta_3 id$ of the model, $\exp(\hat{\beta}_3) = 0.999$).

In the models presented in Equations 5.1 and 5.2, we did not use the number of days since the previous experiment as an explanatory variable, because the value for the control experiments and the first experiment of the conditioning experiments would be arbitrary (i.e. infinity, encoded for instance as 1000 days). Such an arbitrary value would have influenced the significance of the statistical tests performed. We find that there is an effect of the number of days since the previous experiment on the performance over the 20 trials of a focal experiment (Figure 5.3D). Namely, we find a decay of the proportion of escapes that can be modelled as

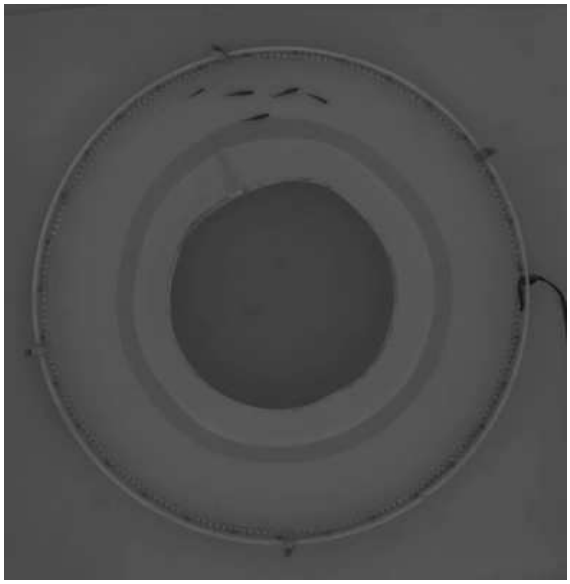
$$p = a \exp(bt), \tag{5.3}$$

where a is the degree of learning, $-b$ is the inverse of the rate of forgetting (in days) and t the time since the previous conditioning experiments (in days). We find $a = 0.64$ and $b = -0.009$. The value for a corresponds to the performance of the group at the beginning of our protocol (blue dots in Figure 5.3B). It is worth to note that we do not find any significant correlation between the number of days since the beginning of the conditioning d and the number of days since the previous conditioning t (p-value $= 0.25$), excluding a confounding effect of d and t on the proportion of escapes p. In other words, it seems that our results indicate both habituation and forgetting.

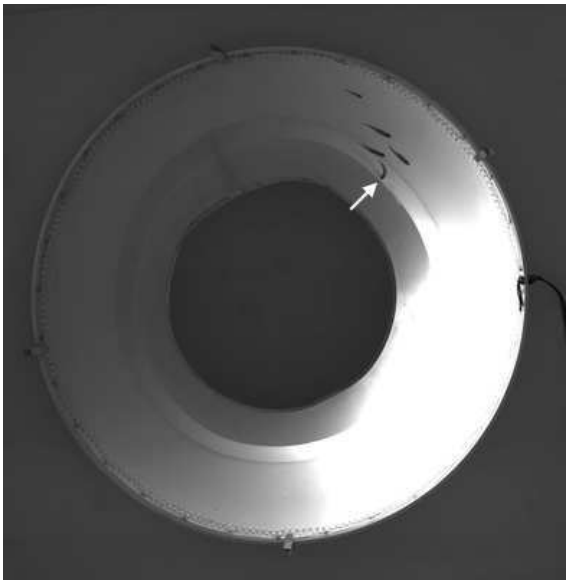
5.3.2 Experiment 2

Figure 5.4 shows a collective escape away from the CS performed by the group of conditioned fish: before the onset of the CS, the group is polarised (Figure 5.4A); at the onset of the CS, one individual reacts instantaneously, performing a U-turn (Figure 5.4B) which propagates to the other group members (Figure

A $t = -1\,\mathrm{s}$

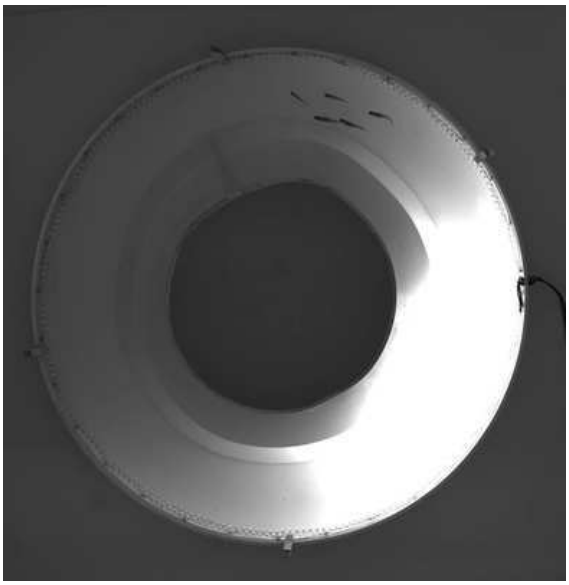


B $t = 0\,\mathrm{s}$



C $t = 1\,\mathrm{s}$

D $t = 2\,\mathrm{s}$



E $t = 3\,\mathrm{s}$

F $t = 4\,\mathrm{s}$

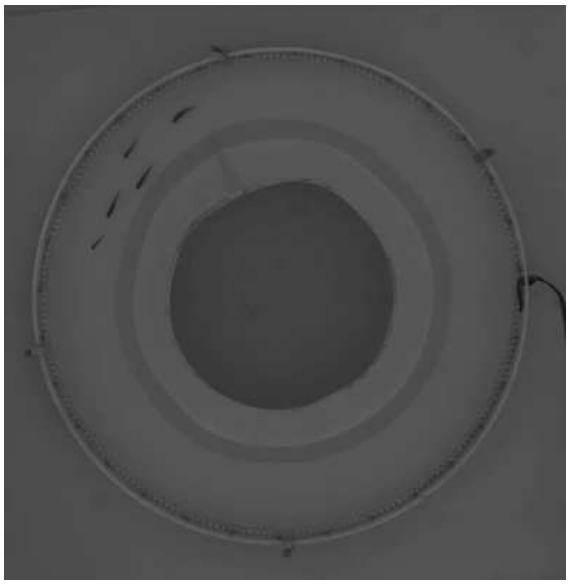


Figure 5.4: Snapshots from the escape behaviour of a group with 5 conditioned fish (page 96). Snapshots are taken every second from 1 second before the onset of the CS to 4 seconds after. The white arrow on (B) shows the fish that responds first to the CS.

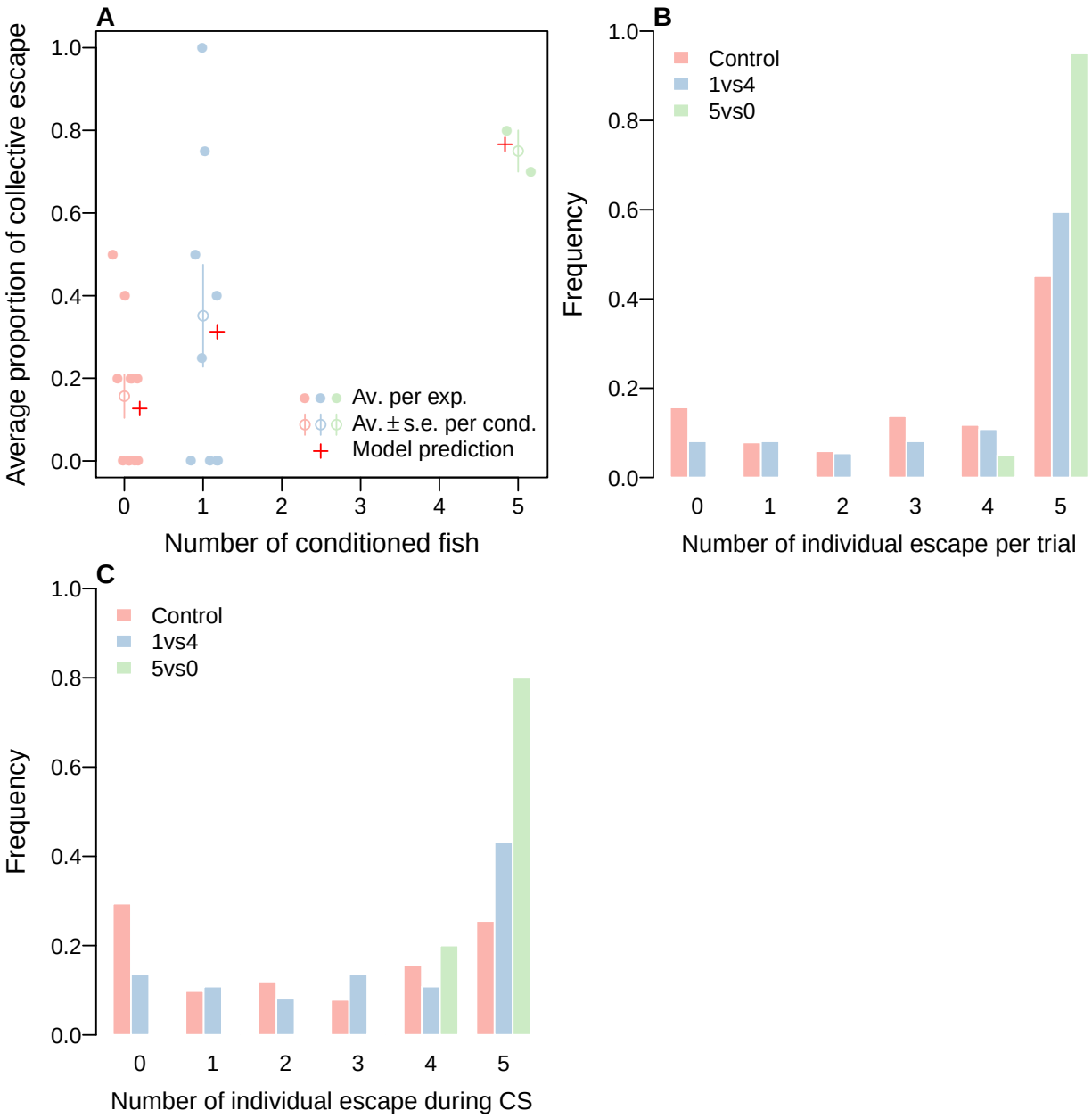


Figure 5.5: The conditioning in the new set-up. A). The average proportion of collective escapes as a function of the number of conditioned fish in the group. We report results averaged over replicates (dots) and conditions (empty circles) as well as the model predictions. The model used for reference is the binomial generalised linear mixed model with only the condition as fixed effect. Noise is added to the x-coordinates of the points, to improve visualisation. B). Bar chart of the number of individual escapes per trial for each condition. C). Bar chart of the number of individual escapes per trial that occur during the CS.

5.4C) and the group swims in the opposite direction in less than 2 s after the light is turned on (Figure 5.4D-F).

We used R (R Core Team, 2016) and the packages `lme4` (Bates et al., 2015) and `lsmeans` (Lenth, 2016) to perform a binomial generalised linear mixed effects analysis of the occurrence of collective escapes in trials as a function of the condition of the replicate (*control*, 1 conditioned versus 4 naïve (*1vs4*) or *0vs5*) and the index of the trial (1 to 10). Namely, we test (i) an effect of the condition, to investigate whether the proportion of conditioned individuals in the group may change the propensity to initiate a collective escape and (ii) an effect of habituation or fatigue to the CS. As random effect, we use intercept for the experiment as well as by-experiment random slopes to account for the non-independence of the responses within a condition, since there are several trials per tested group of fish. P-values were obtained by likelihood ratio tests of the full model with the fixed effects against models with only one of each of the fixed effect. We also perform the likelihood ratio tests of the models with only one of each of the fixed effects against the null model with intercept and random effect only.

Both tests show a significant effect of the condition (*control*, *1vs4* or *0vs5*) on the occurrence of collective escapes and no significant effect of the trial (Table 5.3). We find that the CS is not neutral on the behaviour of the naïve groups. naïve fish perform collective escapes in 16% of all trials (Figure 5.5A). In 71% of all trials there are at least 3 fish that make individual U-turns (Figure 5.5B). However, when all group members are conditioned, groups perform collective escapes in 75% of all trials (Figure 5.5A). Here, the response is in general collective and fast: in 80% of all trials all group members have turned within the three seconds after the CS is turned on (Figure 5.5C). In short, although groups of naïve fish react to the CS, with fluctuations in the heading of individuals they react more slowly and not collectively, in contrast to the groups composed only of conditioned fish.

The occurrence of collective escapes when only one fish is conditioned (condition *1vs4*) does not significantly differ from the occurrence of collective escapes in pure groups of naïve fish (pairwise Tukey method, p-value = 0.3). However, fish in this condition tend to react more and faster than naïve fish (Figures 5.5B and C). We find that the conditioned fish turns in 78% of all trials, turns first in 43% of all trials (always during the CS) and initiates 54% of the collective escapes performed in the *1vs4* condition.

5.4 Discussion

Our experiments are, to our knowledge, the first conditioning experiments ever tried with *H. rhodostomus*. We show that it is possible to train *H. rhodostomus* to escape using an active aversive paradigm. We find that this conditioned behaviour can be successively used in a new environment. The possibility to condition *H. rhodostomus* opens opportunities to investigate collective behaviour with obligate schoolers (in contrast to experiments with zebrafish or goldfish).

The duration of the CS before the onset of electric shocks (3 s) is very short compared to previous studies (7.5 s (Agetsuma et al., 2012), 10 s (Horner et al., 1961; Woodard and Bitterman, 1973; Portavella et al., 2004), 12 s (Pradel et al., 1999; Xu et al., 2007), or 20 s (Piront and Schmidt, 1988, although with a different protocol)). This has to be taken into account to assess the conditioning success of *H. rhodostomus* (which was on average close to 63% during the first half of the year) compared to the other studies, achieving learning criterion of 70% (Portavella et al., 2004) or 80% (Pradel et al., 1999; Piront and Schmidt, 1988). This choice was made to elicit stronger (i.e. fast) aversive reactions in response to the CS.

It seems that our training leads to habituation because the performance of the conditioned group decreased after the first 48 experiments conducted over a period of 4 months. Even if our protocol was not designed to test habituation, our results suggest that this decrease in the performances of the conditioned groups was not the consequence of a lapse in fish memory: fish kept on performing fast aversive responses, although less often during the later conditioning experiments.

To explain the decay of the conditioned escapes as a function of the time since the previous training session, we make the assumption that the proportion of escapes during a replicate p can be modelled as $p = a\exp(bt)$, with a the degree of learning, $-b$ the inverse of the rate of forgetting (in days) and t the time since the previous conditioning experiments. Such a curve is similar to curves referred to as *forgetting curves*, after the seminal work of Herman Ebbinghaus in experimental psychology (Ebbinghaus, 1913, Chapter VII). Interestingly, we find that the exponential fit in our experiments is close to the fits proposed for forgetting curves in humans, modelled as combinations of exponential functions or power laws (Wickelgren, 1974; Rubin et al., 1999; Murre and Dros, 2015). Although our results concern a group of fish in contrast to an individual, it is likely that the collective performance is correlated with the level of learning of individuals. Future research with an experimental protocol dedicated to investigate the forgetting of conditioned behaviour could focus on the dependence of b to the number of times the group has experienced the conditioning m, in order to account for an overlearning effect (we predict that b decreases with m) as well as the dependence

of a to m, to account for long-term habituation (we also predict that a decreases with m, in agreement with the present findings). Such research may be of interest to understand the macroscopic scale of memory in *H. rhodostomus* and its underlying mechanisms.

As for the set of experiments that concerns the new environment, a ring-shaped tank and the propagation of the induced perturbation, we found that the occurrence of fast, collective escape behaviours increases when the proportion of conditioned fish in the group increases. Our results might suggest that there is a critical proportion of trained fish required to elicit a collective escape but this has to be confirmed. Future work will concern the development of a tracking software that would take the change of luminosity into account to keep track of fish positions during the emission of the light, which is badly done by current tools. Given the time resolution of our videos (180 fps after the onset of the CS), data with the positions of fish will help to compare qualitatively and quantitatively the collective behaviours and the propagation of information between the different conditions tested in our experiments. The spatio-temporal propagation of information after the CS will be analysed, using the domino-like propagation of information that occurs during spontaneous collective U-turns as benchmark (see previous chapter). Further experiments with more conditions (i.e. with all possible ratios of conditioned versus naïve fish in a given group size) will help to elucidate the question of the critical proportion.

Our conclusions remain uncertain because of the small number of conditioned fish in our experiments. We made the choice to train only a few fish for ethical reasons, given that it was the first time conditioning experiments were tried with this species. The results of our experiments being promising, we emphasise that our results need to be reproduced experimentally with a greater number of conditioned groups (e.g. 10 groups of 5 conditioned fish).

Chapter 6

General discussion

In this thesis, I have investigated the individual-level interactions and behaviours that underlie the coordination of collective motion and the propagation of information in groups of tropical fish *Hemigrammus rhodostomus*. Here, I summarise and discuss the main results and give an outlook regarding future work.

6.1 Overview of the main results

I discuss the results of this thesis in the light of four topics: 1) the burst-and-coast swimming style, 2) the individual-level interactions and the behavioural rules involved in the coordination of collective swimming, 3) the decision-making processes and the propagation of information within groups and 4) the methodological contributions.

6.1.1 The burst-and-coast swimming behaviour

Burst-and-coast swimming is an intermittent swimming style, adopted by the rummy-nose tetra, *Hemigrammus rhodostomus* (Lopez, 2015). It consists of cyclic bursts of swimming movements followed by a coast phase in which the body of the fish does not move. The quality of the recordings of our experiments (HDTV resolution, 50 Hz) as well as the duration (one to three hours) and the two shapes of our tanks (circular or ring-shaped) have favoured a detailed analysis of this swimming style. Swimming speed of fish in groups of 10 in a ring-shaped tank was on average 7 BL/s, and bursts of speed up to 10 BL/s could be reached in all group sizes. These traits may underlie the efficiency of the burst-and-coast style, as found previously (Weihs, 1974; Videler and Weihs, 1982): fast speeds of 10 BL/s can be reached and fish can sustain such speeds for more than an hour. We have shown that the control of the speed resulted from a subtle interplay between the respective duration of the burst and coast phase and the adjustment of the acceleration, in response to non-social (i.e. the wall) and social information. We developed a model in which fish make decisions (about their direction, speed and duration of the burst-and-coast cycle) only at the onset of the burst phases and in which the speed of fish during coast phases is a viscous drag accounting for water frictions. The model reproduced qualitatively and quantitatively the empirical spatial distribution of a single fish observed in a circular arena (see Appendix B for a discussion on the asynchronous integration of information by *H. rhodostomus*).

This discrete integration of information has also been suggested in a model for another species that has a burst-and-coast behaviour, the zebrafish (*Danio rerio*, Harpaz et al., 2017). Harpaz et al. have found that the burst phases are best predicted by fish that react to social information and that coast phases are best predicted by a passive mode where only drag forces are applied to fish velocity. These results are similar to ours, although obtained with a different method. These findings suggest a close connection between the swimming style of fish and the mechanisms (e.g. the integration of information) that control the coordination of schools.

6.1.2 Individual-level interactions and behavioural rules involved in the coordination of collective swimming

Following the methodological framework described by Gautrais et al. (2012), we used a bottom-up approach to disentangle and measure the non-social (walls of the experimental tank) and social interactions involved in the behaviour of *H. rhodostomus*. A data-driven model has been developed to assess quantitatively the relevance of the measured interactions.

The interaction of fish with the wall consists of avoidance behaviour, of which the intensity is modulated by the perception of fish. We showed that at each decision of the fish regarding its motion, all non-social information available in a circular tank could be summed up by only one point of the wall. In other words, by making decisions only based on the location of the wall closest to the fish (with respect to the radial distance) suffices to reproduce trajectories of one fish swimming spontaneously in circular tanks of three different radii. Although trajectories of fish are concentrated along the walls, interactions with the wall are repulsive: fish turn away from the wall when they are at less than 2 BL from it. The avoidance effect is also weaker when fish are parallel to the wall or when the wall is behind them. Thus, the perception of the wall by a fish depends on (i) its distance to the wall and (ii) its angle to the wall. In particular, how the interaction of avoidance of the wall depends on the angle of the fish to the wall reveals the anisotropic perception of fish, responsible for asymmetric interactions, also found for social stimuli.

We have shown that schools of *H. rhodostomus* were highly cohesive and polarised. We developed a theoretical model to account for the fact that the occurrence of collective U-turns in a ring-shaped tank decreases with increasing group size. The model shows that local social conformity (i.e. the tendency of fish to copy the behaviour of the majority of the neighbours) coupled with anisotropic perception of fish may explain the global properties of the collective U-turns. The social interactions that we measured underlie (at least to some extant) the social conformity implemented in this theoretical model. In particular, we found evidence of repulsion between fish, alignment and attraction. In contrast to previous phenomenological models, these behavioural modes are not set to discrete and arbitrary zones of distances to the neighbouring fish (Aoki, 1982; Huth and Wissel, 1992; Couzin et al., 2002). Instead, there is a continuous combination of attraction and alignment as a function of the distance between fish, as former theoretical studies assumed (Reuter and Breckling, 1994; Kunz and Hemelrijk, 2003). Alignment dominates attraction up to ~ 2.5 BL while attraction dominates for larger distances. At very large distances, in the limit of the field of perception, attraction is obviously expected to decrease as well, although we could not measure it in our experimental tanks that were, at most, of a maximum radius of 35 cm. Experimental evidence for alignment interactions is interesting since (i) no alignment is required to produce most of the collective patterns reported in other models of schools (Strömbom, 2011; Strömbom et al., 2015; Huepe et al., 2015) and (ii) alignment was not found in experiments with other species (Katz et al., 2011; Herbert-Read et al., 2011). Whether this is due to differences related to the species or to the methods is still unclear.

We have also investigated how information from multiple sources may be integrated. In our bottom-up approach, this step is crucial to model larger groups from the social and non-social interactions inferred from experiments with 1 or 2 fish. Our hypothesis that a fish only reacts to the strongest stimulus (either social or non-social) is not confirmed by the model, in that it does not reproduce the empirical properties of the polarised schools of *H. rhodostomus*. However, our simulations suggest that such a mechanism may be responsible for swarming species in which individuals are aggregated and loosely polarised.

6.1.3 Propagation of information and decision-making

In Part II (Chapters 4 and 5), we investigated how information propagates in schools of *H. rhodostomus* and leads to collective decisions.

We showed that spontaneous collective U-turns were initiated by individuals that turned around at the front of the school and that this turning was propagated to the back of the school. This is in agreement with the perception anisotropy detected in Part I. The directionality of the propagation of information in the schools during the collective U-turns is confirmed by two independent methods (Appendices A and B).

We find that collective U-turns are preceded by a period during which fish are slowing down. It has been shown in other fish species that speed affects alignment among individuals (Gautrais et al., 2012), leading slow groups to be less polarised than fast groups (Steven V. Viscido et al., 2004; Hemelrijk and Hildenbrandt, 2008; Tunstrøm et al., 2013; Calovi et al., 2014). In general, at slower speed, there is less inertia to turn, resulting in weaker polarisation (Kunz and Hemelrijk, 2003; Hemelrijk et al., 2010) and more fluctuations in the swimming direction of the fish (Marconi et al., 2008; Calovi et al., 2015). As its speed decreases, the fish school is in a state closer to the transition between schooling (strong alignment) and swarming (weak alignment), i.e. where Calovi et al. (2015) have shown that both fluctuations in fish orientation and the sensitivity of the school to a perturbation increase. It is therefore not surprising that U-turns occur after the group has slowed down.

Within the ranges of group sizes we tested (1 to 10 fish), there was no damping nor positive feedback in the temporal propagation of information. Although it could be argued that this result may be different for larger groups, in sheep (*Ovis aries*), an amplification (i.e. the rate of following behaviours increases as the number of individuals committed in the following behaviour increases) has been reported in groups of at least 6 individuals (Pillot et al., 2011). We showed that for the largest group sizes, a heading opposite to that of the majority is hardly propagated, resulting in a strong decrease in the frequency of the collective

U-turns with increasing group size. Schools of fish are thus less sensitive to internal local perturbations when the number of individuals increases, thus similar to empirical findings in sheep (Toulet et al., 2015).

Our model shows that consensus about the direction chosen at the group-level is achieved by choosing the direction of the local majority. Assuming that fish interact locally and not with all the group-members, even in the small group sizes we consider, is supported by the work presented in Appendix B that suggests that fish mainly interact with a small number of influential neighbours (typically 1 or 2). Temporary leadership is explained by the anisotropy in the perception of fish, that favours aligning with frontal individuals. Preliminary results of our conditioning experiments (Chapter 5) show that the composition of the group (i.e. the ratio between informed and non-informed fish) also affects how the group responds to external and aversive perturbations (here, a conditioned stimulus (green light) triggering an escape behaviour). Namely, our results suggest that a critical proportion of conditioned individuals was necessary to reach a consensus and trigger a collective escape. Although it needs to be confirmed by additional replicates, this finding confirm previous results in appetitive conditioning (Miller et al., 2013). We also expect that this critical proportion depends on group size but it is unknown so far how it exactly does so. This would test previous theoretical predictions (Couzin et al., 2005).

6.1.4 Methodological contributions

In the introduction of this thesis, I have discussed the challenges of the study of complex biological phenomena like collective motion in fish schools. To cope with the complexity, we used a diversity of methods tightly combining experiments, data analyses and computational modelling. These methods may be of interest to others in the field of Collective Behaviour and in related fields such as, for instance, Ecotoxicology, where accurate quantification of the (collective) behaviour of fish is critical to state the effects of toxic chemicals.

In Chapter 2, we used a bottom-up experimental approach to measure the interactions involved in the coordination of schools, as was previously done with another species (Gautrais et al., 2009, 2012). The large amount of experimental data that we have collected has motivated the use of a new method to measure social and non-social interactions. This method limits the number of parameters that we need to fit and provides an explicit model, in contrast to the force-map method used by others (Katz et al., 2011). In our case, to disentangle the social interactions and given the aimed resolution, a three-dimensional map of the angle turned as a function of the distance to the neighbour, the angular position of the neighbour and the relative orientation between the focal fish and its neighbour would have required fitting 36,000 parameters (and thus more empirical data points than the 173,817 burst-and-coast cycles we obtained from 1117 min of videos at 50 Hz), while our method required only fitting of 200 parameters and provided an explicit model.

In the former work with the bottom-up approach (Gautrais et al., 2009, 2012), the interactions of a fish with the walls of the tank and with other fish were phenomenological – the mathematical functions in the model were motivated by the fit to data. Here, we favoured a normative model (Laan et al., 2017), in which the equations in our model are derived from a logical framework based on physical analogies and symmetry considerations. Thus, the equations used in our model to depict non-social and social interactions may have greater generality when applied to new species.

In Chapter 3, we specifically address how fish integrate information from multiple (social and non-social) sources. We developed a new method based on behavioural action maps derived from experimental data. We tested hypotheses regarding the integration of information. Despite the negative outcome regarding the hypothesis tested (namely, that fish would only react to the strongest stimulus when making a decision), the method will still be used to test alternative hypotheses. Because the method does not require any assumption regarding the interaction mechanisms, it is likely that it can be applied to a variety of fish species, even if their swimming behaviour strongly differs from *H. rhodostomus*. However, the method has drawbacks: it assumes that the individual interactions found in groups of 1 and 2 fish are identical to those in larger groups and that the turning angle is a proxy of the importance of a stimulus. Moreover, hypotheses different from ours and involving averaging over several stimuli may sum the error in the measurements (tracking and data treatment) over all stimuli. We emphasise that our results suggest the importance of investigating the mechanism of integration of information, which is rarely studied in the context of collective motion.

In Chapter 4, we used the formalism of the Ising model, developed in Statistical Physics, to investigate the propagation of collective U-turns. We derive a formal definition of local social conformity from the framework of Ising models. This definition is combined with anisotropic perception. Ising models have been widely used to account for collective decisions in humans (Castellano et al., 2009). We think that this formalism, coupled with the interaction mechanisms, may be of interest in animal collective behaviour, to explain patterns where there is consensus at the global level (Ward et al., 2008; Sumpter and Pratt, 2009; Ward et al., 2012).

In Chapter 5, we investigate the relevance of conditioning experiments in studies of collective behaviour of fish. Conditioning experiments may be interesting in the field for the manipulation of behaviour in a

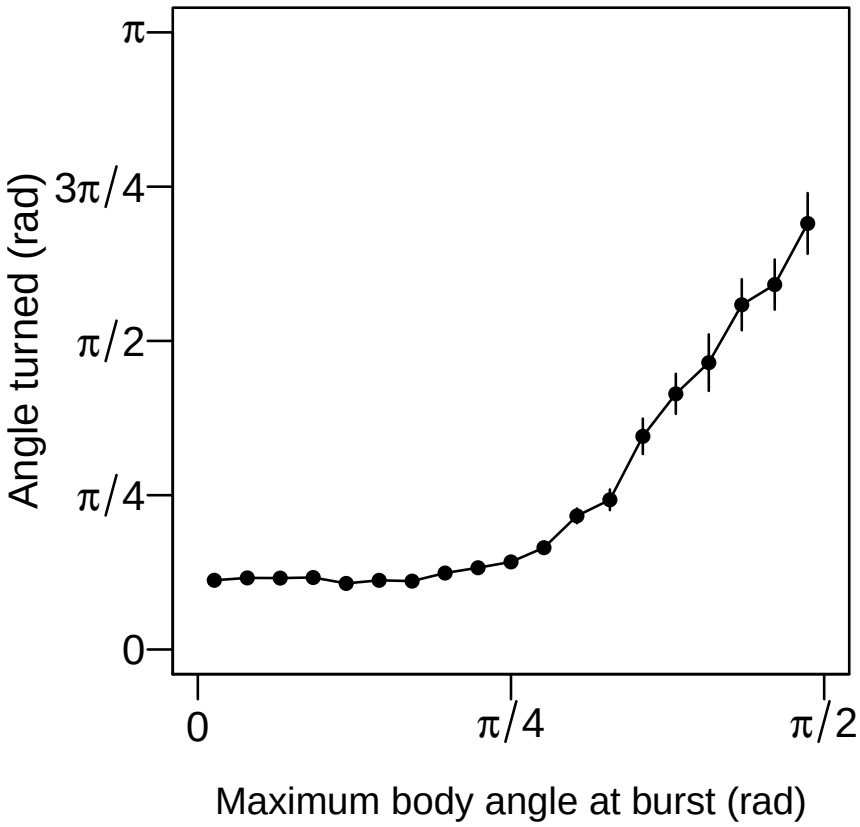


Figure 6.1: Average angle turned versus the maximum body angle of the fish during burst phases. Vertical lines stand for the standard error.

controlled way (Pillot et al., 2010, 2011; Miller et al., 2013; Toulet et al., 2015). In the experiment, we aimed to control (i) which individual would react to a green light (i.e. the conditioned individuals), and (ii) how they would react to it (by escaping away from it). In the present case, we did not entirely succeed to control the identity of the first responder because green light also has an effect on naïve fish. This needs be improved in future experiments, by either changing the conditioning stimulus (or by reducing its intensity) or by familiarising the naive fish to the green light. We did, however, manage to train fish to escape a green light, thus controlling (to some extent) that the green light is perceived as a threat. These results motivate future use of this method to investigate how the proportion of informed individuals among group members affects the propagation of information in the whole group, in particular in the context of escapes.

6.2 Outlook and future work

By focusing on a single species of fish and combining various methods, we have improved our understanding of the behavioural and cognitive mechanisms that govern the schooling behaviour in *H. rhodostomus*. Still, for establishing causal links from the neural scale to the collective scale and providing arguments regarding the fitness of the collective patterns, future work is required. This will be briefly looked at below.

6.2.1 Neural and cognitive scales

We described the burst-and-coast style of fish swimming spontaneously. Future research regarding the motor actions involved in the reactions of fish is required for deriving causal links from the neural scale to the behavioural scale. An interesting method using a virtual reality system has recently been suggested with larvae of zebrafish (Jouary et al., 2016). The authors have made a classification of the possible movements of the tail of the larvae and then monitored the reactions of the larvae in a controlled virtual environment. The changes in the visually induced environment are made thanks to the connection between the motor reaction and the behavioural action. The aim is, eventually, to link the neural activity to the motor reaction inducing a behavioural action, in response to a specific set of stimuli.

In *H. rhodostomus*, we monitored the angle made by the body of the fish and its tail, in particular during the burst phases. We find, for instance, that there is a clear positive correlation between the maximum angle made by the fish during the burst phase and the magnitude of the angle turned (i.e. the stronger the fish bends, the more it turns) (Figure 6.1). Classifying the motor repertoire of the fish (especially regarding the body angle) and mapping it to the behavioural reactions inferred in Chapters 2 and 3 (regarding control of speed and direction) may be a first step to connect the motor reactions of fish to their behavioural responses and thus derive the causal link between the neural activity and the behaviour of the individuals.

6.2.2 Evolutionary perspective

Unfortunately, the animals used in this thesis have all been bred in captivity and were not captured in nature. The ecology of *H. rhodostomus* is not known and we also do not know how our captive animals differ from wild individuals. This prevents any inference about the functions of the schooling behaviour detailed in this manuscript. The general relevance of this study is thus still to be addressed. It is likely that the swimming style of a species constrains the (collective) behavioural repertoire. Characterising in detail the burst-and-coast behaviour of other species (such as the zebrafish, as recently has been done) may provide relevant benchmarks to compare the species and their collective behaviour (Harpaz et al., 2017; Laan et al., 2017). We note that the methods developed in these experimental studies have already been used to examine the impact of predation on schooling behaviour, thanks to measurements based on the burst-and-coast behaviour (Herbert-Read et al., 2017). Such studies, conducted with wild animals and using the quantitative methods developed with controlled experiments to investigate behaviours in detail, are of great interest to learn about the ecological functions of schooling.

Bibliography

Agetsuma, M., Aoki, T., Aoki, R. and Okamoto, H. (2012), Cued fear conditioning in zebrafish (Danio rerio), *in* A. V. Kalueff and A. M. Stewart, eds, 'Zebrafish Protocols for Neurobehavioral Research', Humana Press, Totowa, NJ, pp. 257–264.

Ahl, E. (1924), 'Über eine farbenprächtige Neuheit, Hemigrammus rhodostomus E. AHL, sp. n', *Wschr. Aq. Terr. kde* **2**, 405–406.

Aoki, I. (1982), 'A Simulation Study on the Schooling Mechanism in Fish', *Bulletin of the Japanese Society of Scientific Fisheries* **48**, 1081–1088.

Ashraf, I., Godoy-Diana, R., Halloy, J., Collignon, B. and Thiria, B. (2016), 'Synchronization and collective swimming patterns in fish (Hemigrammus bleheri)', *Journal of The Royal Society Interface* **13**(123).

Atkins, P. and De Paula, J. (2011), *Physical Chemistry for the Life Sciences*, Oxford University Press, New York.

Attanasi, A., Cavagna, A., Del Castello, L., Giardina, I., Grigera, T. S., Jelic, A., Melillo, S., Parisi, L., Pohl, O., Shen, E. and Viale, M. (2014), 'Information transfer and behavioural inertia in starling flocks', *Nature Physics* **10**(9), 691–696.

Attanasi, A., Cavagna, A., Del Castello, L., Giardina, I., Jelic, A., Melillo, S., Parisi, L., Pohl, O., Shen, E. and Viale, M. (2015), 'Emergence of collective changes in travel direction of starling flocks from individual birds fluctuations', *Journal of The Royal Society Interface* **12**(108), 20150319.

Attanasi, A., Cavagna, A., Del Castello, L., Giardina, I., Melillo, S., Parisi, L., Pohl, O., Rossaro, B., Shen, E., Silvestri, E. and Viale, M. (2014), 'Collective behaviour without collective order in wild swarms of midges', *PLoS Comput Biol* **10**(7), e1003697.

Axelsen, B. E., Anker-Nilssen, T., Fossum, P., Kvamme, C. and Nøttestad, L. (2001), 'Pretty patterns but a simple strategy: Predator-prey interactions between juvenile herring and Atlantic puffins observed with multibeam sonar', *Canadian Journal of Zoology* **79**(9), 1586–1596.

Ballerini, M., Cabibbo, N., Candelier, R., Cavagna, A., Cisbani, E., Giardina, I., Lecomte, V., Orlandi, A., Parisi, G., Procaccini, A. and others (2008), 'Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study', *Proceedings of the National Academy of Sciences* **105**(4), 1232–1237.

Ballesta, S., Reymond, G., Pozzobon, M. and Duhamel, J.-R. (2014), 'A real-time 3D video tracking system for monitoring primate groups', *Journal of Neuroscience Methods* **234**, 147 – 152. Measuring Behavior.

Bates, D., Mächler, M., Bolker, B. and Walker, S. (2015), 'Fitting linear mixed-effects models using lme4', *Journal of Statistical Software* **67**(1), 1–48.

Bhattacharya, K. and Vicsek, T. (2010), 'Collective decision making in cohesive flocks', *New Journal of Physics* **12**(9), 093019.

Bialek, W., Cavagna, A., Giardina, I., Mora, T., Pohl, O., Silvestri, E., Viale, M. and Walczak, A. M. (2014), 'Social interactions dominate speed control in poising natural flocks near criticality', *Proceedings of the National Academy of Sciences* **111**(20), 7212–7217.

Bonabeau, E., Theraulaz, G., Deneubourg, J.-L., Aron, S. and Camazine, S. (1997), 'Self-organization in social insects', *Trends in Ecology Evolution* **12**(5), 188 – 193.

Branson, K., Robie, A. A., Bender, J., Perona, P. and Dickinson, M. H. (2009), 'High-throughput ethomics in large groups of Drosophila', *Nat Meth* **6**(6), 451–457.

Brendel, K., Barkema, G. T. and van Beijeren, H. (2003), 'Magnetization reversal times in the two-dimensional Ising model', *Physical Review E* **67**(2), 026119.

Bricard, A., Caussin, J.-B., Desreumaux, N., Dauchot, O. and Bartolo, D. (2013), 'Emergence of macroscopic directed motion in populations of motile colloids', *Nature* **503**(7474), 95–98.

Brown, A. and Poon, W. (2014), 'Ionic effects in self-propelled Pt-coated Janus swimmers', *Soft Matter* **10**(22), 4016–4027.

Brown, C. and Laland, K. N. (2003), 'Social learning in fishes: A review', *Fish and Fisheries* **4**(3), 280–288.

Brown, G. E. (2003), 'Learning about danger: Chemical alarm cues and local risk assessment in prey fishes', *Fish and Fisheries* **4**(3), 227–234.

Brush, S. G. (1967), 'History of the Lenz-Ising Model', *Rev. Mod. Phys.* **39**(4), 883–893.

Buhl, J., Sumpter, D. J. T., Couzin, I. D., Hale, J. J., Despland, E., Miller, E. R. and Simpson, S. J. (2006), 'From disorder to order in marching locusts', *Science* **312**(5778), 1402–1406.

Burgess, H. A., Schoch, H. and Granato, M. (2010), 'Distinct retinal pathways drive spatial orientation behaviors in zebrafish navigation', *Current Biology* **20**(4), 381 – 386.

Buttinoni, I., Volpe, G., Kümmel, F., Volpe, G. and Bechinger, C. (2012), 'Active Brownian motion tunable by light', *Journal of Physics: Condensed Matter* **24**(28), 284129.

Calovi, D. S., Litchinko, A., Lecheval, V., Lopez, U., Pérez-Escudero, A., Chaté, H., Sire, C. and Theraulaz, G. (2017), 'Disentangling and modeling interactions in fish with burst-and-coast swimming', *arXiv:1703.03801* .

Calovi, D. S., Lopez, U., Ngo, S., Sire, C., Chaté, H. and Theraulaz, G. (2014), 'Swarming, schooling, milling: Phase diagram of a data-driven fish school model', *New Journal of Physics* **16**(1), 015026.

Calovi, D. S., Lopez, U., Schuhmacher, P., Chaté, H., Sire, C. and Theraulaz, G. (2015), 'Collective response to perturbations in a data-driven fish school model', *Journal of The Royal Society Interface* **12**(104), 20141362.

Camazine, S., Deneubourg, J.-L., Franks, N., Sneyd, J., Theraulaz, G. and Bonabeau, E. (2001), *Self-Organization in Biological Systems*, Princeton studies in complexity, Princeton University Press.

Carvalho, F. R., Bertaco, V. A. and Jerep, F. C. (2010), 'Hemigrammus tocantinsi: A new species from the upper rio Tocantins basin, Central Brazil (Characiformes: Characidae)', *Neotropical Ichthyology* **8**, 247 – 254.

Castellano, C., Fortunato, S. and Loreto, V. (2009), 'Statistical physics of social dynamics', *Reviews of Modern Physics* **81**(2), 591–646.

Collignon, B., Séguret, A. and Halloy, J. (2016), 'A stochastic vision-based model inspired by zebrafish collective behaviour in heterogeneous environments', *Royal Society Open Science* **3**(1).

Couzin, I. D., Krause, J., Franks, N. R. and Levin, S. A. (2005), 'Effective leadership and decision-making in animal groups on the move', *Nature* **433**(7025), 513–516.

Couzin, I. D., Krause, J., James, R., Ruxton, G. D. and Franks, N. R. (2002), 'Collective memory and spatial sorting in animal groups', *Journal of Theoretical Biology* **218**(1), 1 – 11.

Day, R. L., MacDonald, T., Brown, C., Laland, K. N. and Reader, S. M. (2001), 'Interactions between shoal size and conformity in guppy social foraging', *Animal Behaviour* **62**(5), 917 – 925.

de Chaumont, F., Coura, R. D.-S., Serreau, P., Cressant, A., Chabout, J., Granon, S. and Olivo-Marin, J.-C. (2012), 'Computerized video analysis of social interactions in mice', *Nat Meth* **9**(4), 410–417.

Delcourt, J. and Poncin, P. (2012), 'Shoals and schools: Back to the heuristic definitions and quantitative references', *Reviews in Fish Biology and Fisheries* **22**(3), 595–619.

Dell, A. I., Bender, J. A., Branson, K., Couzin, I. D., de Polavieja, G. G., Noldus, L. P. J. J., Pérez-Escudero, A., Perona, P., Straw, A. D., Wikelski, M. and Brose, U. (2014), 'Automated image-based tracking and its application in ecology', *Trends in Ecology Evolution* **29**(7), 417 – 428.

Delsuc, F. (2003), 'Army ants trapped by their evolutionary history', *PLoS Biology* **1**(2), e7.

Deseigne, J., Dauchot, O. and Chaté, H. (2010), 'Collective motion of vibrated polar disks', *Phys. Rev. Lett.* **105**(9), 098001.

Domenici, P. and Batty, R. S. (1994), 'Escape manoeuvres of schooling Clupea harengus', *Journal of Fish Biology* **45**(sA), 97–110.

Dugatkin, L. A. (2013), *Principles of Animal Behavior: Third International Student Edition*, WW Norton & Company.

Ebbinghaus, H. (1913), *Memory: A Contribution to Experimental Psychology*, Teachers College, Columbia University, New York.

Eddelbuettel, D. (2013), *Seamless R and C++ Integration with Rcpp*, Springer.

Eddelbuettel, D. and Francois, R. (2011), 'Rcpp: Seamless R and C++ Integration', *Journal of Statistical Software, Articles* **40**(8), 1–18.

Efferson, C., Lalive, R., Richerson, P. J., McElreath, R. and Lubell, M. (2008), 'Conformists and mavericks: The empirics of frequency-dependent cultural transmission', *Evolution and Human Behavior* **29**(1), 56 – 64.

Elgar, M. A. (1989), 'Predator vigilance and group size in mammals and birds: A critical review of the empirical evidence', *Biological Reviews* **64**(1), 13–33.

Elgeti, J. and Gompper, G. (2015), 'Run-and-tumble dynamics of self-propelled particles in confinement', *EPL (Europhysics Letters)* **109**(5), 58003.

Elgeti, J., Kaupp, U. B. and Gompper, G. (2010), 'Hydrodynamics of sperm cells near surfaces', *Biophysical Journal* **99**(4), 1018 – 1026.

Eschmeyer, W. N., Fricke, R. and van der Laan, R. (2017), 'Catalog of fishes: Genera, species, references', `http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp`.

Evangelista, D. J., Ray, D. D., Raja, S. K. and Hedrick, T. L. (2017), 'Three-dimensional trajectories and network analyses of group behaviour within chimney swift flocks during approaches to the roost', *Proceedings of the Royal Society of London B: Biological Sciences* **284**(1849).

Farkas, I., Helbing, D. and Vicsek, T. (2002), 'Social behaviour: Mexican waves in an excitable medium', *Nature* **419**(6903), 131–132.

Farkas, I., Helbing, D. and Vicsek, T. (2003), 'Human waves in stadiums', *Physica A: Statistical Mechanics and its Applications* **330**(1–2), 18 – 24.

Faucher, K., Parmentier, E., Becco, C., Vandewalle, N. and Vandewalle, P. (2010), 'Fish lateral system is required for accurate control of shoaling behaviour', *Animal Behaviour* **79**(3), 679–687.

Foster, W. A. and Treherne, J. E. (1981), 'Evidence for the dilution effect in the selfish herd from fish predation on a marine insect', *Nature* **293**(5832), 466–467.

Fox, R. F. and Roy, R. (1987), 'Steady-state analysis of strongly colored multiplicative noise in a dye laser', *Phys. Rev. A* **35**(4), 1838–1842.

Gallup, A. C., Hale, J. J., Sumpter, D. J. T., Garnier, S., Kacelnik, A., Krebs, J. R. and Couzin, I. D. (2012), 'Visual attention and the acquisition of information in human crowds', *Proceedings of the National Academy of Sciences* **109**(19), 7245–7250.

Gautrais, J., Ginelli, F., Fournier, R., Blanco, S., Soria, M., Chaté, H. and Theraulaz, G. (2012), 'Deciphering interactions in moving animal groups', *PLoS Computational Biology* **8**(9), e1002678.

Gautrais, J., Jost, C., Soria, M., Campo, A., Motsch, S., Fournier, R., Blanco, S. and Theraulaz, G. (2009), 'Analyzing fish movement as a persistent turning walker', *Journal of Mathematical Biology* **58**(3), 429–445.

Gerlotto, F., Bertrand, S., Bez, N. and Gutierrez, M. (2006), 'Waves of agitation inside anchovy schools observed with multibeam sonar: A way to transmit information in response to predation', *ICES Journal of Marine Science* **63**(8), 1405–1417.

Gerum, R. C., Fabry, B., Metzner, C., Beaulieu, M., Ancel, A. and Zitterbart, D. P. (2013), 'The origin of traveling waves in an emperor penguin huddle', *New Journal of Physics* **15**(12), 125022.

Giardina, I. (2008), 'Collective behavior in animal groups: Theoretical models and empirical studies', *HFSP Journal* **2**(4), 205–219.

Ginelli, F., Peruani, F., Pillot, M.-H., Chaté, H., Theraulaz, G. and Bon, R. (2015), 'Intermittent collective dynamics emerge from conflicting imperatives in sheep herds', *Proceedings of the National Academy of Sciences* **112**(41), 12729–12734.

Ginot, F., Theurkauff, I., Levis, D., Ybert, C., Bocquet, L., Berthier, L. and Cottin-Bizonne, C. (2015), 'Nonequilibrium Equation of State in Suspensions of Active Colloids', *Phys. Rev. X* **5**(1), 011004.

Glauber, R. J. (1963), 'Time-dependent statistics of the Ising model', *Journal of Mathematical Physics* **4**(2), 294–307.

Gleason, P. E., Weber, P. G. and Weber, S. P. (1977), 'Effect of group size on avoidance learning in zebra fish,Brachydanio rerio (Pisces: Cyprinidae)', *Animal Learning & Behavior* **5**(2), 213–216.

Harpaz, R., Tkacik, G. and Schneidman, E. (2017), 'Discrete modes of social information processing predict individual behavior of fish in a group', *arXiv preprint arXiv:1703.03065* .

Hartman, E. J. and Abrahams, M. V. (2000), 'Sensory compensation and the detection of predators: The interaction between chemical and visual information', *Proceedings of the Royal Society B: Biological Sciences* **267**(1443), 571–575.

Helfman, G., Collette, B. B., Facey, D. E. and Bowen, B. W. (2009), *The Diversity of Fishes: Biology, Evolution, and Ecology*, John Wiley & Sons.

Hemelrijk, C. K. and Hildenbrandt, H. (2008), 'Self-organized shape and frontal density of fish schools', *Ethology* **114**(3), 245–254.

Hemelrijk, C. K., Hildenbrandt, H., Reinders, J. and Stamhuis, E. J. (2010), 'Emergence of oblong school shape: Models and empirical data of fish', *Ethology* **116**(11), 1099–1112.

Hemelrijk, C. K., van Zuidam, L. and Hildenbrandt, H. (2015), 'What underlies waves of agitation in starling flocks', *Behavioral Ecology and Sociobiology* **69**(5), 755–764.

Hemelrijk, C., Reid, D., Hildenbrandt, H. and Padding, J. (2014), 'The increased efficiency of fish swimming in a school', *Fish and Fisheries* pp. n/a–n/a.

Herbert-Read, J. E. (2016), 'Understanding how animal groups achieve coordinated movement', *Journal of Experimental Biology* **219**(19), 2971–2983.

Herbert-Read, J. E., Buhl, J., Hu, F., Ward, A. J. W. and Sumpter, D. J. T. (2015), 'Initiation and spread of escape waves within animal groups', *Royal Society Open Science* **2**(4), 140355.

Herbert-Read, J. E., Krause, S., Morrell, L. J., Schaerf, T. M., Krause, J. and Ward, A. J. W. (2012), 'The role of individuality in collective group movement', *Proceedings of the Royal Society of London B: Biological Sciences* **280**(1752), 20122564.

Herbert-Read, J. E., Perna, A., Mann, R. P., Schaerf, T. M., Sumpter, D. J. and Ward, A. J. (2011), 'Inferring the rules of interaction of shoaling fish', *Proceedings of the National Academy of Sciences* **108**(46), 18726–18731.

Herbert-Read, J. E., Rosén, E., Szorkovszky, A., Ioannou, C. C., Rogell, B., Perna, A., Ramnarine, I. W., Kotrschal, A., Kolm, N., Krause, J. and Sumpter, D. J. T. (2017), 'How predation shapes the social interaction rules of shoaling fish', *Proceedings of the Royal Society of London B: Biological Sciences* **284**(1861).

Hoare, D. J. and Krause, J. (2003), 'Social organisation, shoal structure and information transfer', *Fish and Fisheries* **4**(3), 269–279.

Honkanen, T. and Ekström, P. (1992), 'Comparative study of the olfactory epithelium of the three-spined stickleback (Gasterosteus aculeatus) and the nine-spined stickleback (Pungitius pungitius)', *Cell and Tissue Research* **269**(2), 267–273.

Horner, J. L., Longo, N. and Bitterman, M. E. (1961), 'A Shuttle Box for Fish and a Control Circuit of General Applicability', *The American Journal of Psychology* **74**(1), 114–120.

Howse, J. R., Jones, R. A. L., Ryan, A. J., Gough, T., Vafabakhsh, R. and Golestanian, R. (2007), 'Self-motile colloidal particles: From directed propulsion to random walk', *Phys. Rev. Lett.* **99**(4), 048102.

Huepe, C., Ferrante, E., Wenseleers, T. and Turgut, A. E. (2015), 'Scale-free correlations in flocking systems with position-based interactions', *Journal of Statistical Physics* **158**(3), 549–562.

Hunter, J. R. (1969), 'Communication of velocity changes in jack mackerel (Trachurus symmetricus) schools', *Animal Behaviour* **17, Part 3**(0), 507 – 514.

Huth, A. and Wissel, C. (1992), 'The simulation of the movement of fish schools', *Journal of Theoretical Biology* **156**(3), 365 – 385.

Inada, Y. and Kawachi, K. (2002), 'Order and flexibility in the motion of fish schools', *Journal of Theoretical Biology* **214**(3), 371–387.

Ioannou, C., Tosh, C., Neville, L. and Krause, J. (2007), 'The confusion effect from neural networks to reduced predation risk', *Behavioral Ecology* **19**(1), 126–130.

IUCN (2017), 'Table 1 - Numbers of threatened species by major groups of organisms (1996–2017)', `http://www.iucnredlist.org/about/summary-statistics#Tables_1_2`.

Jiang, L., Giuggioli, L., Perna, A., Escobedo, R., Lecheval, V., Sire, C., Han, Z. and Theraulaz, G. (2017), 'Identifying influential neighbors in animal flocking', *PLOS Computational Biology* **13**(11), 1–32.

Jouary, A., Haudrechy, M., Candelier, R. and Sumbre, G. (2016), 'A 2D virtual reality system for visual goal-driven navigation in zebrafish larvae', **6**, 34015.

Jung, P. and Hänggi, P. (1987), 'Dynamical systems: A unified colored-noise approximation', *Phys. Rev. A* **35**(10), 4464–4466.

Kastberger, G., Hoetzl, T., Maurer, M., Kranner, I., Weiss, S. and Weihmann, F. (2014), 'Speeding up social waves. Propagation mechanisms of shimmering in giant honeybees', *PLoS ONE* **9**(1), e86315.

Kastberger, G., Schmelzer, E. and Kranner, I. (2008), 'Social waves in giant honeybees repel hornets', *PLoS ONE* **3**(9), e3141.

Katz, Y., Tunstrøm, K., Ioannou, C. C., Huepe, C. and Couzin, I. D. (2011), 'Inferring the structure and dynamics of interactions in schooling fish', *Proceedings of the National Academy of Sciences* **108**(46), 18720–18725.

King, A. J., Wilson, A. M., Wilshin, S. D., Lowe, J., Haddadi, H., Hailes, S. and Morton, A. J. (2012), 'Selfish-herd behaviour of sheep under threat', *Current Biology* **22**(14), R561–R562.

Kolpas, A., Moehlis, J. and Kevrekidis, I. G. (2007), 'Coarse-grained analysis of stochasticity-induced switching between collective motion states', *Proceedings of the National Academy of Sciences* **104**(14), 5931–5935.

Krause, J. and Ruxton, G. D. (2002), *Living in Groups*, Oxford University Press.

Kudrolli, A., Lumay, G., Volfson, D. and Tsimring, L. S. (2008), 'Swarming and swirling in self-propelled polar granular rods', *Phys. Rev. Lett.* **100**(5), 058001.

Kunz, H. and Hemelrijk, C. K. (2003), 'Artificial fish schools: Collective effects of school size, body size, and body form', *Artificial Life* **9**(3), 237–253.

Laan, A., Gil de Sagredo, R. and de Polavieja, G. G. (2017), 'Signatures of optimal control in pairs of schooling zebrafish', *Proceedings of the Royal Society of London B: Biological Sciences* **284**(1852).

Latane, B. (1981), 'The psychology of social impact', *American psychologist* **36**(4), 343–356.

Lemasson, B. H., Anderson, J. J. and Goodwin, R. A. (2009), 'Collective motion in animal groups from a neurobiological perspective: The adaptive benefits of dynamic sensory loads and selective attention', *Journal of Theoretical Biology* **261**(4), 501 – 510.

Lemasson, B. H., Anderson, J. J. and Goodwin, R. A. (2013), 'Motion-guided attention promotes adaptive communications during social navigation', *Proceedings of the Royal Society of London B: Biological Sciences* **280**(1754).

Lenth, R. V. (2016), 'Least-squares means: The R package lsmeans', *Journal of Statistical Software* **69**(1), 1–33.

Lett, C., Semeria, M., Thiebault, A. and Tremblay, Y. (2014), 'Effects of successive predator attacks on prey aggregations', *Theoretical Ecology* **7**(3), 239–252.

Lopez, U. (2015), Étude expérimentale et modélisation des déplacements collectifs chez deux espèces de poissons, Khulia mugil et Hemigrammus rhodostomus., PhD thesis, Université Toulouse 3 Paul Sabatier, Toulouse.

Lopez, U., Gautrais, J., Couzin, I. D. and Theraulaz, G. (2012), 'From behavioural analyses to models of collective motion in fish schools', *Interface Focus* **2**(6), 693–707.

Marconi, U. M. B., Puglisi, A., Rondoni, L. and Vulpiani, A. (2008), 'Fluctuation–dissipation: Response theory in statistical physics', *Physics Reports* **461**(4–6), 111 – 195.

Marcos Mirande, J. (2009), 'Weighted parsimony phylogeny of the family Characidae (Teleostei: Characiformes)', *Cladistics* **25**(6), 574–613.

Miller, N., Garnier, S., Hartnett, A. T. and Couzin, I. D. (2013), 'Both information and social cohesion determine collective decisions in animal groups', *Proceedings of the National Academy of Sciences* **110**(13), 5263–5268.

Morgan, T. and Laland, K. (2012), 'The biological bases of conformity', *Frontiers in Neuroscience* **6**, 87.

Moussaïd, M., Helbing, D. and Theraulaz, G. (2011), 'How simple rules determine pedestrian behavior and crowd disasters', *Proceedings of the National Academy of Sciences* **108**(17), 6884–6888.

Murre, J. M. J. and Dros, J. (2015), 'Replication and analysis of Ebbinghaus' forgetting curve', *PLOS ONE* **10**(7), 1–23.

Mwaffo, V., Anderson, R. P., Butail, S. and Porfiri, M. (2015), 'A jump persistent turning walker to model zebrafish locomotion', *Journal of The Royal Society Interface* **12**(102), 20140884–20140884.

Nagy, M., Akos, Z., Biro, D. and Vicsek, T. (2010), 'Hierarchical group dynamics in pigeon flocks', *Nature* **464**(7290), 890–893.

Narayan, V., Ramaswamy, S. and Menon, N. (2007), 'Long-lived giant number fluctuations in a swarming granular nematic', *Science* **317**(5834), 105–108.

Nøttestad, L. and Axelsen, B. E. (1999), 'Herring schooling manoeuvres in response to killer whale attacks', *Canadian Journal of Zoology* **77**(10), 1540–1546.

Palacci, J., Cottin-Bizonne, C., Ybert, C. and Bocquet, L. (2010), 'Sedimentation and effective temperature of active colloidal suspensions', *Phys. Rev. Lett.* **105**(8), 088304.

Palacci, J., Sacanna, S., Steinberg, A. P., Pine, D. J. and Chaikin, P. M. (2013), 'Living crystals of light-activated colloidal surfers', *Science* **339**(6122), 936–940.

Parrish, J. K. and Edelstein-Keshet, L. (1999), 'Complexity, pattern, and evolutionary trade-offs in animal aggregation', *Science* **284**(5411), 99–101.

Parrish, J. K., Viscido, S. V. and Grünbaum, D. (2002), 'Self-organized fish schools: An examination of emergent properties', *The biological bulletin* **202**(3), 296–305.

Partridge, B. L. and Pitcher, T. J. (1980), 'The sensory basis of fish schools: Relative roles of lateral line and vision', *Journal of Comparative Physiology* **135**(4), 315–325.

Pérez-Escudero, A., Vicente-Page, J., Hinz, R. C., Arganda, S. and de Polavieja, G. G. (2014), 'idTracker: Tracking individuals in a group by automatic identification of unmarked animals', *Nature Methods* pp. 743–748.

Pillot, M.-H., Gautrais, J., Arrufat, P., Couzin, I. D., Bon, R. and Deneubourg, J.-L. (2011), 'Scalable rules for coherent group motion in a gregarious vertebrate', *PLOS ONE* **6**(1), 1–8.

Pillot, M. H., Gautrais, J., Gouello, J., Michelena, P., sibbald, A. and Bon, R. (2010), 'Moving together: Incidental leaders and naïve followers', *Behavioural Processes* **83**(3), 235 – 241.

Piront, M.-L. and Schmidt, R. (1988), 'Inhibition of long-term memory formation by anti-ependymin antisera after active shock-avoidance learning in goldfish', *Brain Research* **442**(1), 53 – 62.

Pita, D., Moore, B. A., Tyrrell, L. P. and Fernández-Juricic, E. (2015), 'Vision in two cyprinid fish: Implications for collective behavior', *PeerJ* **3**, e1113.

Pitcher, T. J. and Magurran, A. E. (1983), 'Shoal size, patch profitability and information exchange in foraging goldfish', *Animal Behaviour* **31**(2), 546 – 555.

Pitcher, T. J., Misund, O. A., Fernö, A., Totland, B. and Melle, V. (1996), 'Adaptive behaviour of herring schools in the Norwegian Sea as revealed by high-resolution sonar', *ICES Journal of Marine Science: Journal du Conseil* **53**(2), 449–452.

Pitcher, T. J. and Parrish, J. K. (1993), Functions of shoaling behaviour in teleosts, *in* T. J. Pitcher, ed., 'The Behaviour of Teleost Fishes', 2 edn, number 7 *in* 'Fish & Fisheries Series', Springer Netherlands, pp. 294–337.

Pitcher, T. J. and Wyche, C. J. (1983), Predator-avoidance behaviours of sand-eel schools: Why schools seldom split, *in* 'Predators and Prey in Fishes', Springer, pp. 193–204.

Pitcher, T., Magurran, A. and Winfield, I. (1982), 'Fish in larger shoals find food faster', *Behavioral Ecology and Sociobiology* **10**(2), 149–151.

Polverino, G., Phamduy, P., Facci, A. L., Drago, M., Khan, K., Yang, L. and Porfiri, M. (2013), Analysis of fish and bioinspired robotic fish swimming together in a water tunnel, *in* R. J. Martín-Palma and A. Lakhtakia, eds, 'Proceedings of SPIE', Vol. 8686, SPIE, p. 868606.

Portavella, M., Torres, B. and Salas, C. (2004), 'Avoidance response in goldfish: Emotional and temporal involvement of medial and lateral telencephalic pallium', *Journal of Neuroscience* **24**(9), 2335–2342.

Pradel, G., Schachner, M. and Schmidt, R. (1999), 'Inhibition of memory consolidation by antibodies against cell adhesion molecules after active avoidance conditioning in zebrafish', *Journal of Neurobiology* **39**(2), 197–206.

Procaccini, A., Orlandi, A., Cavagna, A., Giardina, I., Zoratto, F., Santucci, D., Chiarotti, F., Hemelrijk, C. K., Alleva, E., Parisi, G. and Carere, C. (2011), 'Propagating waves in starling, Sturnus vulgaris, flocks under predation', *Animal Behaviour* **82**(4), 759–765.

R Core Team (2016), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.

Radakov, D. (1973), *Schooling in the Ecology of Fish*, A Halstead Press book, John Wiley & Sons, New York.

Reader, S. M. and Laland, K. N. (2000), 'Diffusion of foraging innovations in the guppy', *Animal Behaviour* **60**(2), 175–180.

Reis, R. E., Kullander, S. O. and Ferraris, C. J. (2003), *Check List of the Freshwater Fishes of South and Central America*, Edipucrs.

Reuter, H. and Breckling, B. (1994), 'Selforganization of fish schools: An object-oriented model', *Ecological Modelling* **75**(Supplement C), 147 – 159. State-of-the-Art in Ecological Modelling proceedings of ISEM's 8th International Conference.

Robinson, G. E., Fernald, R. D. and Clayton, D. F. (2008), 'Genes and social behavior', *Science* **322**(5903), 896–900.

Romenskyy, M., Herbert-Read, J. E., Ward, A. J. W. and Sumpter, D. J. T. (2017), 'Body size affects the strength of social interactions and spatial organization of a schooling fish (Pseudomugil signifer)', *Royal Society Open Science* **4**(4), 161056.

Rosenthal, S. B., Twomey, C. R., Hartnett, A. T., Wu, H. S. and Couzin, I. D. (2015), 'Revealing the hidden networks of interaction in mobile animal groups allows prediction of complex behavioral contagion', *Proceedings of the National Academy of Sciences* **112**(15), 4690–4695.

Rubin, D. C., Hinton, S. and Wenzel, A. (1999), 'The precise time course of retention.', *Journal of Experimental Psychology: Learning, Memory, and Cognition* **25**(5), 1161.

Russel, W. M. S. and Burch, R. L. (1959), *The Principles of Humane Experimental Technique.*, London: Methuen Co. Ltd.

Schaerf, T. M., Dillingham, P. W. and Ward, A. J. W. (2017), 'The effects of external cues on individual and collective behavior of shoaling fish', *Science Advances* **3**(6).

Schneider, J. and Levine, J. D. (2014), 'Automated identification of social interaction criteria in Drosophila melanogaster', *Biology Letters* **10**(10).

Seghers, B. H. (1974), 'Schooling behavior in the guppy (poecilia reticulata): An evolutionary response to predation', *Evolution* **28**(3), 486–489.

Shaw, E. (1978), 'Schooling fishes: The school, a truly egalitarian form of organization in which all members of the group are alike in influence, offers substantial benefits to its participants', *American Scientist* pp. 166–175.

Shemesh, Y., Sztainberg, Y., Forkosh, O., Shlapobersky, T., Chen, A. and Schneidman, E. (2013), 'High-order social interactions in groups of mice', *eLife* **2**, e00759.

Smeets, J. B. J. and Brenner, E. (1994), 'The difference between the perception of absolute and relative motion: A reaction time study', *Vision Research* **34**(2), 191 – 195.

Steven V. Viscido, Julia K. Parrish and Daniel Grünbaum (2004), 'Individual behavior and emergent properties of fish schools: A comparison of observation and theory', *Marine Ecology Progress Series* **273**, 239–249.

Strandburg-Peshkin, A., Farine, D. R., Couzin, I. D. and Crofoot, M. C. (2015), 'Shared decision-making drives collective movement in wild baboons', *Science* **348**(6241), 1358–1361.

Strandburg-Peshkin, A., Twomey, C. R., Bode, N. W., Kao, A. B., Katz, Y., Ioannou, C. C., Rosenthal, S. B., Torney, C. J., Wu, H. S., Levin, S. A. and others (2013), 'Visual sensory networks and effective information transfer in animal groups', *Current Biology* **23**(17), R709–R711.

Strömbom, D. (2011), 'Collective motion from local attraction', *Journal of Theoretical Biology* **283**(1), 145 – 151.

Strömbom, D., Siljestam, M., Park, J. and Sumpter, D. (2015), 'The shape and dynamics of local attraction', *The European Physical Journal Special Topics* **224**(17), 3311–3323.

Sumpter, D., Buhl, J., Biro, D. and Couzin, I. (2008), 'Information transfer in moving animal groups', *Theory in Biosciences* **127**(2), 177–186.

Sumpter, D. J. (2010), *Collective Animal Behavior*, Princeton University Press.

Sumpter, D. J. and Pratt, S. C. (2009), 'Quorum responses and consensus decision making', *Philosophical Transactions of the Royal Society B: Biological Sciences* **364**(1518), 743–753.

Sumpter, D. J. T., Mann, R. P. and Perna, A. (2012), 'The modelling cycle for collective animal behaviour', *Interface Focus* **2**(6), 764–773.

Tailleur, J. and Cates, M. E. (2009), 'Sedimentation, trapping, and rectification of dilute bacteria', *EPL (Europhysics Letters)* **86**(6), 60002.

Teichmann, H. (1954), 'Vergleichende untersuchungen an der nase der fische', *Zeitschrift für Morphologie und Ökologie der Tiere* **43**(2), 171–212.

Theurkauff, I., Cottin-Bizonne, C., Palacci, J., Ybert, C. and Bocquet, L. (2012), 'Dynamic clustering in active colloidal suspensions with chemical signaling', *Phys. Rev. Lett.* **108**(26), 268303.

Thutupalli, S., Seemann, R. and Herminghaus, S. (2011), 'Swarming behavior of simple model squirmers', *New Journal of Physics* **13**(7), 073021.

Toulet, S., Gautrais, J., Bon, R. and Peruani, F. (2015), 'Imitation combined with a characteristic stimulus duration results in robust collective decision-making', *PLOS ONE* **10**(10), e0140188.

Treherne, J. E. and Foster, W. A. (1981), 'Group transmission of predator avoidance behaviour in a marine insect: The Trafalgar effect', *Animal Behaviour* **29**(3), 911–917.

Tunstrøm, K., Katz, Y., Ioannou, C. C., Huepe, C., Lutz, M. J. and Couzin, I. D. (2013), 'Collective states, multistability and transitional behavior in schooling fish', *PLoS Computational Biology* **9**(2), e1002915.

Turner, G. F. and Pitcher, T. J. (1986), 'Attack abatement: A model for group protection by combined avoidance and dilution', *American Naturalist* pp. 228–240.

Vicsek, T., Czirók, A., Ben-Jacob, E., Cohen, I. and Shochet, O. (1995), 'Novel type of phase transition in a system of self-driven particles', *Phys. Rev. Lett.* **75**(6), 1226–1229.

Videler, J. J. and Weihs, D. (1982), 'Energetic advantages of burst-and-coast swimming of fish at high speeds', *Journal of Experimental Biology* **97**(1), 169–178.

Vladescu, I. D., Marsden, E. J., Schwarz-Linek, J., Martinez, V. A., Arlt, J., Morozov, A. N., Marenduzzo, D., Cates, M. E. and Poon, W. C. K. (2014), 'Filling an emulsion drop with motile bacteria', *Phys. Rev. Lett.* **113**(26), 268101.

Walther, A. and Muller, A. H. E. (2008), 'Janus particles', *Soft Matter* **4**(4), 663–668.

Wang, X. R., Miller, J. M., Lizier, J. T., Prokopenko, M. and Rossi, L. F. (2012), 'Quantifying and tracing information cascades in swarms', *PLoS ONE* **7**(7), e40084.

Ward, A. J., Sumpter, D. J., Couzin, I. D., Hart, P. J. and Krause, J. (2008), 'Quorum decision-making facilitates information transfer in fish shoals', *Proceedings of the National Academy of Sciences* **105**(19), 6948–6953.

Ward, A. J. W. (2004), 'The effects of habitat- and diet-based cues on association preferences in three-spined sticklebacks', *Behavioral Ecology* **15**(6), 925–929.

Ward, A. J. W., Krause, J. and Sumpter, D. J. T. (2012), 'Quorum decision-making in foraging fish shoals', *PLoS ONE* **7**(3), e32411.

Ward, A. J. W., Schaerf, T. M., Herbert-Read, J. E., Morrell, L., Sumpter, D. J. T. and Webster, M. M. (2017), 'Local interactions and global properties of wild, free-ranging stickleback shoals', *Royal Society Open Science* **4**(7).

Webster, M., Atton, N., Ward, A. and Hart, P. (2007), 'Turbidity and foraging rate in threespine sticklebacks: The importance of visual and chemical prey cues', *Behaviour* **144**(11), 1347–1360.

Weihs, D. (1974), 'Energetic advantages of burst swimming of fish', *Journal of Theoretical Biology* **48**(1), 215 – 229.

Weitz, S., Blanco, S., Fournier, R., Gautrais, J., Jost, C. and Theraulaz, G. (2012), 'Modeling collective animal behavior with a cognitive perspective: A methodological framework', *PLoS ONE* **7**(6), e38588.

Wickelgren, W. A. (1974), 'Single-trace fragility theory of memory dynamics', *Memory & Cognition* **2**(4), 775–780.

Woodard, W. T. and Bitterman, M. (1973), 'Pavlovian analysis of avoidance conditioning in the goldfish (Carassius auratus).', *Journal of Comparative and physiological Psychology* **82**(1), 123.

Xu, X., Scott-Scheiern, T., Kempker, L. and Simons, K. (2007), 'Active avoidance conditioning in zebrafish (Danio rerio)', *Neurobiology of Learning and Memory* **87**(1), 72 – 77.

Zheng, M., Kashimori, Y., Hoshino, O., Fujita, K. and Kambara, T. (2005), 'Behavior pattern (innate action) of individuals in fish schools generating efficient collective evasion from predation', *Journal of Theoretical Biology* **235**(2), 153 – 167.

Appendix A

Informative and misinformative interactions in a school of fish

Emanuele Crosato, Li Jiang, Valentin Lecheval, Joseph T. Lizier, X. Rosalind Wang, Pierre Tichit, Guy Theraulaz, Mikhail Prokopenko

Article published in Crosato, E et al. 2018. "Informative and misinformative interactions in a school of fish" Swarm Intelligence. https://doi.org/10.1007/s11721-018-0157-x.

arXiv:1705.01213v1 [q-bio.QM] 3 May 2017

Informative and misinformative interactions in a school of fish

Emanuele Crosato *[1,5], Li Jiang[2,3], Valentin Lecheval[3,4], Joseph T. Lizier[1], X. Rosalind Wang[5], Pierre Tichit[3], Guy Theraulaz[3], and Mikhail Prokopenko[1]

[1] *Complex Systems Research Group and Centre for Complex Systems, Faculty of Engineering & IT, The University of Sydney, Sydney, NSW 2006, Australia.*
[2] *School of Systems Science, Beijing Normal University, Beijing, 100875, P. R. China.*
[3] *Centre de Recherches sur la Cognition Animale, Centre de Biologie Intégrative (CBI), Centre National de la Recherche Scientifique (CNRS), Université Paul Sabatier (UPS), F-31062 Toulouse Cedex 9, France.*
[4] *Groningen Institute for Evolutionary Life Sciences, University of Groningen, Centre for Life Sciences, Nijenborgh 7, 9747AG Groningen, The Netherlands.*
[5] *CSIRO Data61, PO Box 76, Epping, NSW 1710, Australia.*

Abstract

It is generally accepted that, when moving in groups, animals process information to coordinate their motion. Recent studies have begun to apply rigorous methods based on Information Theory to quantify such distributed computation. Following this perspective, we use transfer entropy to quantify dynamic information flows locally in space and time across a school of fish during directional changes around a circular tank, i.e. U-turns. This analysis reveals peaks in information flows during collective U-turns and identifies two different flows: an informative flow (positive transfer entropy) based on fish that have already turned about fish that are turning, and a misinformative flow (negative transfer entropy) based on fish that have not turned yet about fish that are turning. We also reveal that the information flows are related to relative position and alignment between fish, and identify spatial patterns of information and misinformation cascades. This study offers several methodological contributions and we expect further application of these methodologies to reveal intricacies of self-organisation in other animal groups and active matter in general.

Introduction

Collective motion is one of the most striking examples of aggregated coherent behaviour in animal groups, dynamically self-organising out of local interactions between individuals. It is observed in different animal species, such as schools of fish [59, 73], flocks of birds [58, 54, 6, 11], colonies of insects [14, 29, 16, 4, 15] and herds of ungulates [32]. There is an emerging understanding that information plays a *dynamic* role in such a coordination [73], and that *distributed* information processing is a specific mechanism that endows the group with collective computational capabilities [13, 23, 1].

Information transfer is of particular relevance for collective behaviour, where it has been observed that small perturbations cascade through an entire group in a wave-like manner [62, 63, 34, 3], with these cascades conjectured to embody information transfer [73]. This phenomenon is related to underlying causal interactions, and a common goal is to infer physical interaction rules directly from experimental data [36, 30, 35] and measure correlations within a collective.

*emanuele.crosato@sydney.edu.au

Nagy et al. [55] used a variety of correlation functions to measure directional dependencies between the velocities of pairs of pigeons flying in flocks of up to ten individuals, reconstructing the leadership network of the flock. As has been shown later, this network does not correspond to the hierarchy between birds [56]. Information transfer has been extensively studied in flocks of starlings, by observing the propagation of direction changes across the flocks [20, 19, 2]. More recently, Rosenthal et al. [69] attempted to determine a communication structure of a school of fish during its collective evasion manoeuvres manifested through cascades of behavioural change. A functional mapping between sensory inputs and motor responses was inferred by tracking fish position and body posture, and calculating visual fields.

Rather than consider *semantic* or *pragmatic* information, many contemporary studies employ rigorous information theoretic measures that quantify information as uncertainty reduction, following Shannon [24], in order to deal with the stochastic, continuous and noisy nature of intrinsic information processing in natural systems [28]. Distributed information processing is typically dissected into three primitive functions: the *transmission*, *storage* and *modification of information* [38]. *Information dynamics* is a recent framework characterising and measuring each of the primitives information-theoretically [49, 41]. In viewing the state update dynamics of a random process as an information processing event, this framework performs an *information regression* in accounting for where the information to predict that state update can be found by an observer, first identifying predictive information from the past of the process as *information storage*, then predictive information from other sources as *information transfer* (including both pairwise transfer from single sources, and higher-order transfers due to multivariate effects). The framework has been applied to modelling collective behaviour in several complex systems, such as Cellular Automata [46, 47, 48], Ising spin models [9], Genetic Regulatory Networks and other biological networks [45, 64, 26], and neural information processing [33, 78].

This study proposes a domain-independent, information-theoretic approach to detecting and quantifying individual-level dynamics of information transfer in animal groups using this framework. This approach is based on transfer entropy [70], an information-theoretic measure that quantifies the directed and time-asymmetric predictive effect of one random process on another. We aim to characterize the dynamics of how information transfer is conducted in space and time within a *biological* school of fish (*Hemigrammus rhodostomus* or rummy-nose tetras, Figure 1a).

We stress that the predictive information transfer should be considered from the observer perspective, that is, it is the observer that gains (or loses) predictability about a fish motion, having observed another fish. In other words, notwithstanding possible influences among the fish that could potentially be reflected in their information dynamics, our quantitative analysis focuses on the information flow within the school which is detectable by an external observer, captured by the transfer entropy. This means that, whenever we quantify a predictive information flow from a source fish to a destination fish, we attribute the change of predictability (uncertainty) to a third party, be it another fish in the school, a predator approaching the school or an independent experimentalist. Accordingly, this predictive information flow may or may not account for the causal information flow affecting the source and the destination [5, 40] — however it does typically indicate presence of causality, either within the considered pair or from some common cause.

We focus on collective direction changes, i.e. collective U-turns, during which the directional changes of individuals progress in a rapid cascade, at the end of which a coherent motion is re-established within the school. Sets of different U-turns are comparable across experiments under the same conditions, permitting a statistically significant analysis involving an entire set of U-turns.

By looking at the *pointwise* or *local* values of transfer entropy over time, rather than at its

average values, we are not only able to detect information transfer, but also to observe its dynamics over time and across the school. We demonstrate that information is indeed constantly flowing within the school, and identify the source-destination lag where predictive information flow is maximised (which has an interpretation as an observer-detectable reaction time to other fish). The information flow is observed to peak during collective directional changes, where there is a typical "cascade" of predictive gains and losses to be made by observers of these pairwise information interactions. Specifically, we identify two distinct predictive information flows: (i) an "informative" flow, characterised by positive local values of transfer entropy, based on fish that have already changed direction about fish that are turning, and (ii) a "misinformative" flow, characterised by negative local values of transfer entropy, based on fish that have not changed direction yet about the fish that are turning. Finally, we identify spatial patterns coupled with the temporal transfer entropy, which we call spatio-informational motifs. These motifs reveal spatial dependencies between the source of information and its destination, which shape the directed pairwise interactions underlying the informative and misinformative flows. The strong distinction revealed by our quantitative analysis between informative and misinformative flows is expected to have an impact on modelling and understanding the dynamics of collective animal motion.

Information-theoretic measures for collective motion

The study of Wang et al. [77] introduced the use of transfer entropy to investigations of collective motion. This work quantitatively verified the hypothesis that information cascades within an (artificial) swarm can be spatiotemporally revealed by *conditional transfer entropy* [46, 47] and thus correspond to communications, while the collective memory can be captured by *active information storage* [48].

Richardson et al. [67] applied related variants of conditional mutual information, a measure of non-linear dependence between two random variables, to identify dynamical coupling between the trajectories of foraging meerkats. Transfer entropy has also been used to study the response of schools of zebrafish to a robotic replica of the animal [17, 37], and to infer leadership in pairs of bats [57] and simulated zebrafish [18]. Lord et al. [51] also posed the question of identifying individual animals which are directly interacting with other individuals, in a swarm of insects (*Chironomus riparius*). Their approach used conditional mutual information (called "causation entropy" although it does not directly measure causality [40]), inferring "information flows" within the swarm over moving windows of time.

Unlike the study of Wang et al. [77], the above studies quantified average dependencies over time rather than local dependencies at specific time points; for example, leadership relationships in general rather than their (local) dynamics over time. Local versions of transfer entropy and active information storage have been used to measure pairwise correlations in a "swarm" of soldier crabs, finding that decision-making is affected by the group size [74]. Statistical significance was not reported, presumably due to a small sample size. Similar techniques were used to construct interaction networks within teams of simulated RoboCup agents [22].

In this study we focus on local (or pointwise) transfer entropy [70, 46, 43] for specific samples of time-series processes of fish motion, which allows us to reconstruct the dynamics of information flows over time. Local transfer entropy [46], captures information flow from the realisation of a *source* variable Y to a *destination* variable X at time n. As described in Methods, local transfer entropy is defined as the information provided by the source $\mathbf{y_{n-v}} = \{y_{n-v}, y_{n-v-1}, \ldots, y_{n-v-l+1}\}$, where v is a time delay and l is the history length, about the destination x_n in the context of the

past of the destination $\mathbf{x_{n-1}} = \{x_{n-1}, x_{n-2}, \ldots, x_{n-k}\}$, with a history length k:

$$t_{y \to x}(n, v) = \log_2 \frac{p(x_n | \mathbf{x_{n-1}}, \mathbf{y_{n-v}})}{p(x_n | \mathbf{x_{n-1}})}. \tag{1}$$

Importantly, local values of transfer entropy can be negative, while the average transfer entropy is non-negative. Negative values of the local transfer entropies indicate that the source is *misinformative* about the next state of the destination (i.e. it increases uncertainty). Previous studies that used average measures over sliding time windows in order to investigate how information transfer varies over time [67, 51] cannot detect misinformation because they measure average but not local values.

As an observational measure, transfer entropy does not measure causal effect of the source on the target; this can only be established using interventional measures [5, 40, 21, 71]. Rather, transfer entropy measures the predictive information gained from a source variable about the state transition in a target, which may be viewed as *information transfer* when measured on an underlying causal interaction [40]. It should be noted that while some researchers may be initially more interested in causality, the concept of information transfer reveals much about the dynamics that causal effect does not [40], in particular being associated with emergent local structure in dynamics in complex systems [46, 77] and with changes in behaviour, state or regime [12, 9], as well as revealing the misinformative interactions described above. As a particular example, local transfer entropy spatiotemporally highlights emergent glider entities in cellular automata [46], which are analogues of cascading turning waves in swarms (also highlighted by transfer entropy [77]), while local measures of causality do not differentiate these from the background dynamics [40].

It is well known that the internal dynamics within a school of fish depends on the number of fish. For example, for schools of minnows (*Phoxinus phoxinus*), two fish schools are qualitatively different from schools containing three or more — however, the effects seem to level off by the time the school reaches a size of six individuals [60]. Collective behaviour, as well as a stereotypical "phase transition", when an increase in density leads to the onset of directional collective motion, have also been detected in small groups of six glass prawns (*Paratya australiensis*) [52]. Furthermore, at such intermediate group sizes, it has been observed that multiple fish interactions could often be faithfully factorised into pair interactions in one particular species of fish [30].

In our study we investigated information transfer within a school of fish during specific collective direction changes, i.e., U-turns, in which the school collectively reverses its direction. Groups of five fish were placed in a ring-shaped tank (Figure 1b), a design conceived to constrain fish swimming circularly, with the possibility of undergoing U-turns spontaneously, without any obstacles or external factors. A total of 455 U-turns have been observed during 10 trials of one hour duration each. We computed local transfer entropy between each (directed) pair of fish from time series obtained from fish heading. Specifically, the destination process X was defined as the directional change of the destination fish, while the source process Y was defined as the relative heading of the destination fish with respect to the source fish (see Methods). This allowed us to capture the influence of the source-destination fish alignment on the directional changes of the destination. Such influence is usually delayed in time and we estimated the optimal delay (maximizing $\langle t_{y \to x}(n, v) \rangle_n$ [79], see Methods) at $v = 6$, corresponding to 0.12 seconds.

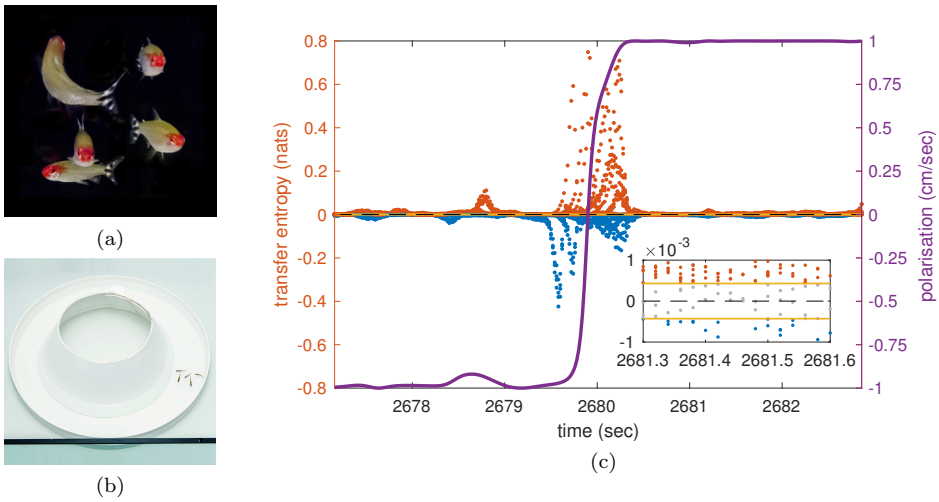


Figure 1: Transfer entropy within the school during a U-turn. Figure 1a is a photo of a spontaneous U-turn initiated by a single fish in a group of five Hemigrammus rhodostomus fish. Figure 1b shows the experimental ring-shaped tank. Figure 1c plots the school's polarisation during a U-turn and the detected transfer entropy over a time interval of approximately 6 seconds. The purple line represents the school's polarisation, while dots represent local values of transfer entropy between all directed pairs of fish: red dots represent positive transfer entropy and blue dots represent negative transfer entropy. Time is discretised in steps of length 0.02 seconds and for each time step 20 points of these local measures are plotted, for the 20 directed pairs formed out of 5 fish. The yellow lines in the inset are the thresholds for considering a value of transfer entropy statistically different from zero ($p < 0.05$ before false discovery rate correction, see Methods). Grey dots between these lines represent values that are not statistically different from zero.

Results

Information flows during U-turns

In order to represent the school's orientation around the tank, we define its polarisation so that it is positive when the school is swimming clockwise and negative when it is swimming anti-clockwise (see Methods). The better the school's average heading is aligned with an ideal circular trajectory around the tank, the higher is the intensity of the polarisation. When the school is facing one of the tank's walls, for example in the middle of a U-turn, the polarisation is zero, and the polarisation flips sign during U-turns. Polarisation allows us to map local values of transfer entropy onto the progression of the collective U-turns.

The analyses of transfer entropy over time reveal that the measure clearly diverges from its baseline in the vicinity of U-turns, as shown in the representative U-turn in Figure 1c (Supplementary Figure S1 shows a longer time interval during which several U-turns can be observed). The figure shows that during regular circular motion, when the school's polarisation is highly pronounced, transfer entropy is low. As the polarisation approaches zero the intensity of transfer entropy grows, peaking near the middle of a U-turn, when polarisation switches its sign.

We clarify that the aim here is *not* to establish transfer entropy as an alternative to polarisation for detecting turn; rather, our aim is to use polarisation to describe the overall progression of

the collective U-turns and then to use transfer entropy to investigate the underlying information flows in the dynamics of such turns. Indeed, transfer entropy is found to be statistically different from zero at many points outside of the U-turns (see Supplementary Figure S1), although the largest values and most concentrated regions of these are during the U-turns. This indicates that information transfer occurs even when fish school together without changing direction; we know that the fish are not executing precisely uniform motion during these in-between periods, and so interpret these small amounts of information transfer as sufficiently underpinning the dynamics of the group maintaining its collective heading.

We also see in Figure 1c that both positive and negative values of transfer entropy are detected. In order to understand the role of the positive and negative information flows during collective motion, in the next section we show the dynamics of transfer entropy for individual pairwise interactions.

Informative and misinformative flows

Our analysis revealed a clear relationship between positive and negative values of transfer entropy and the sequence of individual fish turning, which is illustrated in Figure 2. Figure 2a shows the trajectories of individual fish during the same U-turn depicted in Figure 1. These trajectories are retraced in Figure 2d in terms of polarisation of each fish. It is quite clear that there is a well-defined sequence of individual U-turns during the collective U-turn. Moreover, Figure 2 shows how the transfer entropy maps onto the fish trajectories, both from the fish whose trajectory is traced as a source to the other four fish — i.e. *outgoing* transfer entropy — and, vice versa, from the other four fish to the traced one as a destination — i.e. *incoming* transfer entropy.

The incoming transfer entropy clearly peaks during the destination fish's individual turns and its local values averaged over all sources go from negative, for the first (destination) fish that turns, to positive for the last fish turning (Figures 2b and 2e). In the opposite direction, the outgoing transfer entropy (averaged over all destinations) displays negative peaks only before the source fish has turned, and positive peaks only afterwards (Figures 2c and 2f). Figure 2 suggests that predictive information transfer intensifies only when a destination fish is turning, with this transfer being informative based on source fish that have already turned and misinformative based on source fish that have not turned yet.

This phenomenon can be observed very clearly in Figures 3a and 3b, which show the transfer entropy in both directions for a single fish (the second fish turning in Figures 1 and 2). One positive peak of incoming transfer entropy (indicating informative flow) and three negative ones (misinformative flows) are detected when this fish, as a destination, is undergoing the U-turn (Figure 3a). No other peaks are detected for this fish as a destination. On the other hand, one negative peak of outgoing transfer entropy is detected before the fish, this time as a source, has turned, and three positive peaks are detected after the fish has turned (Figure 3b). These four peaks occur respectively when the first, the third, the fourth and the fifth fish undergo the U-turn, as is evident by comparing Figures 3b and 2d. A movie of the fish undergoing this specific U-turn is provided in Supplementary Video S1, while a detailed reconstruction of the U-turn, showing the dynamics of transfer entropy over time for each directed pair of fish, is provided in Supplementary Video S2.

In order to demonstrate that the phenomenon described here holds for U-turns in general, and not only for the representative one shown in Figure 2, we performed an aggregated analysis of all 455 U-turns observed during the experiment. Since the order in which fish turn is not the same in every U-turn, in this analysis, we refer not to single fish as individuals, but rather to fish in the

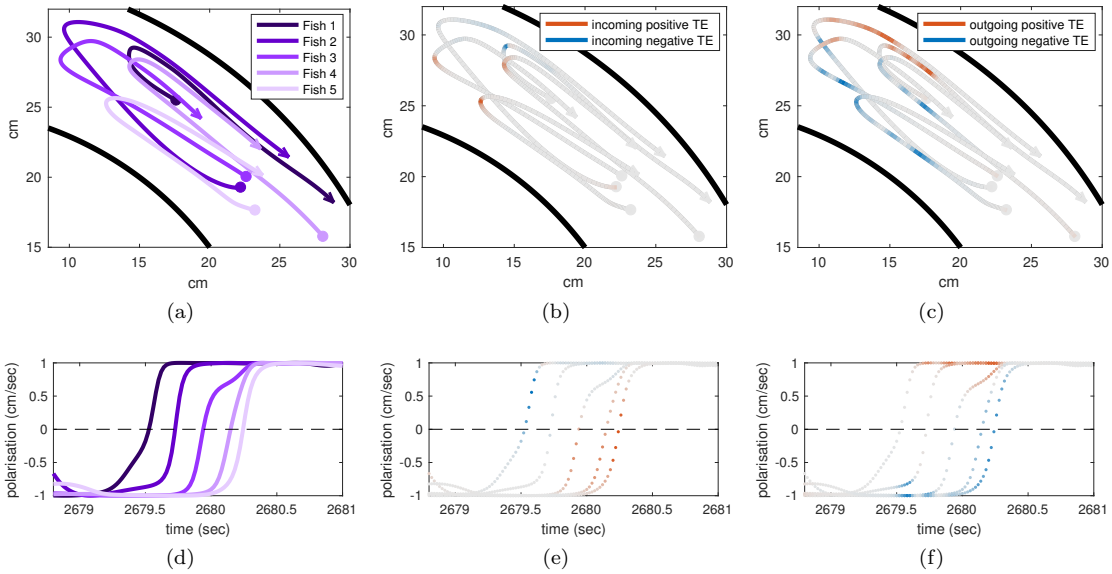


Figure 2: Positive and negative information flows during a U-turn. Figure 2a shows the trajectories of the five fish during the U-turn shown in Figure 1. The two black lines are the inner and the outer walls of the tank, and each of the five trajectories coloured in different shades of purple correspond to a different fish: from darkest purple for the first fish turning (Fish 1), to the lightest purple for the last (Fish 5). The total time interval is approximately 2 seconds, during which all fish turn from swimming anti-clockwise to clockwise. Figure 2d depicts the polarisations of the five fish, showing the temporal sequence of fish turns. Figure 2b shows the fish trajectories again, but this time indicates the value of the incoming local transfer entropy to each fish as a destination, averaged over the other four fish as sources. The colour of each trajectory changes as the fish turn: strong red indicates intense positive transfer entropy; strong blue indicates intense negative transfer entropy; intermediate grey indicates that transfer entropy is close to zero. Figure 2e is obtained analogously to Figure 2b, but the polarisations of the individual fish are shown rather than their trajectories. Figures 2c and 2f mirror Figures 2b and 2e, but where the direction of the transfer entropy has been inverted: the colour of each trajectory or polarisation now indicates the value of the outgoing local transfer entropy from each fish as a source, averaged over the other four fish as destinations.

order in which they turn. Thus, when we refer, for instance, to "the first fish that turns", we may be pointing to a different fish at each U-turn.

The aggregated results are presented in Figures 3c and 3d. Figure 3c shows that incoming transfer entropy peaks for each fish in turning order and gradually grows, from a minimum negative peak corresponding to the first fish turning, to a maximum positive peak corresponding to the last fish turning. Vice versa, Figure 3d shows that outgoing transfer entropy peaks only positively for the first fish turning, which is an informative source about all other fish turning afterwards. For the last fish that turns the peak is negative, since this fish is misinformative about all other fish that have already turned. The second, third and fourth fish present both a negative and a positive peak. The intensity of the negative peaks increases from the second fish to the fourth, while the

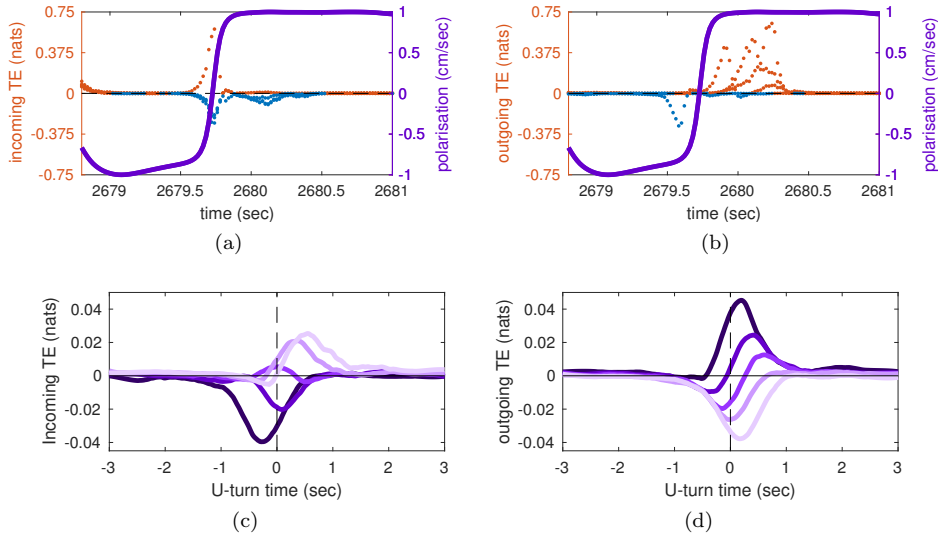
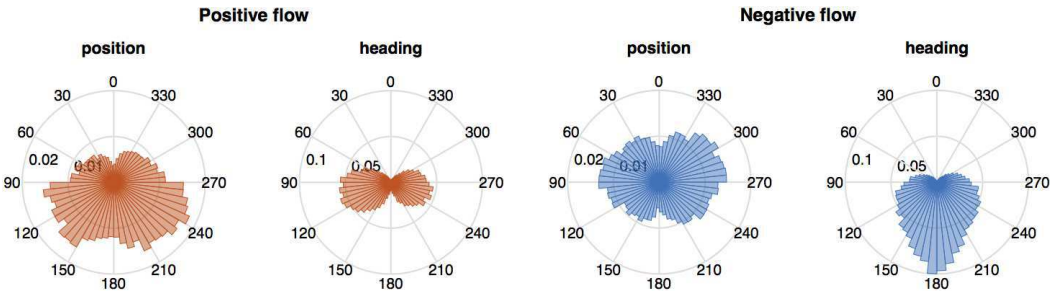


Figure 3: Figure 3a shows the polarisation of the second *fish turning, together with the incoming transfer entropy to that fish as the destination, with the other four fish as the sources: red dots represent positive values and blue dots represent negative values. Figure 3b mirrors Figure 3a, but with the outgoing transfer entropy from that fish as the source, and the other four fish as destinations. In Figure 3c, each purple line corresponds to a fish, with the shade again reflecting the order in which the fish turn (darkest for first fish to turn, and lightest for the last). Now however (in Figure 3c), rather than corresponding to a single U-turn event, the incoming local transfer entropy (to each fish as a destination, averaged over the other four fish as sources) is averaged over all 455 observed U-turns and is shown as a function of time. The horizontal axis is the relative time of the U-turns, with zero being the time when the average polarisation of the school changes sign. Figure 3d mirrors Figure 3c, but where the direction of the transfer entropy has been inverted (showing outgoing transfer from each fish in turning order).*

intensity of the positive peak decreases.

In general, the source fish is informative about all destination fish turning after it and misinformative about any destination fish turning before it. This is because the prior turn of a source helps the observer to predict the later turn of the destination, whereas examining a source which has not turned yet itself is actively unhelpful (misinformative) in predicting the occurrence of such a turn. This also explains why, for a source, the negative peaks come before positives.

The sequential cascade-like dynamics of information flow suggests that the strongest sources of predictive information transfer are fish that have already turned. Moreover our analyses reveal that once a fish has performed a U-turn, its behaviour in general ceases to be predictable based on the behaviour of other fish that swim in opposite direction (in fact such fish would provide misinformative predictions). This suggests an asymmetry of predictive information flows based on and about an individual fish during U-turns.



(a) (b)

Figure 4: Spatio-informational motifs. Each diagram is a circle centred on a source fish with zero heading, providing a reference. In each diagram space is divided into 60 angular sectors measuring 6°. Within each circle we group all pairwise samples from all 455 U-turns such that the source fish is placed in the centre and the destination fish is placed within the circle in one of the sectors. The left circles in Figures 4a and 4b aggregate the relative positions of destination fish, while the right circles aggregate the relative headings of destination fish. The value of each radial sector (for both position and heading) represents the average of the corresponding values of either positive (Figure 4a) of negative (Figure 4b) transfer entropy. For example, the value in each sector of the left diagram of 4a represents the average positive transfer entropy for a destination fish, given it has relative position in that sector with respect to the source fish: all positive values of transfer entropy corresponding to each sector are summed and divided by the total number of values corresponding to that sector. The value in each sector of the right diagram of 4a represents the average positive transfer entropy for a destination fish, given that its heading diverges from the one of the source by an angle in that sector. Figure 4b mirrors Figure 4a this negative transfer entropy. The source fish data are taken from the time points corresponding to the time delay v with respect to the source.

Spatial motifs of information transfer

It is reasonable to assume that predictive information transfer in a school of fish results from spatial interactions among individuals. We investigated the role of pairwise spatial interactions in carrying the positive and negative information flows that we detected in the previous section, looking for spatial patterns of information and misinformation transfer.

In particular we established the statistics of the relative position and heading of the destination fish relative to the source fish, at times when the transfer entropy from the source to the destination is more intense. For this purpose we used radial diagrams (see Figure 4) representing the relative data in terms transfer entropy, focusing separately on their positive (informative) and negative (misinformative) values. In each diagram we aggregate data from all 455 U-turns and all pairs. The diagrams show clear spatial patterns coupled with the transfer entropy, which we call spatio-informational motifs.

We see that positive information transfer is on average more intense from source fish to: a. other fish positioned behind them (Figure 4a, left), and b. to fish with headings closer to perpendicular rather than parallel to them (Figure 4a, right). We know from Figures 2 and 3 that positive transfer entropy is detected from source fish that have already turned to destination fish that are turning. Thus, Figure 4a suggests that a source is more informative about destination fish that are left behind it after a turn, most intensely when the destination fish are executing their own turning manoeuvre to follow the source. Directional relationships from individuals in front towards others

that follow were observed in previous works on birds [55], bats [57] and fish [36, 35, 69].

For negative information transfer (Figure 4b) we see a different spatio-informational motif. Negative information transfer is on average more intense to fish generally positioned at the side and with opposite heading. This aligns with Figures 2 and 3 in that negative transfer entropy typically flows from fish that have not turned yet to those which are turning.

In summary, transfer entropy has a clear spatial signature, showing that the spatiotemporal dependencies in the studied school of fish are not random but reflect specific interactions.

Discussion

Information transfer within animal groups during collective motion is hard to quantify because of implicit and distributed communication channels with delayed and long-ranged effects, selective attention [68] and other species-specific cognitive processes. Here we presented a rigorous framework for detecting and measuring predictive information flows during collective motion, by attending to the dynamic statistical dependence of directional changes in destination fish on relative heading of sources. This predictive information flow should be interpreted as a change (gain or loss) in predictability obtained by an observer.

We studied *Hemigrammus rhodostomus* fish placed in a ring-shaped tank which effectively only allowed the fish to move straight ahead or turn back to perform a U-turn. The individual trajectories of the fish were recorded for hundreds of collective U-turns, enabling us to perform a statistically significant information-theoretical analysis for multiple pairs of source and destination fish.

Transfer entropy was used in detecting pairwise time delayed dependencies within the school. By observing the local dynamics of this measure, we demonstrated that predictive information flows intensify during collective direction changes — i.e. the U-turns — a hypothesis that until now was not verified in a real biological system. Furthermore, we identified two distinct predictive information flows within the school: an informative flow based on fish that have already preformed the U-turn about fish that are turning, and a misinformative flow based on fish that have not preformed the U-turn yet about the fish that are turning.

We also explored the role of spatial dynamics in generating the influential interactions that carry the information flows, another well-known problem. In doing so, we mapped the detected values of transfer entropy against fish relative position and heading, identifying clear spatio-informational motifs. Importantly, the positive and negative predictive information flows were shown to be associated with specific spatial signatures of source and destination fish. For example, positive information flow is detected when the source fish is in front of the destination, similarly to what was already observed in previous works on animals [55, 36, 35, 69, 57].

Local transfer entropy as it was applied in this study reveals the dynamics of *pairwise* information transfer. It is well-known that multivariate extensions to the transfer entropy, e.g. conditioning on other information sources, can be useful in terms of eliminating redundant pairwise relationships whilst also capturing higher-order relationships beyond pairwise (i.e. synergies) [46, 47, 40, 75, 81], and as such the identification of *effective* neighbourhoods cannot be accurately inferred using pairwise relationships alone. Improvements are possible by adapting algorithms for deciding when to include higher-order multivariate transfer entropies (and which variables to condition on), developed to study effective networks in brain imaging data sets [50, 25, 53, 72], to collective animal behaviour, as such methods can eliminate redundant connections and detect synergistic effects. Whether or not such algorithms will prove useful for swarm dynamics is an open research question,

with conflicting findings that first suggest that multiple fish interactions could be faithfully factorised into simply pair interactions in one species [30] but conversely that this may not necessarily generalise [36].

In any case, such adaptations to capture multivariate effects will be non-trivial, as it must handle the short-term and dynamic structure of interactions across the collective. Early attempts have been made using (a similar measure to) conditional TE – on average over time windows – in collectives under such algorithms [51], however it remains to be seen what such measures reveal about the collective dynamics on a local scale.

In summary, we have proposed a novel information-theoretic framework for studying the dynamics of information transfer in collective motion and applied it to a school of fish, without making any specific assumptions on fish behavioural traits and/or rules of interaction. This framework can be easily applied to studies of other biological collective phenomena, such as swarming and flocking, artificial multi-agent systems and active matter in general.

Methods

Ethics statement

All experiments have been approved by the Ethics Committee for Animal Experimentation of the Toulouse Research Federation in Biology N1 and comply with the European legislation for animal welfare.

Experimental procedures

70 *Hemigrammus rhodostomus* (rummy-nose tetras) were purchased from Amazonie Labège (`http://www.amazonie.com`) in Toulouse, France. Fish were kept in 150 L aquariums on a 12:12 hour, dark:light photoperiod, at 27.7°C (±0.5°C) and were fed *ad libitum* with fish flakes. Body lengths of the fish used in these experiments were on average 31 mm (± 2.5 mm).

The experimental tank measured 120 × 120 cm, was made of glass and set on top of a box to isolate fish from vibrations. The setup, placed in a chamber made by four opaque white curtains, was surrounded by four LED light panels giving an isotropic lighting. A ring-shaped tank made from two tanks (an outer wall of radius 35 cm and an inner wall, a cone of radius 25 cm at the bottom, both shaping a corridor of 10 cm) was set inside the experimental tank 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 28.1°C (±0.7°C). The conic shape of the inner wall has been chosen to avoid the occlusion on videos of fish swimming too close to the inner wall that would occur with straight walls.

Five fish were randomly sampled from their breeding tank for a trial. Fish were ensured to be used only in one experiment per day at most. Fish were let for 10 minutes to habituate before the start of the trial. A trial consisted in one hour of fish swimming freely (i.e. without any external perturbation).

Data extraction and pre-processing

Fish trajectories were recorded by a Sony HandyCam HD camera filming from above the setup at 50Hz (50 frames per second) in HDTV resolution (1920×1080p). Videos were converted from MTS

to AVI files with the command-line tool FFmpeg 2.4.3. Positions of fish on each frame were tracked with the tracking software idTracker 2.1 [61].

When possible, missing positions of fish have been manually corrected, only during the collective U-turn events detected by the sign changes of polarisation of the fish groups. The corrections have involved manual tracking of fish misidentified by idTracker as well as interpolation or merging of positions in the cases where only one fish was detected instead of several because they were swimming too close from each others for a long time. All sequences less or equal than 50 consecutive missing positions were interpolated. Larger sequence of missing values have been checked by eye to check whether interpolating was reasonable or not — if not, merging positions with closest neighbors was considered.

Time series of positions have been converted from pixels to meters and the origin of the coordinate system $\mathcal{O}(0,0)$ has been set to the centre of the ring-shaped tank. The resulting data set contains 9,273,720 data points (1,854,744 for each fish) including all the ten trials. Velocity was numerically derived from position using the symmetric difference quotient two-point estimation [39]. Heading was then computed as the four-quadrant inverse tangent of velocity and used to compute transfer entropy.

Polarisation

The polarisation is used to represent the orientation of a fish or of the whole school around the tank, which can be clockwise or anti-clockwise. Let Z and $\dot{Z}$ be the two-dimensional position and normalised velocity of a fish, defined as Cartesian vectors with the centre of the tank being the origin — in case of the whole school, Z and $\dot{Z}$ are averaged over all fish. The fish direction along an ideal circular clockwise rotation is described by a unit vector $z = \frac{\omega \times Z}{|\omega \times Z|}$, where ω is a vector orthogonal to plane of the rotation, chosen using the left-hand rule.

The polarisation is defined as $\dot{Z} \cdot z$, so that it is positive when the fish is swimming clockwise and negative when it is swimming anti-clockwise. Also, the better $\dot{Z}$ is aligned with z or $-z$, the higher is the intensity of the polarisation. On the contrary, as $\dot{Z}$ deviates from z or $-z$, the polarisation decreases and eventually reaches zero when $\dot{Z}$ and z are orthogonal. As a consequence, during a U-turn the intensity of the polarisation decreases and becomes zero at least once, before it increases again with the opposite sign.

Local transfer entropy

Transfer entropy [70] is defined in terms of Shannon entropy, a fundamental measure in Information Theory [24] that quantifies the uncertainty of random variables. Shannon entropy of a random variable X is $H(X) = -\sum_{x \in X} p(x) \log_2 p(x)$, where $p(x)$ is the probability of a specific instance x of X. $H(X)$ can be interpreted as the minimal expected number of bits required to encode a value of X without losing information. The joint Shannon entropy between two random variables X and Y is $H(X,Y) = -\sum_{x \in X} \sum_{y \in Y} p(x,y) \log_2 p(x,y)$, where $p(x,y)$ is the joint probability of instances x of X and y of Y. This quantity allows the definition of conditional Shannon entropy as $H(X|Y) = H(X,Y) - H(X)$, which represents the uncertainty of X knowing Y.

In this study we are interested in local (or pointwise) transfer entropy [27, 43] for specific instances of time-series processes of fish motion, which allows us to reconstruct the dynamics of information flows over time. Shannon information content of an instance x_n of process X at time n is defined as $h(x_n) = -\log_2 p(x_n)$. The quantity $h(x_n)$ is the information content attributed

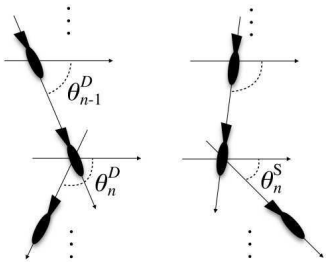
to the specific instance x_n, or the information required to encode or predict that specific value. Conditional Shannon information content of an instance x_n of process X given an instance y_n of process Y is defined as $h(x_n|y_n) = h(x_n, y_n) - h(x_n)$.

Local transfer entropy is defined as the information provided by the source $\mathbf{y_{n-v}} = \{y_{n-v}, y_{n-v-1}, \ldots, y_{n-v-l+1}\}$, where v is a time delay and l is the history length, about the destination x_n in the context of the past of the destination $\mathbf{x_{n-1}} = \{x_{n-1}, x_{n-2}, \ldots, x_{n-k}\}$, with a history length k:

$$\begin{aligned} t_{y\to x}(n,v) &= h(x_n|\mathbf{x_{n-1}}) - h(x_n|\mathbf{x_{n-1}}, \mathbf{y_{n-v}}) \\ &= \log_2 \frac{p(x_n|\mathbf{x_{n-1}}, \mathbf{y_{n-v}})}{p(x_n|\mathbf{x_{n-1}})}. \end{aligned} \tag{2}$$

Transfer entropy $T_{Y\to X}(v)$ is the average of the local transfer entropies $t_{y\to x}(n,v)$ over samples (or over n under a stationary assumption). The transfer entropy is asymmetric in Y and X and is also a dynamic measure (rather than a static measure of correlations) since it measures information in state transitions of the destination.

In order to compute transfer entropy here, the source variable Y and destination variable X are defined in terms of the fish heading. Specifically, X is the first-order divided difference (Newton's difference quotient) of the destination fish heading, while Y is the difference between the two fish headings at the same time. Let Θ_S and Θ_D be respectively the heading time series of the source and the destination fish. We then construct variables X and Y as follows, for all time points n:



$$x_n = \Theta^D_n - \Theta^D_{n-1} \tag{3}$$

$$y_n = \Theta^D_n - \Theta^S_n. \tag{4}$$

Thus, y_n represents the relative heading of the destination fish with respect to the source fish, while x_n represents the directional change of the destination fish. The variables were so defined in order to capture directional changes of the destination fish in relation to its alignment with the source fish, which is considered an important component of movement updates in swarm models [66].

Given the definition of the variables (3) and (4), we computed local transfer entropy $t_{y\to x}(n,v)$ using Equation (2), where v was determined as described in section "Parameters optimisation" that follows. The past state $\mathbf{x_{n-1}}$ of the destination in transfer entropy was defined as a vector of an embedding space of dimensionality k and delay τ, with $\mathbf{x_{n-1}} = \{x_{n-1-j\tau}\}$, for $j = \{0, 1, \ldots, k-1\}$. Finding optimal values for k and τ is also described in section "Parameters optimisation". The state of the source process $\mathbf{y_{n-v}}$ was also defined as a vector of an embedding space whose the dimensionality l and delay τ' were similarly optimised. The local transfer entropy $t_{y\to x}(n,v)$ computed on these variables therefore tells us how much information (l time steps of) the heading of the destination relative to the source adds to our knowledge of the directional change in the destination (some v time steps later), in the context of k past directional changes of the destination. We note that while turning dynamics of the destination may contain more entropy (as rare events), there will only be higher transfer entropy at these events if the source fish is able to add to the prediction of such dynamics.

Computing transfer entropy requires knowledge of the probabilities of x_n and y_n defined in (3) and (4). These are not known a priori, but the measures can be estimated from the data samples using existing techniques. In this study, this was accomplished assuming that the probability distribution function for the observations is a multivariate Gaussian distribution (making the transfer entropy proportional to the Granger causality [7]), using the JIDT software implementation [42].

Also, we assume stationarity of behaviour and homogeneity across the fish, such that we can pool together all pairwise samples from all time steps, for all trials, maximising the number of samples available for the calculation of each measure. For performance efficiency, we make calculations of the local measures using 10 separate sub-sampled sets (sub-sampled evenly across the trials), then recombine into a single resultant information-theoretic data set.

Parameter optimisation

The embedding dimensionality and delay for the source and the past state of the destination need to be appropriately chosen in order to optimise the quality of transfer entropy. The combination (k, τ) for the past state of the destination, as well as the combination (l, τ') for the source, have been optimised separately by minimising the global self-prediction error, as described in [65, 80]. In the case of Markov processes, the optimal dimensionality of the embedding is the order of the process. Lower dimensions do not provide the same amount of predictive information, while higher dimensions add redundancy that weaken the prediction. For non-Markov processes, the algorithm selects the highest dimensionality found to contribute to self-prediction of the destination whilst still being supported by the finite amount of data that we have. Values of the dimensionality between 1 and 10 have been explored in combination with values of the delay between 1 and 5. The optimal combinations were found to be the same for both the source and the past of the destination: $k = l = 3$, $\tau = \tau' = 1$.

The lag v was also optimised. This was done by maximising the average transfer entropy (after the optimisation of k, τ, l and τ') as per [79], over lags between 0.02 and 1 second, at time steps of 0.02 seconds. The average transfer entropy was observed to grow and reach a local maximum at $v = 6$ (0.12 seconds), and then decrease for higher values (see Figure 5). This result might have a biological interpretation: it is plausible for a fish to have a minimum reaction time, which delays the response to behaviour of other fish.

Statistical significance of estimates of local transfer entropy

Theoretically, transfer entropy between two independent variables is zero. However, a non-zero bias (and a variance of estimates around that bias) is likely to be observed when, as in this study, transfer entropy is numerically estimated from a finite number of samples. This leads to the problem of determining whether a non-zero estimated value represents a real relationship between two variables, or is otherwise not statistically significant [80].

There are known statistical significance tests for the average transfer entropy [76, 44, 42], involving comparing the measured value to a null hypothesis that there was no (directed) relationship between the variables. For an average transfer entropy estimated from N samples, one surrogate measurement is constructed by resampling the corresponding $\mathbf{y_{n-v}}$ for each of the N samples of $\{x_n, \mathbf{x_{n-1}}\}$ and then computing the average transfer entropy over these new surrogate samples. This process retains $p(x_n|\mathbf{x_{n-1}})$ and $p(\mathbf{y_{n-v}})$, but not $p(x_n|\mathbf{y_{n-v}}, \mathbf{x_{n-1}})$. Many surrogate measurements are repeated so as to construct a surrogate distribution under this null hypothesis of no directed

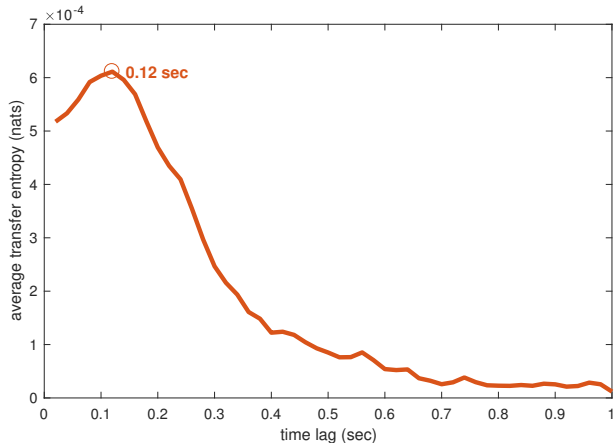


Figure 5: Time lag optimisation. The red line represents the average transfer entropy (with $k = l = 3$, $\tau = \tau' = 1$) over all samples, as a function of the time delay between the source variable and the destination variable, for time delays between 0.02 to 1 seconds (1 to 50 time cycles).

relationship, and the transfer entropy estimate can then be compared in a statistical test against this distribution. For the average transfer entropy measured via the linear-Gaussian estimator, it is known that analytically the surrogates (in nats, and multiplied by $2 \times N$) asymptotically follow a χ^2 distribution with l degrees of freedom [31, 8]. We use this distribution to confirm that the transfer entropy at the selected lag of 0.12 seconds (and indeed all lags tested) is statistically significant compared to the null distribution (at $p < 0.05$ plus a Bonferroni correction for the multiple comparisons across the 50 candidate lags).

Next, we introduce an extension of these methods in order to assess the statistical significance of the *local* values. This simply involves constructing surrogate transfer entropy measurements as before, however this time retaining the local values within those surrogate measurements and building a distribution of those surrogates. Measured local values are then statistically tested against this null distribution of local surrogates to assess their statistical significance.

We generated ten times as many surrogate local values as the number of actual local estimates, with a total of approximately 371 million local surrogates. This large set of surrogate local values was used to estimate p-values of actual local values of the transfer entropy. If p-value is sufficiently small, then the test fails and the value of the transfer entropy is considered significant (the value represents an actual relationship). The Benjamini-Hochberg [10] procedure was used to select the p-value cutoff whilst controlling for the false discovery rate under (N) multiple comparisons.

Acknowledgements

E.C. was supported by the University of Sydney's "Postgraduate Scholarship in the field of Complex Systems" from Faculty of Engineering & IT and by a CSIRO top-up scholarship. L.J. was supported by a grant from the China Scholarship Council (CSC NO.201506040167). V.L. was supported by a doctoral fellowship from the scientific council of the University Paul Sabatier. This study was supported by grants from the Centre National de la Recherche Scientifique and University Paul

Sabatier (project Dynabanc). J.L. was supported through the Australian Research Council DECRA grant DE160100630. The University of Sydney HPC service provided computational resources that have contributed to the research results reported within this paper.

Author contributions statement

G.T. designed research; V.L., P.T. and G.T. performed research; V.L., L.J., P.T., R.W. and G.T. analysed data. E.C., J.L., R.W. and M.P. developed information dynamics methods, performed information-theoretic analysis, and identified information flows and motifs. E.C. designed, developed and run software for the information-theoretic analysis. G.T., J.L., E.C. and M.P. conceived and analysed information cascade. E.C., J.L and M.P. wrote the paper. G.T. and V.L. edited the manuscript and contributed to the writing.

References

[1] Larissa Albantakis, Arend Hintze, Christof Koch, Christoph Adami, and Giulio Tononi. Evolution of integrated causal structures in animats exposed to environments of increasing complexity. *PLOS Computational Biology*, 10(12):1–19, 12 2014.

[2] Alessandro Attanasi, Andrea Cavagna, Lorenzo Del Castello, Irene Giardina, Tomas S Grigera, Asja Jelić, Stefania Melillo, Leonardo Parisi, Oliver Pohl, Edward Shen, et al. Information transfer and behavioural inertia in starling flocks. *Nature physics*, 10(9):691–696, 2014.

[3] Alessandro Attanasi, Andrea Cavagna, Lorenzo Del Castello, Irene Giardina, Asja Jelic, Stefania Melillo, Leonardo Parisi, Oliver Pohl, Edward Shen, and Massimiliano Viale. Emergence of collective changes in travel direction of starling flocks from individual birds' fluctuations. *Journal of The Royal Society Interface*, 12(108), 2015.

[4] Alessandro Attanasi, Andrea Cavagna, Lorenzo Del Castello, Irene Giardina, Stefania Melillo, Leonardo Parisi, Oliver Pohl, Bruno Rossaro, Edward Shen, Edmondo Silvestri, and Massimiliano Viale. Collective behaviour without collective order in wild swarms of midges. *PLoS Comput Biol*, 10(7):1–10, 07 2014.

[5] Nihat Ay and Daniel Polani. Information flows in causal networks. *Advances in complex systems*, 11(01):17–41, 2008.

[6] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic. Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study. *Proceedings of the National Academy of Sciences*, 105(4):1232–1237, 2008.

[7] Lionel Barnett, Adam B. Barrett, and Anil K. Seth. Granger causality and transfer entropy are equivalent for gaussian variables. *Physical Review Letters*, 103(23):238701+, 2009.

[8] Lionel Barnett and Terry Bossomaier. Transfer Entropy as a Log-Likelihood Ratio. *Physical Review Letters*, 109:138105+, 2012.

[9] Lionel Barnett, Joseph T. Lizier, Michael Harré, Anil K. Seth, and Terry Bossomaier. Information flow in a kinetic ising model peaks in the disordered phase. *Physical Review Letters*, 111(17):177203, October 2013.

[10] Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1):289–300, 1995.

[11] William Bialek, Andrea Cavagna, Irene Giardina, Thierry Mora, Edmondo Silvestri, Massimiliano Viale, and Aleksandra M. Walczak. Statistical mechanics for natural flocks of birds. *Proceedings of the National Academy of Sciences*, 109(13):4786–4791, 2012.

[12] Joschka Boedecker, Oliver Obst, Joseph T Lizier, N Michael Mayer, and Minoru Asada. Information processing in echo state networks at the edge of chaos. *Theory in Biosciences*, 131(3):205–213, 2012.

[13] Eric Bonabeau, Marco Dorigo, and Guy Theraulaz. *Swarm intelligence: from natural to artificial systems*. Oxford university press, 1999.

[14] J. Buhl, D. J. T. Sumpter, I. D. Couzin, J. J. Hale, E. Despland, E. R. Miller, and S. J. Simpson. From disorder to order in marching locusts. *Science*, 312(5778):1402–1406, 2006.

[15] Jerome Buhl and Stephen Rogers. Mechanisms underpinning aggregation and collective movement by insect groups. *Current Opinion in Insect Science*, 15:125 – 130, 2016. Pests and resistance * Behavioural ecology.

[16] Jerome Buhl, Gregory A. Sword, Fiona J. Clissold, and Stephen J. Simpson. Group structure in locust migratory bands. *Behavioral Ecology and Sociobiology*, 65(2):265–273, 2010.

[17] Sachit Butail, Fabrizio Ladu, Davide Spinello, and Maurizio Porfiri. Information flow in animal-robot interactions. *Entropy*, 16(3):1315–1330, 2014.

[18] Sachit Butail, Violet Mwaffo, and Maurizio Porfiri. Model-free information-theoretic approach to infer leadership in pairs of zebrafish. *Physical Review E*, 93(4):042411, 2016.

[19] Andrea Cavagna, Irene Giardina, and Francesco Ginelli. Boundary information inflow enhances correlation in flocking. *Physical review letters*, 110(16):168107, 2013.

[20] Andrea Cavagna, SM Duarte Queirós, Irene Giardina, Fabio Stefanini, and Massimiliano Viale. Diffusion of individual birds in starling flocks. In *Proc. R. Soc. B*, volume 280, page 20122484. The Royal Society, 2013.

[21] Daniel Chicharro and Anders Ledberg. When two become one: The limits of causality analysis of brain dynamics. *PLoS ONE*, 7(3):e32466+, 2012.

[22] Oliver M. Cliff, Joseph T. Lizier, X. Rosalind Wang, Peter Wang, Oliver Obst, and Mikhail Prokopenko. Quantifying long-range interactions and coherent structure in multi-agent dynamics. *Artificial Life*, 23(1):34–57, 2017.

[23] Iain D. Couzin. Collective cognition in animal groups. *Trends in Cognitive Sciences*, 13(1):36 – 43, 2009.

[24] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley-Interscience, July 2006.

[25] Luca Faes, Giandomenico Nollo, and Alberto Porta. Information-based detection of nonlinear Granger causality in multivariate processes via a nonuniform embedding technique. *Physical Review E*, 83:051112+, 2011.

[26] Luca Faes and Alberto Porta. Conditional Entropy-Based evaluation of information dynamics in physiological systems. In Michael Wibral, Raul Vicente, and Joseph T. Lizier, editors, *Directed Information Measures in Neuroscience*, Understanding Complex Systems, pages 61–86. Springer Berlin Heidelberg, 2014.

[27] Robert M. Fano. *Transmission of information: A statistical theory of communications*. M.I.T. Press, Cambridge, MA, USA, 1961.

[28] David P Feldman, Carl S McTague, and James P Crutchfield. The organization of intrinsic computation: Complexity-entropy diagrams and the diversity of natural information processing. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 18(4):043106, 2008.

[29] Vincent Fourcassié, Audrey Dussutour, and Jean-Louis Deneubourg. Ant traffic rules. *The Journal of Experimental Biology*, 213(14):2357–2363, 2010.

[30] Jacques Gautrais, Francesco Ginelli, Richard Fournier, Stèphane Blanco, Marc Soria, Hugues Chaté, and Guy Theraulaz. Deciphering interactions in moving animal groups. *PLoS Comput Biol*, 8(9):1–11, 09 2012.

[31] John Geweke. Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association*, 77(378):304–313, 1982.

[32] Francesco Ginelli, Fernando Peruani, Marie-Helène Pillot, Hugues Chaté, Guy Theraulaz, and Richard Bon. Intermittent collective dynamics emerge from conflicting imperatives in sheep herds. *Proceedings of the National Academy of Sciences*, 112(41):12729–12734, 2015.

[33] Carlos Gómez, Joseph T Lizier, Michael Schaum, Patricia Wollstadt, Christine Grützner, Peter Uhlhaas, Christine M Freitag, Sabine Schlitt, Sven Bölte, Roberto Hornero, et al. Reduced predictable information in brain signals in autism spectrum disorder. *Frontiers in Neuroinformatics*, 8(PMC3924322), 2014.

[34] James E Herbert-Read, Jerome Buhl, Feng Hu, Ashley JW Ward, and David JT Sumpter. Initiation and spread of escape waves within animal groups. *Royal Society open science*, 2(4):140355, 2015.

[35] James E. Herbert-Read, Andrea Perna, Richard P. Mann, Timothy M. Schaerf, David J. T. Sumpter, and Ashley J. W. Ward. Inferring the rules of interaction of shoaling fish. *Proceedings of the National Academy of Sciences*, 108(46):18726–18731, 2011.

[36] Yael Katz, Kolbjørn Tunstrøm, Christos C. Ioannou, Cristián Huepe, and Iain D. Couzin. Inferring the structure and dynamics of interactions in schooling fish. *Proceedings of the National Academy of Sciences*, 108(46):18720–18725, 2011.

[37] Fabrizio Ladu, Violet Mwaffo, Jasmine Li, Simone Macrì, and Maurizio Porfiri. Acute caffeine administration affects zebrafish response to a robotic stimulus. *Behavioural brain research*, 289:48–54, 2015.

[38] Chris G. Langton. Computation at the edge of chaos: phase transitions and emergent computation. *Physica D*, 42(1-3):12–37, 1990.

[39] Lee Larson. The symmetric derivative. *Transactions of the American Mathematical Society*, 277(2):589–599, 1983.

[40] J. T. Lizier and M. Prokopenko. Differentiating information transfer and causal effect. *The European Physical Journal B*, 73(4):605–615, 2010.

[41] Joseph T. Lizier. *The Local Information Dynamics of Distributed Computation in Complex Systems*. Springer Theses. Springer, Berlin / Heidelberg, 2013.

[42] Joseph T. Lizier. JIDT: An information-theoretic toolkit for studying the dynamics of complex systems. *Frontiers in Robotics and AI*, 1:11, 2014.

[43] Joseph T. Lizier. Measuring the dynamics of information processing on a local scale in time and space. In Michael Wibral, Raul Vicente, and Joseph T. Lizier, editors, *Directed Information Measures in Neuroscience*, Understanding Complex Systems, pages 161–193. Springer, Berlin/Heidelberg, 2014.

[44] Joseph T. Lizier, Jakob Heinzle, Annette Horstmann, John-Dylan Haynes, and Mikhail Prokopenko. Multivariate information-theoretic measures reveal directed information structure and task relevant changes in fMRI connectivity. *Journal of Computational Neuroscience*, 30(1):85–107, 2011.

[45] Joseph T. Lizier, Siddharth Pritam, and Mikhail Prokopenko. Information dynamics in small-world boolean networks. *Artificial Life*, 17(4):293–314, 2016/05/03 2011.

[46] Joseph T. Lizier, Mikhail Prokopenko, and Albert Y. Zomaya. Local information transfer as a spatiotemporal filter for complex systems. *Physical Review E*, 77(2):026110, 2008.

[47] Joseph T. Lizier, Mikhail Prokopenko, and Albert Y. Zomaya. Information modification and particle collisions in distributed computation. *Chaos*, 20(3):037109, 2010.

[48] Joseph T. Lizier, Mikhail Prokopenko, and Albert Y. Zomaya. Local measures of information storage in complex distributed computation. *Information Sciences*, 208:39–54, 2012.

[49] Joseph T. Lizier, Mikhail Prokopenko, and Albert Y. Zomaya. A framework for the local information dynamics of distributed computation in complex systems. In Mikhail Prokopenko, editor, *Guided Self-Organization: Inception*, volume 9 of *Emergence, Complexity and Computation*, pages 115–158. Springer Berlin Heidelberg, 2014.

[50] Joseph T. Lizier and Mika Rubinov. Multivariate construction of effective computational networks from observational data. Technical Report Preprint 25/2012, Max Planck Institute for Mathematics in the Sciences, 2012.

[51] W. M. Lord, J. Sun, N. T. Ouellette, and E. M. Bollt. Inference of causal information flow in collective animal behavior. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, 2(1):107–116, June 2016.

[52] Richard P Mann, Andrea Perna, Daniel Strömbom, Roman Garnett, James E Herbert-Read, David JT Sumpter, and Ashley JW Ward. Multi-scale inference of interaction rules in animal groups using bayesian model selection. *PLoS Comput Biol*, 9(3):e1002961, 2013.

[53] D. Marinazzo, M. Pellicoro, and S. Stramaglia. Causal information approach to partial conditioning in multivariate data sets. *Computational and Mathematical Methods in Medicine*, 2012:303601+, 2012.

[54] Robert M May. Flight formations in geese and other birds. *Nature*, 282:778–780, 1979.

[55] Máté Nagy, Zsuzsa Ákos, Dora Biro, and Tamás Vicsek. Hierarchical group dynamics in pigeon flocks. *Nature*, 464(7290):890–893, 2010.

[56] Máté Nagy, Gábor Vásárhelyi, Benjamin Pettit, Isabella Roberts-Mariani, Tamás Vicsek, and Dora Biro. Context-dependent hierarchies in pigeons. *Proceedings of the National Academy of Sciences*, 110(32):13049–13054, 2013.

[57] N. Orange and N. Abaid. A transfer entropy analysis of leader-follower interactions in flying bats. *The European Physical Journal Special Topics*, 224(17):3279–3293, 2015.

[58] Carl A. Shollenberger P. B. S. Lissaman. Formation flight of birds. *Science*, 168(3934):1003–1005, 1970.

[59] Julia K Parrish, Steven V Viscido, and Daniel Grünbaum. Self-organized fish schools: an examination of emergent properties. *The biological bulletin*, 202(3):296–305, 2002.

[60] BL Partridge. The effect of school size on the structure and dynamics of minnow schools. *Animal Behaviour*, 28(1):68–IN3, 1980.

[61] Alfonso Pérez-Escudero, Julián Vicente-Page, Robert C Hinz, Sara Arganda, and Gonzalo G de Polavieja. idtracker: tracking individuals in a group by automatic identification of unmarked animals. *Nature methods*, 11(7):743–748, 2014.

[62] Wayne K. Potts. The chorus-line hypothesis of manoeuvre coordination in avian flocks. *Nature*, 309(5966):344–345, 05 1984.

[63] Andrea Procaccini, Alberto Orlandi, Andrea Cavagna, Irene Giardina, Francesca Zoratto, Daniela Santucci, Flavia Chiarotti, Charlotte K Hemelrijk, Enrico Alleva, Giorgio Parisi, et al. Propagating waves in starling, sturnus vulgaris, flocks under predation. *Animal Behaviour*, 82(4):759–765, 2011.

[64] Mikhail Prokopenko, Joseph T. Lizier, Oliver Obst, and X. Rosalind Wang. Relating Fisher information to order parameters. *Physical Review E*, 84(4):041116, 2011.

[65] Mario Ragwitz and Holger Kantz. Markov models from data by simple nonlinear time series predictors in delay embedding spaces. *Phys. Rev. E*, 65:056201, Apr 2002.

[66] Craig W. Reynolds. Flocks, herds and schools: A distributed behavioral model. In *SIGGRAPH '87 Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques*, volume 21, pages 25–34, New York, NY, USA, 1987. ACM.

[67] Thomas O Richardson, Nicolas Perony, Claudio J Tessone, Christophe AH Bousquet, Marta B Manser, and Frank Schweitzer. Dynamical coupling during collective animal motion. *arXiv preprint arXiv:1311.1417*, 2013.

[68] Donald A Riley and Charles R Leith. Multidimensional psychophysics and selective attention in animals. *Psychological Bulletin*, 83(1):138, 1976.

[69] Sara Brin Rosenthal, Colin R. Twomey, Andrew T. Hartnett, Hai Shan Wu, and Iain D. Couzin. Revealing the hidden networks of interaction in mobile animal groups allows prediction of complex behavioral contagion. *Proceedings of the National Academy of Sciences*, 112(15):4690–4695, 2015.

[70] Thomas Schreiber. Measuring information transfer. *Physical Review Letters*, 85(2):461–464, 2000.

[71] Dmitry A. Smirnov. Spurious causalities with transfer entropy. *Phys. Rev. E*, 87:042917, Apr 2013.

[72] S. Stramaglia, Guo-Rong Wu, M. Pellicoro, and D. Marinazzo. Expanding the transfer entropy to identify information circuits in complex systems. *Physical Review E*, 86(6):066211+, 2012.

[73] David Sumpter, Jerome Buhl, Dora Biro, and Iain Couzin. Information transfer in moving animal groups. *Theory in Biosciences*, 127(2):177–186, 2008.

[74] Takenori Tomaru, Hisashi Murakami, Takayuki Niizato, Yuta Nishiyama, Kohei Sonoda, Toru Moriyama, and Yukio-Pegio Gunji. Information transfer in a swarm of soldier crabs. *Artificial Life and Robotics*, pages 1–4, 2016.

[75] Vasily A. Vakorin, Olga A. Krakovska, and Anthony R. McIntosh. Confounding effects of indirect connections on causality estimation. *Journal of Neuroscience Methods*, 184(1):152–160, 2009.

[76] Raul Vicente, Michael Wibral, Michael Lindner, and Gordon Pipa. Transfer entropy—a model-free measure of effective connectivity for the neurosciences. *Journal of Computational Neuroscience*, 30(1):45–67, 2011.

[77] X. Rosalind Wang, Jennifer M. Miller, Joseph T. Lizier, Mikhail Prokopenko, and Louis F. Rossi. Quantifying and tracing information cascades in swarms. *PLoS ONE*, 7(7):e40084, 07 2012.

[78] Michael Wibral, Joseph T. Lizier, and Viola Priesemann. Bits from brains for biologically-inspired computing. *Frontiers in Robotics and AI*, 2(5), 2015.

[79] Michael Wibral, Nicolae Pampu, Viola Priesemann, Felix Siebenhühner, Hannes Seiwert, Michael Lindner, Joseph T. Lizier, and Raul Vicente. Measuring information-transfer delays. *PLoS ONE*, 8(2):e55809+, 2013.

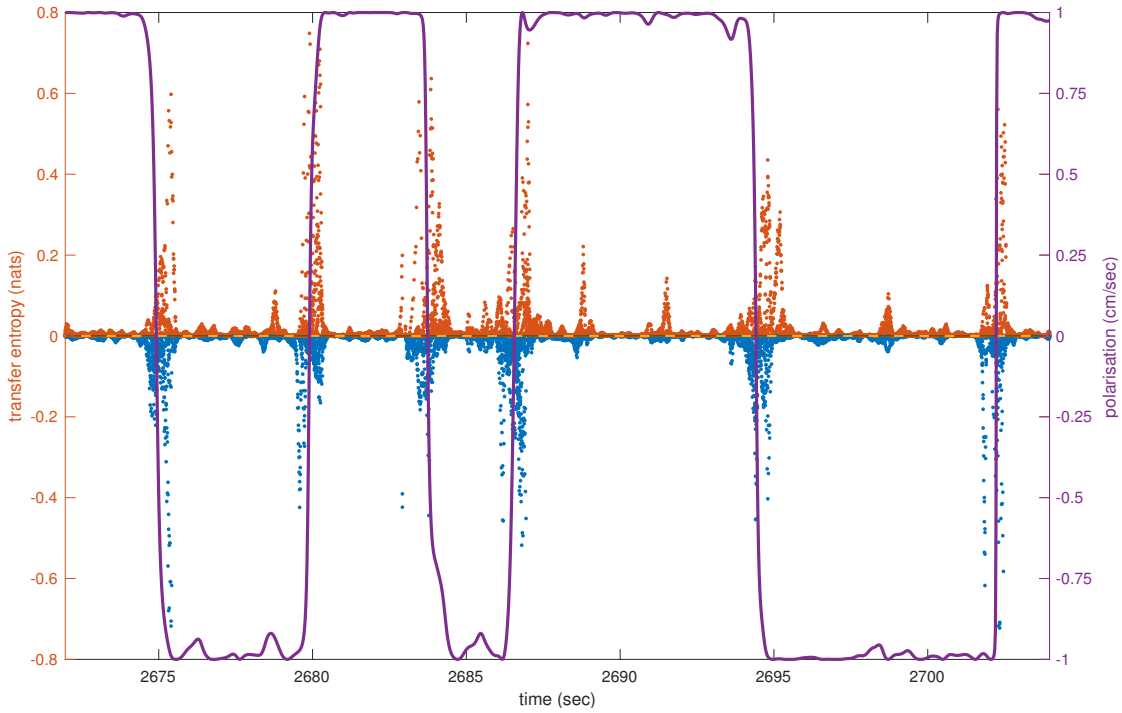
[80] Michael Wibral, Raul Vicente, and Michael Lindner. Transfer entropy in neuroscience. In Michael Wibral, Raul Vicente, and Joseph T. Lizier, editors, *Directed Information Measures in Neuroscience*, Understanding Complex Systems, pages 3–36. Springer, Berlin/Heidelberg, 2014.

[81] Paul L. Williams and Randall D. Beer. Generalized measures of information transfer, 2011. arXiv:1102.1507.

Informative and misinformative interactions in a school of fish

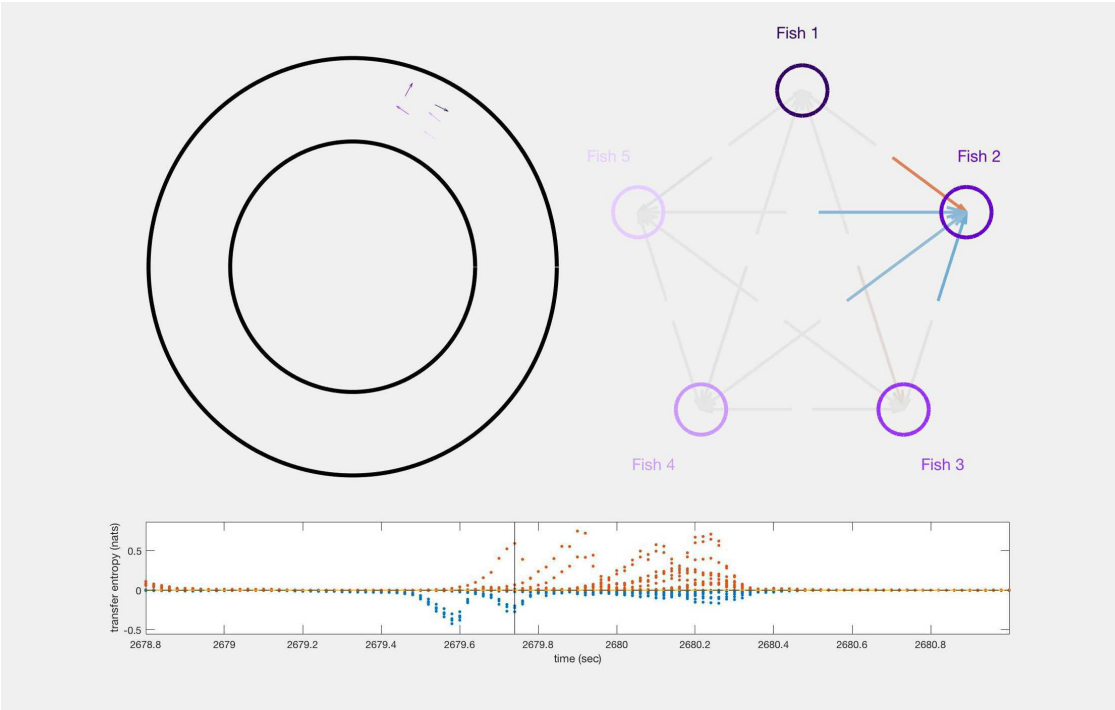
Emanuele Crosato, Li Jiang, Valentin Lecheval, Joseph T. Lizier,
X. Rosalind Wang, Pierre Tichit, Guy Theraulaz, Mikhail Prokopenko

Supplementary information



Supplementary Figure S1. Transfer entropy within the school during several U-turns. The figure plots the school's polarisation during a U-turn and the detected transfer entropy over a time interval of approximately 35 seconds. The purple line represents the school's polarisation, while dots represent local values of transfer entropy between all directed pairs of fish: red dots represent positive transfer entropy and blue dots represent negative transfer entropy. Time is discretised in steps of length 0.02 seconds and for each time step 20 points of these local measures are plotted, for the 20 directed pairs formed out of 5 fish.

Supplementary Video S1. Fish undergoing the representative U-turn (click on the image to open the video). The movie shows five Hemigrammus rhodostomus swimming in the ring-shaped tank for approximately 6 seconds, during which they undergo the U-turn presented in the main article.



Supplementary Video S2 (click on the image to open the video). Animation of the representative U-turn showing transfer entropy dynamics. The movie shows an animation of the representative U-turn over a time interval of approximately 2 seconds. On the top-left is the ring-shaped tank with the five fish, represented by arrows of different shades of purple. On the bottom is the transfer entropy between any directed pair of fish over the time interval: red dots represent positive transfer entropy and blued dots represent negative transfer entropy. Time is discretised in steps of length 0.02 seconds and for each time step 20 points of transfer entropy are plotted, for the 20 directed pairs that can be formed out of 5 fish. On the top-right is the network of transient neighbours changing over time. Each node represents a fish and each directed edge entering a node indicates the transfer entropy to that fish from the other four (the source fish is easily identifiable from the angle of the edges). The colour of the edges changes during the U-turn: strong red indicates intense positive transfer entropy; strong blue indicates intense negative transfer entropy; intermediate grey indicates that transfer entropy is close to zero.

Appendix B

Identifying influential neighbors in animal flocking

Li Jiang, Luca Giuggioli, Andrea Perna, Ramón Escobedo, Valentin Lecheval, Clément Sire, Zhangang Han, Guy Theraulaz

Article published in Jiang, L et al. 2017. "Identifying influential neighbors in animal flocking" PLOS Computational Biology 13(11): e1005822. https://doi.org/10.1371/journal.pcbi.1005822.

PLOS | COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

Identifying influential neighbors in animal flocking

Li Jiang[1,2], Luca Giuggioli[3], Andrea Perna[4], Ramón Escobedo[2], Valentin Lecheval[2,5], Clément Sire[6], Zhangang Han[1], Guy Theraulaz[2]*

1 School of Systems Science, Beijing Normal University, Beijing, China, **2** Centre de Recherches sur la Cognition Animale, Centre de Biologie Intégrative (CBI), Centre National de la Recherche Scientifique (CNRS), Université Paul Sabatier (UPS), Toulouse, France, **3** Bristol Centre for Complexity Sciences, Department of Engineering Mathematics and School of Biological Sciences, University of Bristol, Bristol, United Kingdom, **4** Life Sciences, Roehampton University, London, United Kingdom, **5** Groningen Institute for Evolutionary Life Sciences, University of Groningen, Centre for Life Sciences, Groningen, The Netherlands, **6** Laboratoire de Physique Théorique, CNRS & Université de Toulouse (UPS), Toulouse, France

* guy.theraulaz@univ-tlse3.fr

Abstract

Schools of fish and flocks of birds can move together in synchrony and decide on new directions of movement in a seamless way. This is possible because group members constantly share directional information with their neighbors. Although detecting the directionality of other group members is known to be important to maintain cohesion, it is not clear how many neighbors each individual can simultaneously track and pay attention to, and what the spatial distribution of these influential neighbors is. Here, we address these questions on shoals of *Hemigrammus rhodostomus*, a species of fish exhibiting strong schooling behavior. We adopt a data-driven analysis technique based on the study of short-term directional correlations to identify which neighbors have the strongest influence over the participation of an individual in a collective U-turn event. We find that fish mainly react to one or two neighbors at a time. Moreover, we find no correlation between the distance rank of a neighbor and its likelihood to be influential. We interpret our results in terms of fish allocating sequential and selective attention to their neighbors.

Author summary

Schooling fish exhibit impressive group-level coordination in which multiple individuals move together in a seamless way. This is possible because each individual in the group responds to the movement of other group members. But how many individuals does each fish pay attention to? Which are the influential neighbors? It is necessary to answer these questions in order to understand how directional information propagates across a group. Our research shows that in the rummy-nose tetra species there is a limited number of influential neighbors which are not necessarily the closest ones.

OPEN ACCESS

Citation: Jiang L, Giuggioli L, Perna A, Escobedo R, Lecheval V, Sire C, et al. (2017) Identifying influential neighbors in animal flocking. PLoS Comput Biol 13(11): e1005822. https://doi.org/10.1371/journal.pcbi.1005822

Editor: Philip K Maini, Oxford, UNITED KINGDOM

Received: February 19, 2017

Accepted: October 16, 2017

Published: November 21, 2017

Copyright: © 2017 Jiang et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The data are available at: https://figshare.com/s/93c8bafcb301f8c81b4f. All other relevant data are within the paper and its Supporting Information files.

Funding: LJ was funded by a grant from the China Scholarship Council (CSC NO. 201506040167). LG acknowledges support from EPSRC grant EP/I013717/1. RE has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 655235 "SmartMass". VL was supported by doctoral fellowships from the scientific council of the

University Paul Sabatier. ZH was supported by the National Natural Science Foundation grants 61374165, 31261160495. This study was supported by grants from the Centre National de la Recherche Scientifique and University Paul Sabatier (project Dynabanc). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Collective motion phenomena such as swarming, flocking and schooling behavior have been observed in a large variety of animal species ranging from bacteria to humans [1]. Several theoretical models have been proposed to explain how such large scale coordination patterns emerge from "microscopic level" interaction rules among individual animals [2–7]. These models have been instrumental in improving our understanding of collective motion in real animal groups by providing an indication of which interaction mechanisms are sufficient to reproduce realistic patterns of collective behavior. In particular, most models agree on the fact that two types of interaction are responsible for maintaining group cohesion to achieve coherent collective motion: attraction and alignment.

More recent improvements in remote sensing and video-tracking technologies [8–10] have made possible to automate data collection and test directly theoretical models against highly resolved empirical movement data in various species. Generally, these studies have confirmed the importance that attraction and alignment behavior play in the formation and maintenance of collective movement patterns [11–15]. However, there is a less clear scientific consensus about how these interaction rules are implemented in the sensory-motor responses of individuals. This lack of agreement underscores the importance of answering the following question: how do individuals mediate interactions with multiple neighbors? [16].

Specifically, theoretical studies have postulated a number of factors that are likely to affect the probability and intensity of interactions: distance (metric neighborhood) [2–7], position rank (topological neighborhood) [17], projected size (visual neighborhood) [18–20], and spatial arrangement around a focal individual (Voronoi neighborhood) [13]. Each of these different definitions of influential neighborhood is supported to some extent by computational models and empirical observations.

Rather than siding with one or more of the proposed neighborhood definitions, we adopt a fully data-driven approach with minimalist modeling assumptions. The simplest hypothesis consists of assuming that fish copy the actions of their neighbors, but not instantaneously: the fish reaction takes time to process sensory information and to trigger the appropriate behavioral response. Those assumptions impose a temporal constraint given by the sequential occurrence of the perception of the neighbors' actions, and the movement response [21, 22]. We thus assume that animals following a particular neighbor in a new direction are subject to a time-delay when copying the heading of influential neighbors.

Considerable work has already appeared on the identification of these time-delays. The delays with which individuals align with each other have in fact been exploited to determine social hierarchies in animal groups, as shown, *e.g.*, for pigeon flocks [23], where the leadership network is constructed with link weights given by the delay for which pairwise angle correlation is maximal. Improvements on how to identify such delays from movement data have proposed the use of time-dependence in pairwise angle correlation [24]. A computational analysis, based on similarities between trajectories (Fréchet distance), has also been proposed and implemented in a visual analytic tool [25]. A different approach has made use of a time-ordering procedure on the pairwise angle correlation to determine temporary leader/follower relations in foraging pairs of echolocating bats [26]. The analysis of the bat trajectories was instrumental in identifying transient leadership and coupling it to sensory biases of the species. However, only pairs of individuals were considered and group influence on individual behavior was not investigated.

Since identifying influential neighbors is key to unravel the mechanisms of interaction, there is a need in collective behavior studies to establish transient leadership from the dynamics of the individual trajectories. One way to bridge this gap consists of determining who are

those influential individuals whose heading is being copied more closely by others, how many of such influential neighbors exist, and where are located in the group.

Fish have the ability to choose not only when to copy the heading of another individual, but also the extent to which this heading is copied, that is the similarity and the pace at which fish match the trajectory's curvature of another individual [11, 27]. The closer two (or more) fish are to this matching, the more aligned they are (even if with some delay), and the more faithfully they are following the movement path of the transient leader.

Here, we introduce a procedure that allows us to identify the influential neighbors of fish moving in a group, and we test it along a series of experiments in groups of two and five individuals of the freshwater tropical fish *Hemigrammus rhodostomus* swimming in a ring-shaped tank (see details in Materials and methods). In this set-up, fish swim in a highly synchronized and polarized manner, and can only head in two directions, clockwise or anticlockwise, regularly switching from one to the other. We base our procedure for identifying influential neighbors on time-dependent directional correlations between fish, focussing our analysis on the interactions that occur during these collective U-turns. Indeed, during U-turns, fish have to make a substantial change of direction to reverse their heading, making easier the extraction of the correlation resulting from the direct interactions between individuals rather than other incidental correlations, *e.g.*, their channeled motion in the ring-shaped tank. Moreover, as correlation does not imply causal influence, we need to control for potential spurious correlations. We do so by constructing a null model of collective U-turns to show that the patterns of interaction observed in the experiments are not due to random processes.

Results

Dynamics of collective U-turns

Hemigrammus rhodostomus performs burst-and-coast swimming behavior that consists of sudden heading changes combined with brief accelerations followed by quasi-passive, straight decelerations [15]. Moreover, fish spend most of their time swimming in a single group along the wall of the tank. Fish regularly change their position within the group [28], so that every individual fish can be found at the front of the group.

A typical collective U-turn event starts with the spontaneous turnaround of a single fish (hereafter called the initiator), mostly located at the front of the group [28]. This sudden change of behavior triggers a collective reaction in which all the other individuals in the group make a U-turn themselves, so that, after a short transient, all individuals adopt the same final direction of motion as the initiator. Overall, we analyzed 1586 U-turns of which 1111 were observed in groups of 2 fish and 475 in groups of 5 fish. Fig 1 shows two examples of collective U-turns in groups of $N = 2$ (left column, panels ABC) and $N = 5$ fish (right column, panels DEF; see also supplementary S8 Fig and supplementary S1 and S2 Videos in the Supplementary Information).

Fig 1A shows a first fish F_1 (red color) swimming close to the upper-left region of the tank, followed by a second fish F_2 (purple color) at a distance $d_{12} \approx 8.5$ cm, swimming in the same direction. Right before the U-turn starts (Fig 1A), fish F_1 reduces its speed (circles become closer to each other), the distance d_{12} decreases (to ≈ 5.1 cm), and F_2 also reduces its speed. Then, both fish perform a change of direction which lasts about 1 second and during which fish F_2 clearly follows fish F_1 (see the corresponding circles at each instant of time in Fig 1B). Once the U-turn is completed (Fig 1C), F_1 accelerates again, and so does F_2, which also adopts the direction of motion of F_1. The distance d_{12} increases again (≈ 9.5 cm), due to the larger velocities, and remains of the same order along the depicted trajectory.

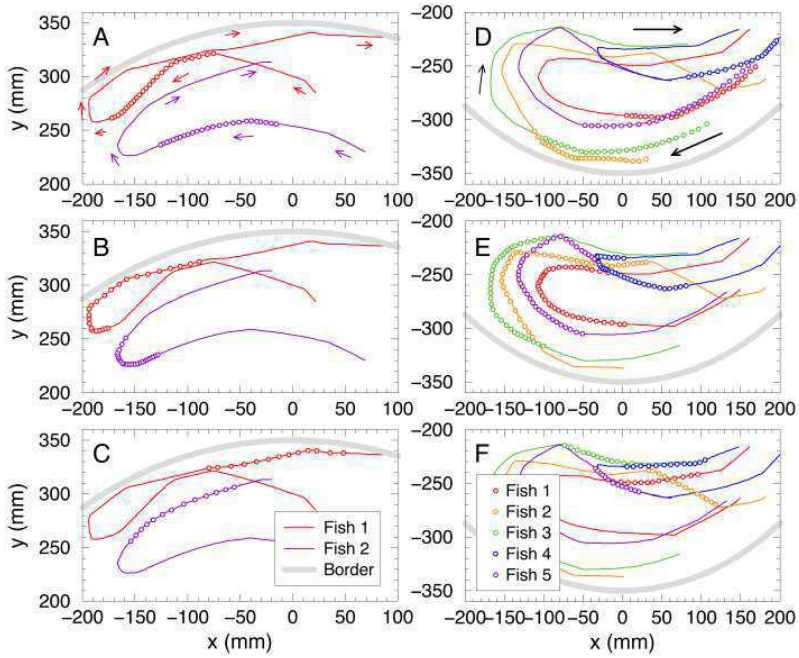


Fig 1. Collective U-turns in groups of two and five fish. Fish trajectories (solid lines) with successive positions (circles) equispaced in time every 0.04s. (ABC): $N = 2$, (DEF): $N = 5$. The top row (AD) displays the collective U-turn one second before it starts, $t \in [t_s - 1\text{ s}, t_s]$, where t_s denotes the time at which the collective U-turn starts. The middle row (BE) displays the collective U-turn, $t \in [t_s, t_e]$, where t_e denotes the end time of the collective U-turn. The bottom row (CF) displays the movement data 0.5 s after collective U-turn's end, $t \in [t_e, t_e + 0.5\text{ s}]$. For visual convenience solid lines indicate the actual fish trajectories before $t_s - 1$ s and after $t_e + 0.5$ s. Arrows indicate the direction of motion. The grey thick line represents the tank border of radius 35 cm.

https://doi.org/10.1371/journal.pcbi.1005822.g001

The situation is less clear when we try to describe collective U-turns in larger groups. Fig 1D, 1E and 1F show a collective U-turn for the case where $N = 5$. Before the U-turn, fish F_2 (orange) seems to be the fish that the rest of the group follows, the first circle of its trajectory being the most advanced one in the direction of motion. In fact, a position order can be inferred from Fig 1D: F_2, F_3, F_5, F_1 and F_4. However, it is rather complicated to extract from Panel E a precise information about which fish is the initiator of the U-turn, in which order the other fish follow, and therefore, who is influencing whom, especially if time-delays and reaction times are taken into account. The same happens with the information about fish's positions after the U-turn, provided by Panel F.

In order to describe rigorously the individual behavior of the N fish during a U-turn, we introduce the angle $\phi_i(t)$ as an instantaneous measure of the direction of motion of a fish F_i; see Fig 2. We assume that the instantaneous heading of a fish F_i can be defined in terms of the velocity vector $\vec{v}_i(t)$, so that $\vec{v}_i = (\cos\phi_i, \sin\phi_i)\,\|\vec{v}_i\|$. The heading of a fish ϕ_i allows us to characterize the angle of incidence of the fish relative to the wall, $\theta_{wi} = \phi_i - \psi_i$, where ψ_i is the angle formed by the position vector of the fish with the horizontal line (see Fig 2). The angle of incidence θ_{wi} is an individual measure that doesn't depend on the heading of another fish. When a fish F_i is swimming along the wall, the value of θ_{wi} is around $\pm 90^\circ$ (we choose, by convention, the positive sign for the anticlockwise angle). In our experiments, most of the time the absolute value of the angle of incidence is close to 90°; equivalently, $|\sin(\theta_{wi}(t))| \approx 1$. When the

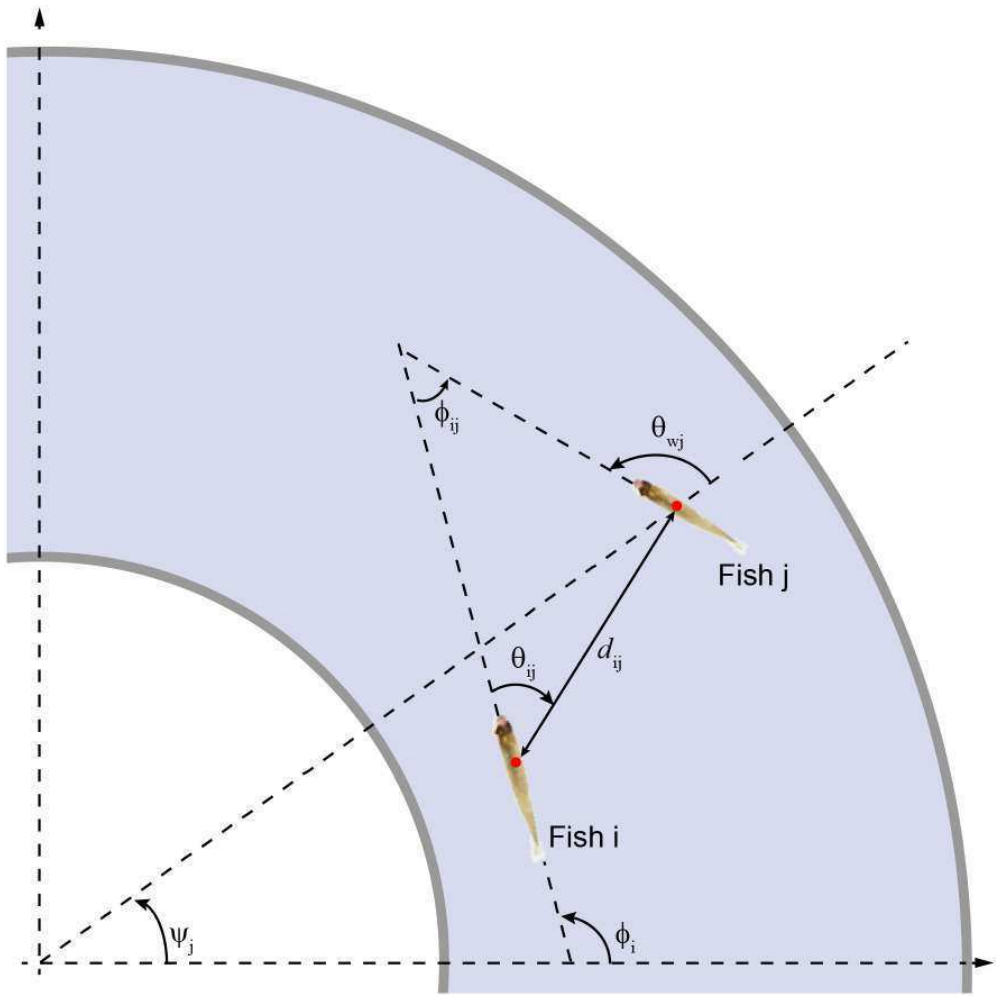


Fig 2. Angles and lengths characterizing the relative position of two fish. Angle ψ_j denotes the angular position of fish F_j with respect to the horizontal (positive values fixed in the anticlockwise direction); angle ϕ_i is the heading of fish F_i; θ_{wi} is the angle of incidence of fish F_i with respect to the outer wall; d_{ij} is the distance between F_i and F_j; θ_{ij} is the viewing angle of F_i with respect to F_j (not necessarily equal to θ_{ji}), and $\phi_{ij} = \phi_j - \phi_i$ is the heading difference of F_i with respect to F_j.

https://doi.org/10.1371/journal.pcbi.1005822.g002

motion is perpendicular to the wall, the incidence is zero if the fish points towards the wall ($\theta_{wi} = 0^\circ$), and maximal if the fish points towards the center of the tank ($\theta_{wi} = 180^\circ$); in both cases, $\sin(\theta_{wi}(t)) = 0$.

The change of sign of angle θ_{wi} can serve as an indicator that a U-turn has taken place. In fact, this allows us to delimit the individual U-turns with precision and, consequently, to determine the start and the end of a collective U-turn.

We define the *start* and *end* times $t_{s,i}$ and $t_{e,i}$ of the individual U-turn of fish F_i in terms of the absolute value of the angle of incidence, $|\theta_{wi}(t)|$. Once a U-turn has been detected, we obtain the time $t_{s,i}$ at which $|\theta_{wi}(t)|$ has decreased (from approximately 90˚) below a given

threshold $\bar{\theta}_s$, and the time $t_{e,i}$ at which $|\theta_{wi}(t)|$ has increased again and is above another given threshold $\bar{\theta}_e$ (see Materials and methods for more details).

Thus, the start of a collective U-turn is determined by the time t_s at which the first individual U-turn starts, while the end of a collective U-turn is given by the time t_e at which the last individual U-turn finishes. That is:

$$t_s = \min_{i=1,\ldots,N}\{t_{s,i}\}, \quad t_e = \max_{i=1,\ldots,N}\{t_{e,i}\}. \tag{1}$$

For each collective U-turn, we have made a convenient time shift so that $t_s = 0$. Then, t_e denotes not only the end time but also the duration of the collective U-turn.

We also introduce an instantaneous measure of how similar the direction of motion of individual fish are across the group. We define the instantaneous group polarization $P(t)$ as the following function of normalized fish velocity vectors:

$$P(t) = \frac{1}{N}\left|\left|\sum_{i=1}^{N}\vec{e}_i(t)\right|\right|, \tag{2}$$

where $\vec{e}_i = \vec{v}_i / \|\vec{v}_i\|$. When all the fish have the same direction then the polarization is maximal and $P(t) = 1$. The minimum value $P(t) = 0$ is reached instead when the velocity vectors cancel.

Figs 3 and 4 depict the two U-turns introduced in Fig 1, in terms of the polarization $P(t)$ and the sine of the angle of incidence of each fish with respect to the outer wall $\theta_{wi}(t)$. The duration of the two illustrated collective U-turns is $t_e = 0.94$ s for $N = 2$ and $t_e = 1.5$ s for $N = 5$.

For both group sizes, the group polarization (Figs 3B and 4B) before and after the U-turn is quite close to 1, showing that before and after the collective U-turn, all individual fish maintain essentially the same common direction. During the U-turn, the polarization decreases, describing a sharp V-form with a minimum at $P(t) \approx 0.27$ for $N = 2$ and $P(t) \approx 0.60$ for $N = 5$. The minimum is reached at approximately half the duration of the collective U-turn, $t_m = (t_s + t_e)/2$: $t_m = 0.47$ s for $N = 2$ and $t_m = 0.75$ s for $N = 5$.

Figs 3C and 4C show the change of direction individually for each fish in both U-turns: from anticlockwise to clockwise direction for $N = 2$, and vice versa for $N = 5$. Fig 3C clearly indicates that at $t \approx 0.3$ s, the fish F_1 has almost completed its individual U-turn, while F_2 has just started to change direction: $\sin(\theta_{w2}(0.3)) \approx 0.98$, while $\sin(\theta_{w1}(0.3)) \approx -0.5$.

In Fig 4C, a similar ordering can be inferred from the times of departure from the bottom line at ordinate $\sin(\theta_{wi}) = -1 + \delta$, where $\delta > 0$ is a small parameter with respect to the range of ordinate values; we used $\delta = 0.1$. Thus, the order is 2-3-1-5-4. However, the order in which individual fish change the sign of their angle of incidence θ_{wi} is different, 2-1-3-5-4, and also different is the arrival order to the top line at ordinate $\sin(\theta_{wi}) = 1 - \delta$: 2-5-1-4-3. Moreover, some of these departure and arrival times are almost identical (see, *e.g.*, F_1 and F_4), and the behavior of the fish during the U-turn is completely different. These difficulties in establishing a consistent order show that another criterion is necessary to identify the relation of influence between fish.

We have based our criterion to decide if a fish is an influential neighbor of another fish on the average value of the time-dependent directional correlation between the two fish along a time window.

For each pair of fish F_i and F_j, we define the directional correlation H_{ij} as a function of the heading of F_i evaluated at time t and the heading of F_j evaluated at a delayed time $t - \tau$, where τ

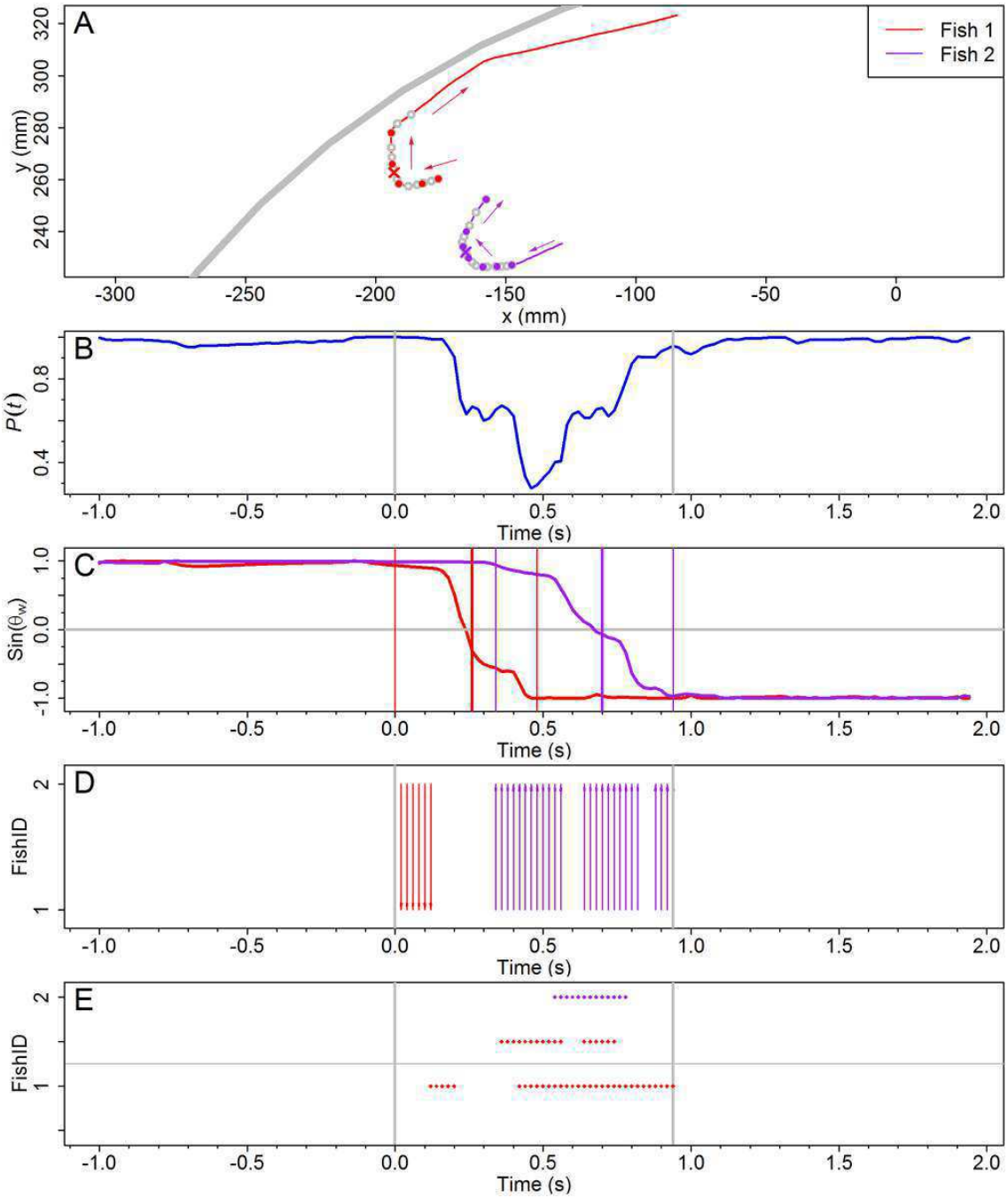


Fig 3. Spatial and temporal dynamics of a collective U-turn for $N = 2$. (A) Individual fish trajectories in the tank during the U-turn. Each individual is represented by a unique color. The temporal sequence is indicated by circles equally spaced over time with a time-step of 0.04 s (empty circles) and 0.1 s (filled circles). Arrows denote direction of motion. Grey wide line is the tank's border. (B) Group polarization $P(t)$, with a minimum value $P_{\min} \approx 0.27$ reached at $t \approx 0.46$ s. (C) Sine of the angle of incidence of fish to the wall: when parallel to the wall, $\sin(\theta_w) = 1$ (anti-clockwise direction) or $\sin(\theta_w) = -1$ (clockwise). The three vertical lines of each color indicate for each fish the beginning, the middle and

the end of its U-turn, with the middle representing the time when a fish has finally reversed its original direction. (D) Interaction with influential neighbors: arrows point from influential neighbors to the focal fish and with the same color as the focal fish. (E) Fish bursting activity and their influential neighbors. Dots at $i = 1, 2$ correspond to bursting activity, blank corresponds to coasting. Dots at $i - 0.5$ represent bursting activity of the neighbor influencing fish i.

https://doi.org/10.1371/journal.pcbi.1005822.g003

is the time-delay [26]:

$$H_{ij}(t, \tau) = \vec{e}_i(t) \cdot \vec{e}_j(t - \tau). \tag{3}$$

The function $H_{ij}(t, \tau)$ is in fact the cosine of the angle formed by the headings $\vec{e}_i(t)$ and $\vec{e}_j(t - \tau)$, and is a measure of how aligned is fish F_i at time t with fish F_j at time $t - \tau$. The values of $H_{ij}(t, \tau)$ are between −1 (when fish swim in opposite directions) and 1 (when fish have the same direction), and equals zero when fish have perpendicular directions.

By averaging $H_{ij}(t, \tau)$ along a time-window of length $(2w + 1)\Delta t$, we are able to quantify how much the focal fish F_i is copying the moving direction of its neighbor with a time-delay τ by means of the following function [26]

$$C_{ij}(t, \tau, w) = \frac{1}{2w + 1} \sum_{k=-w}^{w} H_{ij}(t + t_k, \tau), \tag{4}$$

where $t_k = k\Delta t$ (the time-step in our experiments is $\Delta t = 0.02$s). The time-window parameter length w has been determined by means of a sensitivity analysis (pairwise similarity matrix), finding that $w = 2$ yields the more satisfactory results; see Section "Parameter selection" in Materials and methods and S5 Fig.

The average directional correlation $C_{ij}(t, \tau, w)$ allows us to characterize a fish F_j as an influential neighbor of a focal fish F_i at time t with time-delay τ, if the value of $C_{ij}(t, \tau, w)$ is larger than a given threshold $C_{\min}$. Details on how w and $C_{\min}$ are obtained are given in Sections "Optimal setting parameters for influential neighbors identification" and "Parameter selection" in Material and Methods.

Fig 5 shows the directional correlation H_{12} and its time-average C_{12} between fish F_1 and F_2 along the collective U-turn depicted in Fig 3. Left (resp. right) panels aim to indicate the alignment of fish F_1 (resp. F_2) at each time t with respect to the alignment of fish F_2 (resp. F_1) at an earlier time $t - \tau$. Panels A and C show respectively that for all τ, there is always an interval of time during which $H_{12}(t, \tau) \approx -1$ and $C_{12}(t, \tau) \approx -1$ (dark region), meaning that for all time-delays there is always an interval of time in which fish have opposite directions. Moreover, the larger the time-delay, the wider the black region where the direction of F_1 is opposite to the direction of F_2 at the previous time.

On the other hand, the figures of the directional correlation of F_2 with F_1, especially Panel D, show a connected region in which the correlation $C_{21}(t, \tau)$ remains positive and above the threshold (yellow in the figure) around $\tau \approx 0.42$ s where $H_{21} \approx 1$ during all the time interval [−0.5, 2 s]. This strongly suggests that, during this time interval, F_2 is copying the behavior of F_1 with a 0.42 s time-delay, denoted $\tau_{2,1}$ for this specific U-turn. Thus, one can consider that F_1 is influencing F_2 with time-delay $\tau_{2,1}$, while F_2 is not influencing F_1 in this specific case. This influence dynamics is illustrated in Fig 3D by drawing an arrow at time t from F_j to F_i when F_j satisfies the condition $C_{ij}(t, \tau, w) > C_{\min}$ for being an influential neighbor of F_i at time t, which in turn receives this influence and responds by copying the exhibited heading with a time-delay τ.

Using the same procedure for the $N = 5$ case depicted in Fig 4, we draw Fig 6 that shows F_1 copying F_2 with a time-delay $\tau_{1,2} \approx 0.5$ s (Panels A and E). F_1 also copies F_3 and F_5 with, respectively, $\tau_{1,3} \approx 0.2$ s (Panels B and F) and $\tau_{1,5} \approx 0.1$ s (Panels D and H), but it doesn't copy

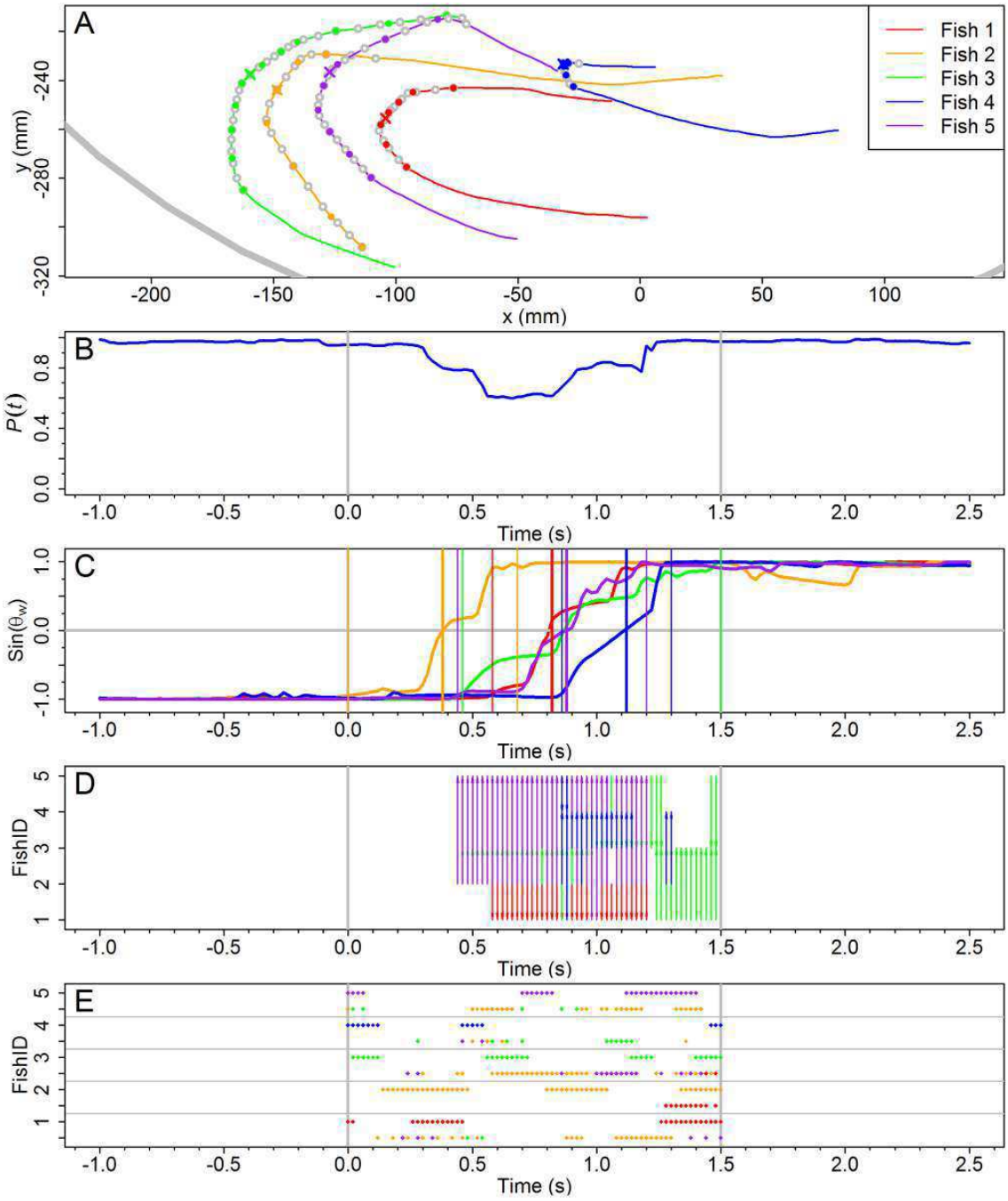


Fig 4. Spatial and temporal dynamics of a collective U-turn for $N = 5$. The displayed temporal sequence is drawn from the fish trajectories one second before the U-turn begins till one second after its end. Symbols in all panels are the same as in Fig 3. (A) Individuals trajectories in the tank during the U-turn. (B) Group polarization with a minimum value $P_{min} \approx 0.59$ reached at $t \approx 0.66$ s. (C) Sine of the angle of incidence of fish to the wall θ_w. The three vertical lines of each color indicate for each fish the beginning, the middle and the end of its U-turn. Here the middle time means the instant where $\sin(\theta_w) = 0$. (D) Interactions with influential neighbors: arrows point from influential neighbors to the focal

fish and with the same color as the focal fish. (E) Fish bursting activity and their influential neighbors. If there is more than one influential neighbor, F_j with largest index value j is shown. Grey lines in Panels BCDE denote the start and end of the collective U-turn.

https://doi.org/10.1371/journal.pcbi.1005822.g004

F_4 (Panels C and G). The influential neighbors of F_1 are thus F_2, F_3 and F_5, at different times and with different time-delays. We have calculated the rest of the correlations for all pairs of fish (see S1 Fig for an overview of all the heading correlations). As for the $N = 2$ case, these relations are illustrated by arrows going from the influential neighbors to the reacting fish in Fig 4D.

Effect of bursting on the transmission of information

The specific behavior of *H. rhodostomus*, namely, the successive alternation of bursts and coasts [15], leads us to ask whether these abrupt changes of acceleration and speed can provide information that other fish could use to adjust their own movement. To address this aspect we study whether there is any correlation between the bursting activity of one fish at time t and the fact that this fish is an influential neighbor of another fish shortly after time t.

A burst corresponds to a brief phase of acceleration during which most changes in fish heading occur [15]. Panels E in Figs 3 and 4 show the bursting activity of each fish F_i, $i = 1, \ldots, N$, and that of its influential neighbors. For each fish F_i, we draw a dot at time t and ordinate i if fish F_i is displaying a burst precisely at time t. Dot color at ordinate i corresponds to fish F_i's color. The absence of a dot at a given time denotes that the fish is in a coasting phase at that time.

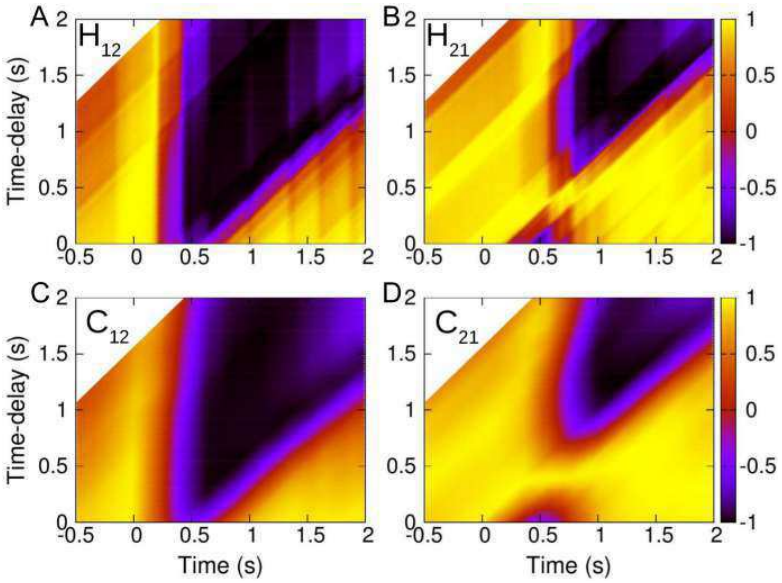


Fig 5. Directional correlations between fish F_1 and F_2. (A) Directional correlations $H_{12}(t, \tau)$ and (B) $H_{21}(t, \tau)$ for $t \in [-0.5, 2]$ and $\tau \in [0, 2]$ and their corresponding average over a time-window of width $2w = 0.4$ s: (C) $C_{12}(t, \tau, w)$, (D) $C_{21}(t, \tau, w)$. Yellow regions: $H_{ij} \approx 1$ and $C_{ij} \approx 1$, *i.e.*, fish have the same direction with a time-delay τ. Dark regions: $H_{ij} \approx -1$ and $C_{ij} \approx -1$, *i.e.*, fish have opposite directions. The white upper-left corners indicate the τ is larger than the minimal time considered in this data set.

https://doi.org/10.1371/journal.pcbi.1005822.g005

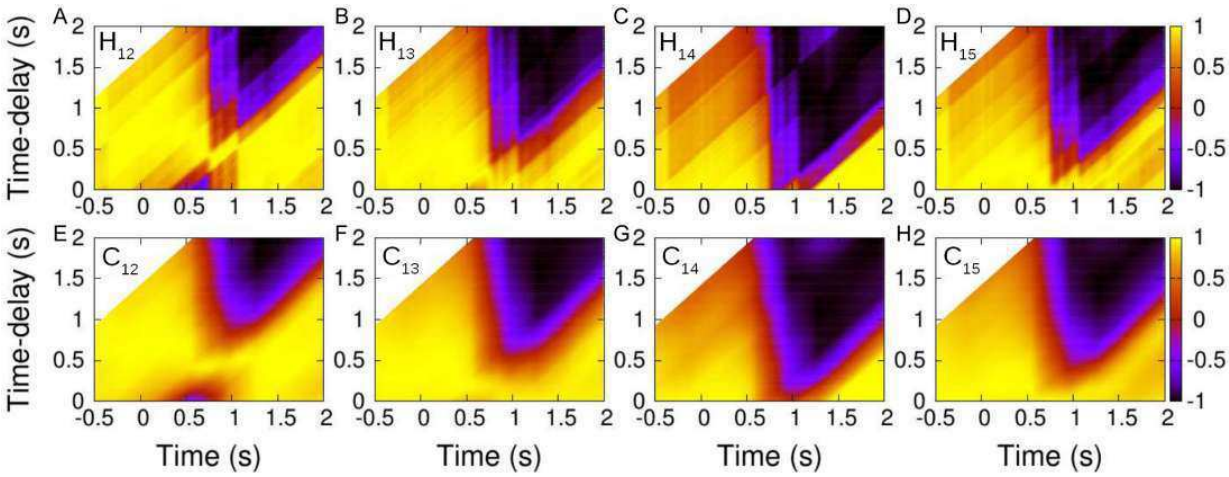


Fig 6. Directional correlation of fish F_1 with the other fish F_j, $j = 2, \ldots, 5$. Directional correlations $H_{1j}(t, \tau)$ (panels ABCD) for $t \in [-0.5, 2]$ and $\tau \in [0, 2]$, and their corresponding average $C_{1j}(t, \tau)$ (panels EFGH) over a time-window of width $2w = 0.4$ s. Yellow regions: $H_{ij} \approx 1$ and $C_{ij} \approx 1$ (fish have the same direction with a time-delay τ). Dark regions: $H_{ij} \approx -1$ and $C_{ij} \approx -1$ (fish have opposite directions). In the upper-left corners the white color indicates that τ is larger than the minimal time considered in this data set.

https://doi.org/10.1371/journal.pcbi.1005822.g006

A second row of colored dots is drawn at ordinate $i - 0.5$ for some values of t when two conditions are met: (1) Fish F_i is being influenced at those times by one or more fish F_j, $j \in \{1, \ldots, N\}$, $j \neq i$, whose identity is given by the color of the dots, and (2) the influential fish F_j was bursting when it was influencing F_i at time $t - \tau$ earlier. If F_i has more than one influential neighbor at time t, the dot drawn at time t in row $i - 0.5$ has the color of the F_j fish with the highest index j.

In Fig 3E, red dots at $i = 1$ mean that fish F_1 is bursting at those time-steps and coasting at the other time-steps, and red dots at $i - 0.5 = 1.5$ indicate that, first, F_1 is the influential fish of F_2 at those time-steps, and second, F_1 was bursting when it was earlier influencing F_2. In turn, there are two possible reasons to explain the absence of red dots at $i - 0.5 = 1.5$ for certain time values: either F_2 has no influential neighbor, or F_1 was coasting. To assess which of the two explanations is valid, one needs to look at Fig 3D. For example, the absence of dots at $i - 0.5 = 1.5$ during 0.57 s and 0.62 s is due to F_2 having no influential neighbors, while the absence of dots in the same row between 0.75 s and 0.81 s results from the fact that F_1, which is the influential neighbor of F_2, is in a coasting phase at time $t - \tau$ (in this example the delay was found to be $\tau = 0.42$ s).

Fig 3E shows that the bursting activities of both the focal fish and its influential neighbor are not directly correlated, suggesting that the primary source of information for fish to adjust their movements is the distance, orientation and angular position of their neighbors [15]. The same conclusion is obtained for $N = 5$. By focusing on fish F_2 for example, Fig 4E shows that there is no systematic overlap between the yellow dots at $i = 2$ and those at $i - 0.5$ for $i \neq 2$, suggesting that the correlation between the bursting activity of a fish and that of their influential neighbors is marginal.

Number of influential neighbors

For all U-turns, we have counted the number of frames in which a fish is an influential neighbor, that is, the number of frames where the above described condition for identifying influential

neighbors is met. When there are only two fish, a fish is found to be the influential neighbor 30% of the time spent in a U-turn. In groups of five fish, this proportion grows up to 62%.

We have counted the number of influential neighbors N_{if} a fish F_i has during a U-turn in groups of five fish, finding that in most cases, a fish has only one or two influential neighbors (for 58% of the time spent in a U-turn $N_{if} = 1$ or 2); see Fig 7A. The most frequent case is $N_{if} = 1$ (43%). Having more than one influential neighbor is frequent (19%), but less than

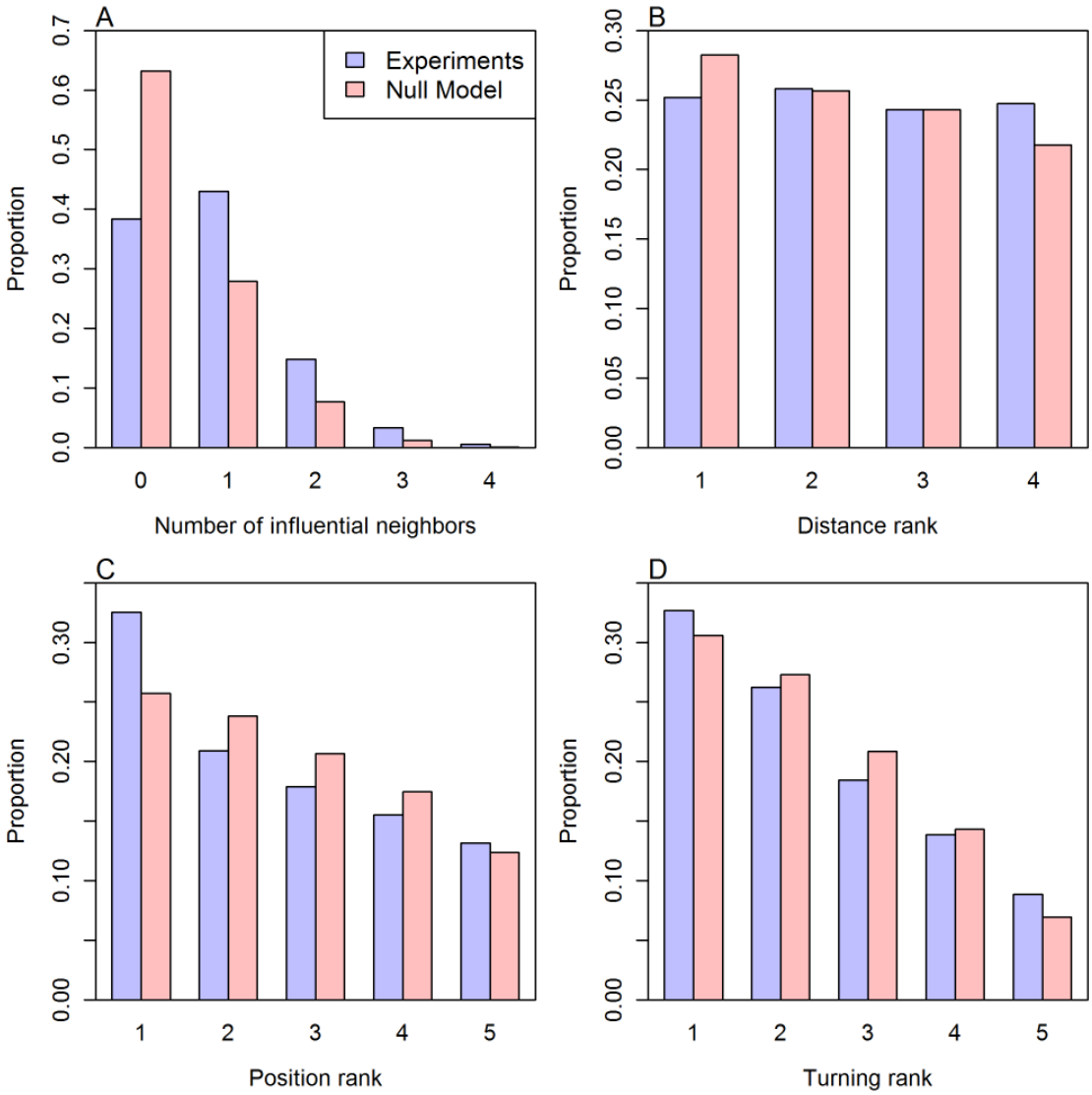


Fig 7. Number, location and temporal occurrence of influential neighbors. Cumulative analysis of collective U-turns of over 475 experimental (blue) and 1000 artificial (red) observations in groups of $N = 5$ fish. (A) Number of influential neighbors.(B) Distance rank of influential neighbors with respect to the focal fish. (C) Position rank of influential neighbors in the group. (D) Turning rank of influential neighbors. Histograms represent the proportion of time during which influential neighbors have been observed in a given class. The procedure to construct the artificial observations is presented later and in the section "Null model" in Materials and Methods.

https://doi.org/10.1371/journal.pcbi.1005822.g007

having no influential neighbors (38%). The cases where there are more than two influential neighbors are negligible (less than 4% of the total time spent in U-turns).

For each fish F_i, we have calculated the respective distance $d_{ij}(t)$ at which the other $N-1$ fish F_j are from F_i during the U-turns, thus establishing a rank order among the neighbors influenced by F_i. We have then compared the influence of close neighbors with those of distant neighbors, finding no correlation between the distance rank of a neighbor and the influence it exerts on the focal fish. This is shown in Fig 7B, where we have depicted the distribution of the distance rank of influential neighbors with respect to a focal fish. The figure shows that fish spent the same proportion of time ($\approx 25\%$) being an influential neighbor of a focal fish independently of their distance rank. In other words, influential neighbors are not necessarily the closest ones.

When trying to identify events of causal influence by means of correlations, it is crucial to keep in mind that correlation does not imply causation. We thus have controlled the effects of potential chains of influence, where *e.g.* fish F_1 is highly correlated with F_3 not because F_1 is directly influencing F_3, but because F_1 is influencing fish F_2, which in turn is influencing F_3. To check the impact of these chains of influence on our results, we have removed from our data all the pairwise influence data that correspond to the following situation: if F_1 is influenced by both F_2 and F_3 and F_2 is simultaneously influenced by F_3 (or F_3 is influenced by F_2), then we removed the pairwise correlation (focal fish, influential neighbor) corresponding to (F_1, F_2) (or (F_1, F_3)). After removing 7172 out of 69703 data points and recomputing the results with the remaining data, we found that our results remain practically unchanged.

We have also calculated the position rank that each fish occupies in the group during a collective U-turn, finding that influential neighbors are mostly located in the front region of the group: 32% in the leading most advanced position, and 20% in the second place; see Fig 7C. Noticeably, influential neighbors can be found in the back of the group (in 29% of the cases in the fourth or fifth position), and even in the last position (a non-negligible 13% of cases).

We also paid attention to the order in which each fish starts its individual U-turn during a collective U-turn, finding that influential neighbors are those that most frequently turn earlier (32% of the cases), and that this relation decreases linearly; see Fig 7D. It is again noticeable that influential fish can be found to be the last turning fish (in 8% of the cases).

The apparently surprising fact that influential fish can be found in the back of the group and that the last fish turning can be an influential fish is due to the anisotropic perception of fish and their relative orientations during U-turns. But these findings have to be understood in the light of our specific time-dependent characterization of influential neighbor. If, for instance, F_1 turns first and influences F_2, F_2 will turn with some time-delay after F_1. Then, when F_2 is at half of its individual turning process, F_2 can be rotating in the same direction as F_1 in such a way that F_1, influenced by F_2, slightly adjusts its direction. We would then say that F_2, which is the last turning fish, has influenced F_1, the first turning fish.

In order to compare different collective U-turns, we define a normalized time $\bar{t} = (t - t_s)/(t_e - t_s)$ in terms of the actual time t and the starting and ending time of each U-turn, so that the duration of a U-turn is now $\bar{t} = 1$. Thus, $\bar{t} = -1$ corresponds to a time as long as the U-turn duration previous to the start of the U-turn, and $\bar{t} = 2$ corresponds to a time as long as the U-turn duration after the end of the U-turn. We have calculated the instantaneous value of the average speed $\mathcal{V}(t) = \langle \| \vec{v}(t) \| \rangle$, the average group polarization $\mathcal{P}(t) = \langle P(t) \rangle$ and the average number of influential neighbors $\mathcal{N}(t) = \langle N_{\mathrm{if}}(t) \rangle$. Here, angle brackets refer to the average across all fish in the U-turn along a time-window containing the collective U-turn.

Fig 8A and 8B show respectively the time evolution of $\mathcal{V}(t)$ and $\mathcal{P}(t)$ during the collective U-turns in groups of 5 fish. The description of the specific U-turn presented in Fig 4 is also

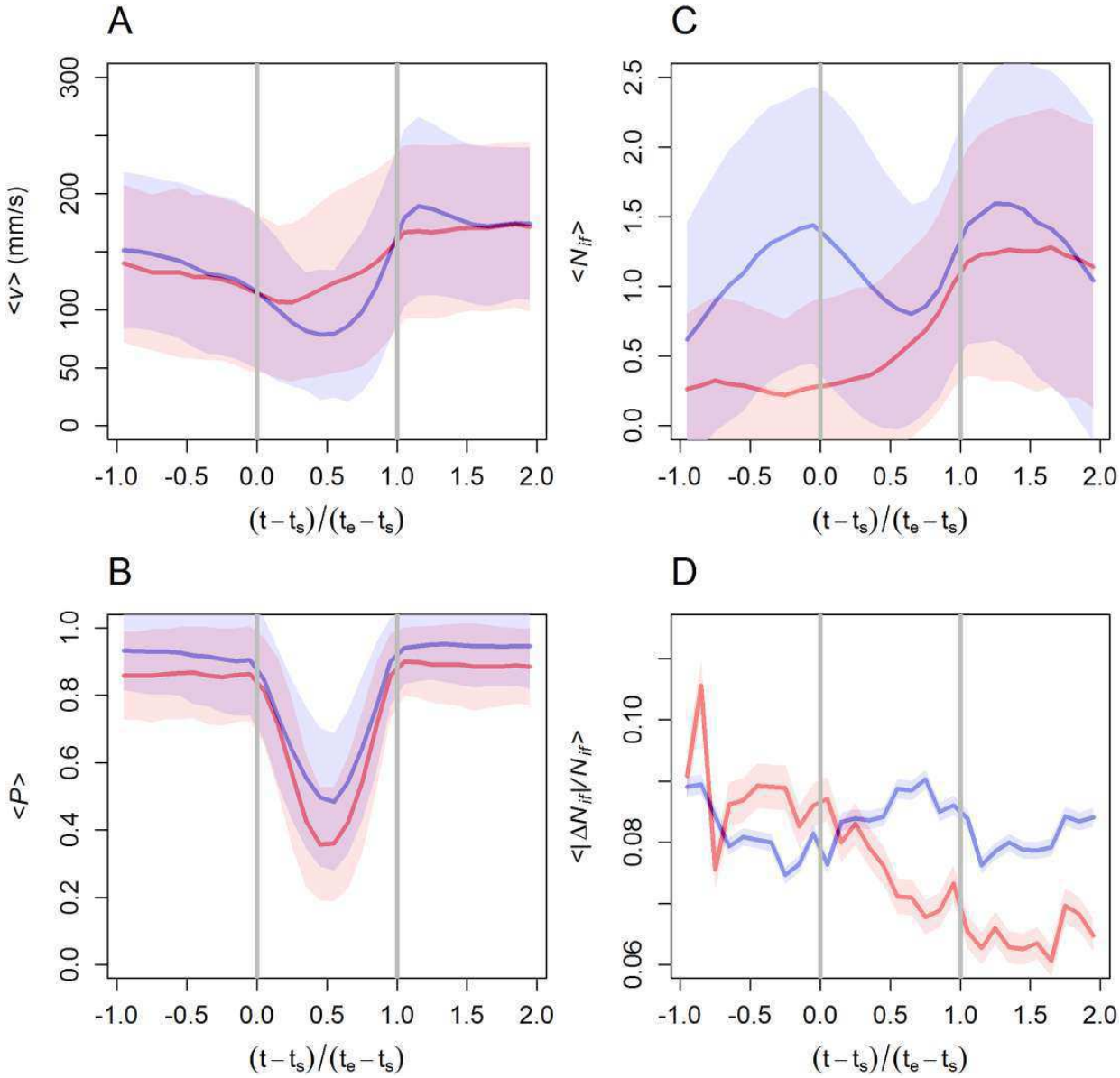


Fig 8. U-turn dynamics in groups of $N = 5$ fish. We depict here the temporal dynamics for the average velocity, average polarization, number of influential neighbors and its variation in over 475 experimental (blue) and 1000 artificial (red) recordings of collective U-turns. (A) Average speed $\mathcal{V}(t)$. (B) Average group polarization $\mathcal{P}(t)$. (C) Average number of influential neighbor $\mathcal{N}(t)$ per focal fish. (D) Average of the absolute variation in the number of influential neighbors $|\Delta N_{if}|$ divided by the number of influential neighbors N_{if}, defined in Eq (5): $\langle \eta(\bar{t}) \rangle$. Horizontal axis denotes normalized time $\bar{t}$, where t_s and t_e denote the start and end of the collective U-turn respectively. The procedure to construct the artificial observations is presented later and in the section "Null model" in Materials and Methods.

https://doi.org/10.1371/journal.pcbi.1005822.g008

valid for the general case: the speed decreases before the U-turn (from $\mathcal{V}(-1) \approx 150$ mm/s to $\mathcal{V}(0) \approx 115$ mm/s), it reaches a minimum at half the U-turn duration $\bar{t} = 0.5$ ($\mathcal{V}(0.5) \approx 70$ mm/s), and it then grows to a higher value after the U-turn ($\mathcal{V}(1.5) \approx 165$ mm/s). A very similar behavior was found in groups of 2, 4, 8 and 10 fish of the same species in [28]. At the same time, the polarization is very high and almost constant outside the U-turn ($\mathcal{P}(\bar{t}) \approx 0.95$), and exhibits a perfect V-shape during the U-turn, with the high values ($\mathcal{P}(\bar{t} = \{0, 1\}) \approx 0.93$) reached at exactly the instants where the start and end of the U-turn takes place $\bar{t} = 0$ and $\bar{t} = 1$, and the minimum value ($\mathcal{P}(0.5) \approx 0.48$) at the middle of the U-turn. As expected, the average group polarization $\mathcal{P}(\bar{t})$ significantly decreases during the U-turn to almost half the value it has outside the U-turn. Right after reaching this minimum, there is a sharp increase of speed and polarization as more fish adopt the new direction of motion.

Fig 8C shows that before the U-turn the average number of influential neighbors $\mathcal{N}(t)$ increases until a maximum value is reached right before the start of the U-turn ($\mathcal{N}(-0.1) \approx 1.45$). During more than one half of the U-turn, $\mathcal{N}(t)$ decreases until a minimum ($\mathcal{N}(0.6) \approx 0.8$), and grows again beyond the end of the U-turn until a second maximum ($\mathcal{N}(1.2) \approx 1.6$, twice the height of the minimum). After that, all fish have completed their U-turns and $\mathcal{N}(t)$ decreases again.

When the polarization is very high, the time-delay with which influential neighbors are detected is often too small in comparison with biologically realistic reaction times τ_R, so that these influential neighbors are not taken into account (we used $\tau_R = 0.04$ s; see Section "Optimal setting parameters for influential neighbors identification" in Materials and methods). This is the reason why the average number of influential neighbors $\mathcal{N}(t)$ appears to be smaller in regions outside the U-turn, than when the U-turn is just about to start ($\bar{t} \approx -0.1$) or slightly after its end ($\bar{t} \approx 1.2$). Meanwhile, the decrease of $\mathcal{N}(t)$ in the middle of the U-turn has a different origin: once a fish has started to turn around, there is no real need of updating its alignment according to all its neighbors. That fish can safely reverse its motion by keeping the alignment with only one of those neighbors and even not paying attention to them for some period of time.

Another indicator of how fish make decisions while turning is how frequently a focal fish pays attention to other individuals. We define the relative variation of the number of influential neighbor per fish $N_{\mathrm{if}}(t)$ between two successive time-steps as follows:

$$\eta(t) = \frac{|N_{\mathrm{if}}(t + \Delta t) - N_{\mathrm{if}}(t)|}{N_{\mathrm{if}}(t)}, \tag{5}$$

denoting by Δt the time-step between frames ($\Delta t = 0.02$ s).

We have depicted the time-evolution of the average $\langle \eta(t) \rangle$ in Fig 8D, finding that $\langle \eta(t) \rangle$ remains essentially constant before, during and after the U-turn event, the amplitude of its variation being smaller than 10% of the signal (0.007 and 0.08, respectively).

Since the average number of influential neighbors $\mathcal{N}(t)$ is smaller when fish are engaged in the U-turn than right before or right after the U-turn, a constant average $\langle \eta(t) \rangle$ suggests that fish adjust their heading more frequently during the U-turn than outside the U-turn. Indeed, in the middle of a U-turn, no real common direction of motion exists ($\mathcal{P}(t) \approx 0.5$), that is, there is a high diversity of headings, so that fish have to frequently update their direction by paying attention to different neighbors.

Spatial organization of influential neighbors

We are now interested in determining the dynamical spatial organization of the influential neighbors of a focal fish. The relative state of a fish F_j with respect to a focal fish F_i is

characterized by several parameters: the relative position of the neighbor $\vec{u}_{ij} = \vec{u}_j - \vec{u}_i$, where $\vec{u}_i$ is the vector position of F_i in cartesian coordinates, the distance between them $d_{ij} = \|\vec{u}_{ij}\|$, the viewing angle of F_j relative to the direction of F_i [26], which is the angle θ_{ij} with which F_i perceives F_j (note that θ_{ij} is not necessarily equal to θ_{ji}), the relative velocity $\vec{v}_{ij} = \vec{v}_j - \vec{v}_i$, and the relative heading $\phi_{ij} = \phi_j - \phi_i$. All these quantities are time-dependent. We have calculated their average value for all the U-turns in a uniform spatial grid of square cells to facilitate the interpretation of the vector field of these continuous variables. Each square cell, of side 20 mm, shows the average of the arbitrarily different number of values contained in the cell.

Fig 9A shows the density map of the relative position of the influential neighbor with respect to the focal fish when $N = 2$. The intensity of color is proportional to the frequency of occupation of the grid cell, showing that the influential neighbor is mostly located in front of the focal fish and at a distance of one to three body lengths from the focal fish. The same information is quantified in Panel B with a heat map in polar coordinates, highlighting the most frequent location of the influential neighbor.

The average relative velocity $\langle \vec{v}_{ij} \rangle$ is shown in Fig 9A (arrows), superimposed to the density map. The vector field shows that when the influential neighbor is in front of or behind the focal fish ($\sin\langle\theta_{ij}\rangle \approx 0$), both fish move at similar speed although the focal fish is a little bit faster (the small black arrows are pointing in the opposite direction to the red one) and the difference in heading is also small. However, when the influential neighbor is on the sides of the focal fish, relative speed and heading difference tend to vary more as the distance between them increases.

The distributions of distances d_{ij} and exposure angles θ_{ij} between a focal fish and its neighbors are depicted in Panels C and D of Fig 9 respectively. We find, on the one hand, that their most frequent separation is 62.6 mm ± 29.7 mm (mean and standard deviation of histogram in Fig 9C), a value that is consistent with previous results where it was shown that the behavioral reactions of a fish depend on the angular position of its neighbors, as a consequence of the anisotropic perception of the environment [15].

On the other hand, the distribution of the exposure angle of fish F_j to the focal fish F_i is narrower when F_j is influencing F_i than when F_j is a neighbor of F_i, not necessarily influencing F_i. As both distributions are centered on $\theta_{ij} = 0$, this shows that F_j is more frequently located in front of F_i when F_j is an influential neighbor of F_i than in the case when F_j is just a neighbor of F_i.

Fig 10 shows similar results for groups of $N = 5$ fish. Influential neighbors are more frequently located in front of the focal fish (although with a slight shift to the right; see Panels A and B) and at a mean distance of 67.5 mm ± 40.6 mm (Panel C).

In turn, the velocity field has a smaller intensity and is much more homogeneous than in the case where $N = 2$. A slight asymmetry can also be observed (not noticed when $N = 2$) with fish located in front and slightly to the right of the focal fish having a higher velocity than those located elsewhere. Moreover, the distribution of exposure angles is more dispersed than in the case of two fish, meaning that influential neighbors are exposed to the focal fish with a larger diversity of angles, something that is simply due to the higher number of fish.

The difference in the homogeneity of the velocity field between groups of 5 and 2 individuals is not necessarily the result of averaging over a larger number of individuals. Although averaging over fish data pairs may reduce the uncertainty in the extracted parameter values, it is well-known that the level of homogeneity in the direction of motion of the school increases with group size [29]. But one also ought to consider that specific values of delay and curvature the individuals adopt during the U-turns could help to limit variability in coordinating the group. Some theoretical studies support this idea: simplified models of velocity alignment with

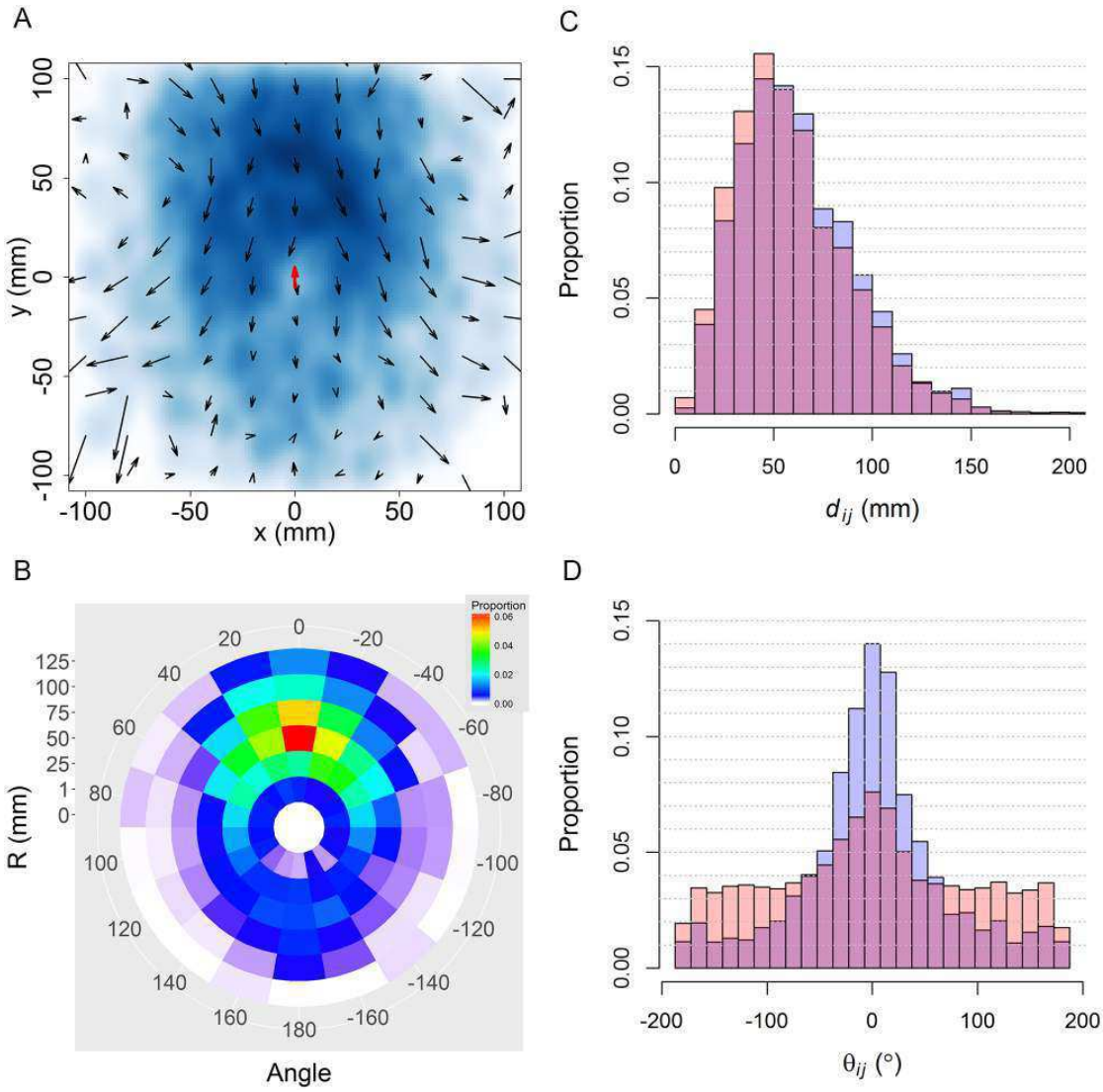


Fig 9. Spatial and velocity distributions of influential neighbors around a focal fish in groups of 2 individuals. (A) Density map of influential neighbors' location (blue) and their average relative velocity field (arrows) with respect to the focal fish (red arrow). (B) Average spatial distribution of influential neighbors in polar coordinates (red: highest frequency; dark blue: low frequency; white: frequency equals zero). (C) and (D): Distributions of the distance d_{ij} and the angle of exposure θ_{ij} respectively. Blue histograms: F_j is an influential neighbor of F_i; orange: F_j is a neighbor of F_i, not necessarily influencing F_i; dark pink: overlap between the two.

https://doi.org/10.1371/journal.pcbi.1005822.g009

additive noise have shown semi-analytically the existence of delay and rate of turn values that minimise the fluctuations in the variance of the individual speed [30], and flocking models of self-propelled particles have also shown that delay can be tuned to increase stability and alignment of the group [31].

Finally, we have analyzed the variation of the time-delay τ as a function of both the distance between the focal fish and its influential neighbors d_{ij} and the difference of heading ϕ_{ij}, finding

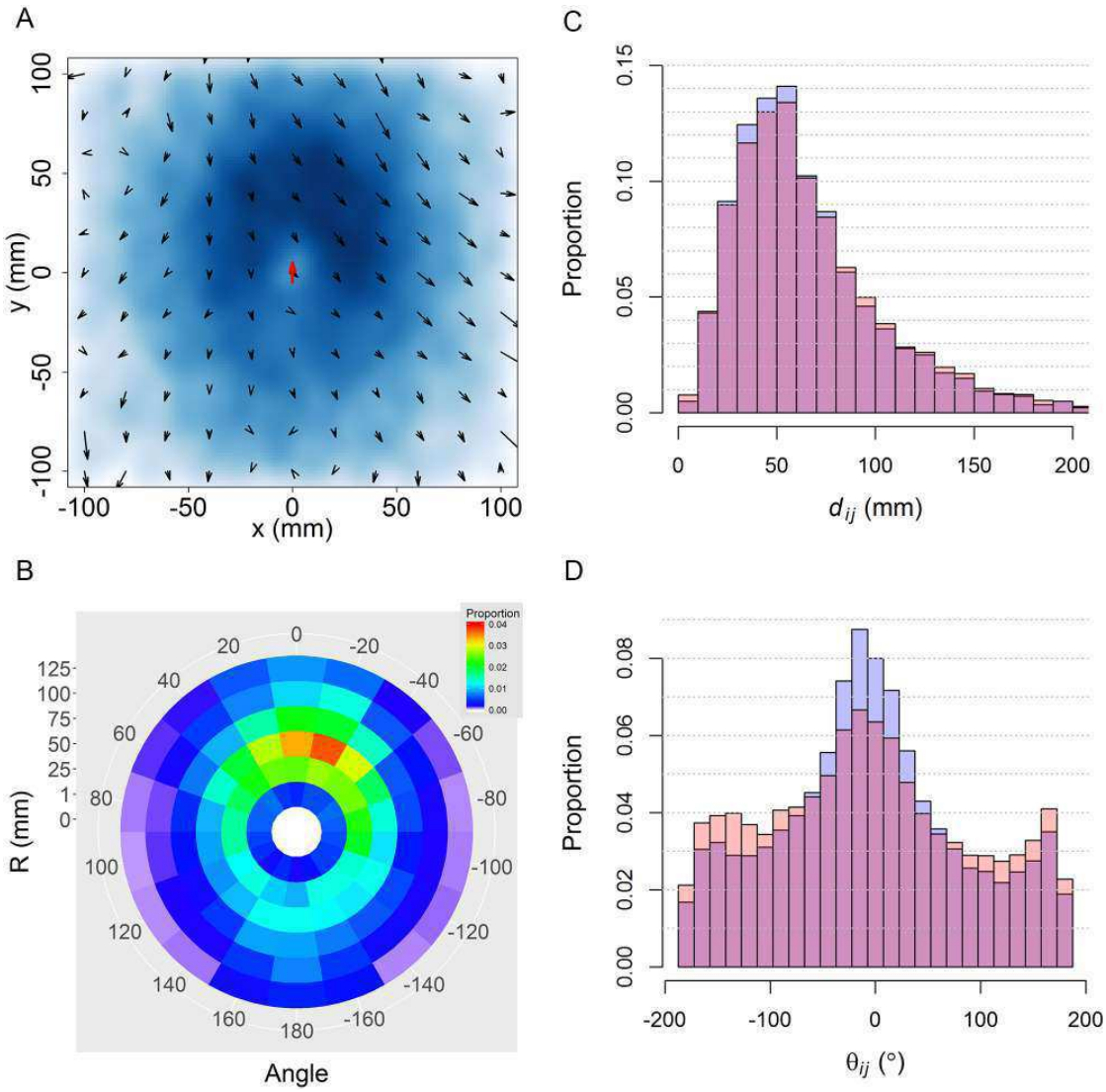


Fig 10. Spatial and velocity distribution of influential neighbors around a focal fish in groups of 5 individuals. (A) Density map of influential neighbors' location (blue) and their average relative velocity field (arrows) with respect to the focal fish (red arrow). (B) Average spatial distribution of influential neighbors in polar coordinates (red: highest frequency; dark blue: low frequency; white: frequency equals zero). (C) and (D): Distributions of the distance d_{ij} and the angle of exposure θ_{ij} respectively. Blue histograms: F_j is an influential neighbor of F_i; orange: F_j is a neighbor of F_i, not necessarily influencing F_i; dark pink: overlap between the two.

https://doi.org/10.1371/journal.pcbi.1005822.g010

that in both cases $N = 2$ and $N = 5$, the time-delay increases with respect to both the distance d_{ij} and the heading difference ϕ_{ij} (see Fig 11). This result can be understood because during a U-turn the fish speed is decreasing and two fish can display larger reaction times the more separated they are and the less aligned they are.

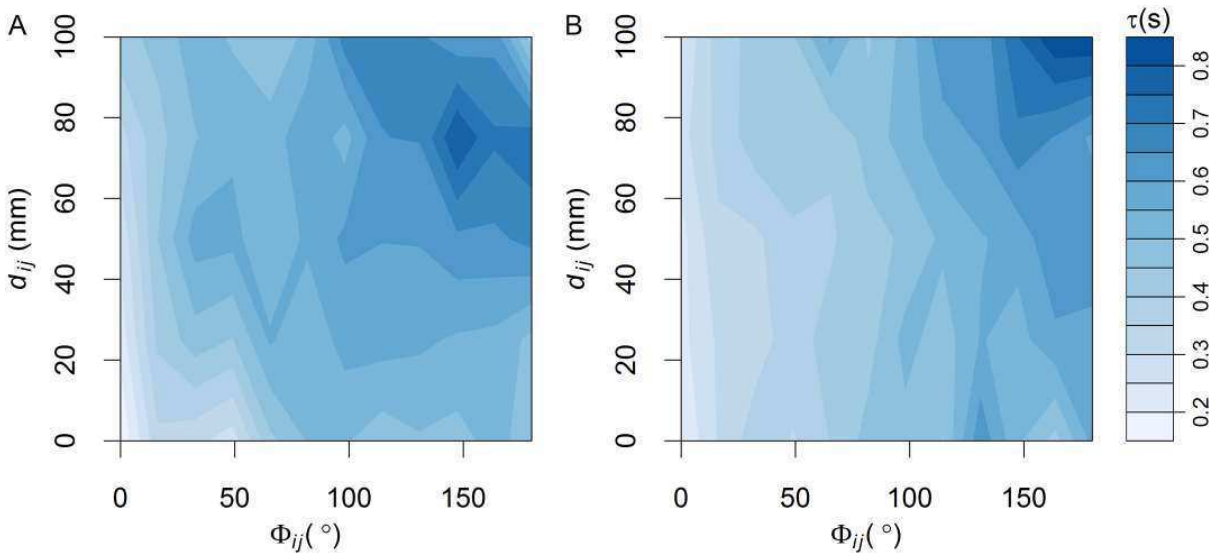


Fig 11. Time-delay dependence on heading difference and separation distance. Time delay τ extracted from the empirical observations as a function of heading difference ϕ_{ij} and separation distance d_{ij}. (A) $N = 2$, (B) $N = 5$. In both cases, the larger the heading difference and the distance, the longer the time-delay.

https://doi.org/10.1371/journal.pcbi.1005822.g011

A null model to detect spurious correlations

As already mentioned in the introduction, establishing causal influence on the basis of correlation measures requires controlling for spurious effects. Although our experimental data correspond to a specific collective behavior in which individuals influence each other, the relatively short time-windows over which cross-correlation are averaged and the use of several parameters through sensitivity analysis can weaken the accuracy of our results. To demonstrate that the particular detections of influential neighbors are not purely due to chance, we generated random artificial U-turns events by bootstrapping the data and applying the same procedure used to analyze collective U-turns in our experiments.

The null model is built for groups of 5 fish, for which our experimental data provide $M = 2375$ individual trajectories (5×475 collective U-turns). For every fish F_i, $i = 1, \ldots, M$, the trajectory is rotated so that the individual turning point of the fish (where $\sin(\theta_{wi}) = 0$) is located in the upper part of the tank, by randomly sampling the new angular position ψ_i in the interval $[\pi/2 - \xi, \pi/2 + \xi]$, where ξ is a small angle (we used $\xi = \pi/12$). Similarly, the time scale of each fish is shifted by sampling the instant of turning in the time interval $[-\zeta, \zeta]$, where ζ is a short time (we have used $\zeta = 1$ s). Then, five trajectories are randomly sampled, each one from a different randomly sampled collective U-turn, and mirrored if necessary so that the five individual U-turns are done in the same direction, clockwise or anti-clockwise. This way, the five fish of the artificial U-turn make their individual U-turn approximately at the same place and approximately the same time. For more details, see the section "Null model" in Materials and methods.

We have produced 1000 artificial collective U-turns; S9 Fig shows a collection of 10 of them. The results of our analysis are shown in red in Figs 7 and 8. As expected, they reveal clear differences between artificial and experimental U-turns.

Fig 7A shows that in artificial U-turns the proportion of time during which a focal fish has no influential neighbor is more than 63% of the time, while in the experiments it was less than 39%. The analysis also reveals that in artificial U-turns a focal fish has one influential neighbor for less than 28% of the time, while in the experiments, the proportion raises to 43%. Similarly, Fig 8C shows that the average number of influential neighbors $\mathcal{N}(t) = \langle N_{\mathrm{if}}(t) \rangle$ is much smaller in artificial U-turns (≈ 0.4) than in real U-turns, where $\mathcal{N}(t)$ is almost always greater than 1. Note that the increase of $\mathcal{N}(t)$ during U-turns in artificial data is the consequence of the channeled motion of fish by the corridor. Moreover, the variation of $\mathcal{N}(t)$ along time, including the transients preceding and following the U-turn, decreases in artificial U-turns while it remains constant and with a higher value in experiments.

Fig 7B shows that distance rank has no significant effect on which fish is the influential one, both in experiments and in artificial U-turns. The decreasing number of influential neighbors comes from the fact that the tank is circular and the method we use. If the tunnel had been a straight corridor, we should have detected no decrease in our null model. However, in a circular tank, because of the geometrical constraints imposed by the curvature, even when two fish are both swimming in the same direction (*i.e.*, clockwise or anti-clockwise), as the distance between fish increases, our method will detect a decrease of correlation. While Fig 7C confirms that influential neighbors are slightly more often ranked in the first position of the group, this effect is much more pronounced in the experiments. In fact, Figs 7B, 7C and 7D and 8A and 8B show that the selected null model satisfactorily reproduces the typical spatiotemporal behavioral patterns of real U-turns: the position and turning ranks are almost identical, as well as the variation of the average speed and the average group polarization, although the V-shape of the average polarization in real U-turns is significantly sharper than in artificial U-turns.

An additional, albeit expected, result of our null model is the homogeneous (isotropic) spatial distribution of "influential neighbors", while in real collective U-turns influential neighbors are mostly located in front of the focal fish; see S10A and S10B Fig, compared with Fig 10A and 10B.

Discussion

By sharing information with other group members, schooling fish and other collectively moving animals can potentially improve their navigational accuracy (*e.g.* the many wrongs principle [32]), take better decisions (*e.g.* to avoid a predator [33]), or improve their abilities to sense the environment [34]. However, there are both physical and practical reasons why information is expected to be shared with only a few neighbors. Physical reasons involve material limitations, such as visual occlusions. Practical reasons often refer to trade-offs between sharing information, so that the group collectively selects a direction of motion, and deciding independently [35, 36].

Assuming that correlations between fish behavior rely to some extent on a causal influence, our analysis reveal that in groups of *H. rhodostomus*, during a collective U-turn, at any moment in time each fish only pays attention to a small number of neighbors whose identity regularly changes. We also find that the phases during which a focal fish is affected by one or two influential neighbors are interspersed with other phases during which its movement appears uninfluenced by the movement of neighbors. Moreover, influential fish are mostly located in front of the focal fish. The distance between a focal fish and its influential neighbors is about two body-lengths and the relative exposure angle is smaller than 60 degrees.

Our results bring insights on the way information on the neighborhood is processed by fish. Instead of having a synchronous update based on a fixed number of neighbors (topological neighborhood) or on all neighbors located within a fixed distance (metric neighborhood),

our results suggest an asynchronous updating that does not depend on the distance between a focal fish and its influential neighbors. A similar asynchronous updating scheme has been previously introduced by Bode et al. [37] in a flocking model showing that it can give rise to emergent topological interactions consistent with the measures done on starling flocks [38].

It is however worth noting that our experiments, performed on small group sizes, may have prevented us from detecting any influence of the distance, since each of the four neighbors are located between one and three body lengths. In larger groups of fish moving in an unconstrained space, we expect the effective neighborhood of fish to result from the interplay between an asynchronous updating on a small number of neighbors and a modulation of the strength of interactions with the distance between fish [15].

Previous studies on the number and the spatial arrangement of influential neighbors led to different results depending on the species and on the procedure used to analyse the data. The work by Ballerini *et al.* [39] provides evidence that each bird within a starling flock (*Sturnus vulgaris*) coordinates its motion with a fixed number of closest neighbors, irrespective of their distance, while in mosquitofish (*Gambusia holbrooki*), one single nearest neighbor was sufficient to account for the large majority of the observed interaction responses [12]. In barred flagtails (*Kuhlia mugil*), it has been shown that different kinds of neighborhoods (Voronoi neighborhood and the k nearest neighbors ($k \approx 6 \sim 8$) were compatible with experimental data in a tank [13]. Our study points to a low number of influential neighbors. There are multiple possible explanations for the differences in the number of interacting neighbors found across the scientific literature. (*i*) It is possible that different animal groups interact with different numbers of neighbors. (*ii*) Temporal factors are also important [37], as interactions can be integrated in time to produce effectively larger neighborhoods. Here, we propose a third explanation (*iii*) based on the consideration that interaction responses such as attraction, alignment and avoidance are qualitatively different mechanisms that rely on different sensory-motor responses and, consequently, on different interacting neighborhoods. In particular, attraction and repulsion require to process information about the position of neighbors, while alignment is intrinsically a response dependent on orientation and velocity. These different interactions are likely to rely on different neural circuits (motion and form are typically processed by different brain areas in many animal groups [40, 41]) and hence might depend on different sets of influential neighbors: for instance, a focal individual could avoid collisions with its Voronoi neighbors, be attracted towards a different neighborhood of visually salient individuals and only process alignment information for one or two selected neighbors. It might also depend on different sets of influential neighbors: for instance a focal individual could avoid collisions with its Voronoi neighbors, be attracted towards a different neighborhood of visually salient individuals and only process alignment information for one or two selected neighbors.

It is thus natural to suggest that influential neighbors are intrinsically associated with different interaction mechanisms, which might also explain why fish point to different neighborhoods.

Our method for identifying influential neighbors is based on the computation of the time-dependent directional correlation between a focal fish and its neighbors. Of course, correlation does not imply causation, so that inferring causal influence between fish from directional correlation requires an extremely cautious methodology.

The methodology we proposed here is based on two solid procedural cornerstones. First, the data used in our study were carefully selected from a clearly recognizable behavior, the collective U-turns, where influence from neighbors undoubtedly exists, and thus should be, to some extent, responsible for a fundamental part of the correlations detected by our method. Time-delay between individuals' direction choices has already been used to measure the interactions between group members in animal flocking. Specifically, Nagy *et al.* [23] used

correlation delay times to reconstruct flight hierarchies in flocks of pigeons. Their approach consisted in integrating delay times over the entire trajectory to obtain a "leadership mark" for each individual. Our assumption is instead that the time-delay results from the individuals' behavior and their environment, which varies in time depending on the information being gathered. To detect the response delay of each individual, we have instead followed the approach employed in [26] that allows for a change of delay over time. In fact, it is easy to show that the time delay between the same pair of fish is not constant, as revealed by our analysis of pair of fish (see Material and methods). Applying Nagy *et al.*' method to different subsets of data in the same experiment, we found that the time delays between the same pair of fish vary substantially (see S2 Fig). The second methodological cornerstone is provided by the results of the null model that clearly show that the correlations we detected come from causal influence between neighbors and not from spurious random coincidences. The results of the null model also confirm that distance rank has no effect.

Identifying the number and position of influential neighbors is an essential step towards reconstructing behavioral cascades of information propagation across a group. Our method provides an accurate basis for mapping interaction network that does not rely on any assumption about the channel (*e.g.*, vision, sound or hydrodynamic interactions) mediating information transfer. We are confident that by adopting our technique to map interactions in different species and different experimental contexts we will gain a much more detailed understanding of the distributed information processing taking place in fish schools.

Materials and methods

Ethics statement

Our experiments have been approved by the Ethics Committee for Animal Experimentation of the Toulouse Research Federation in Biology *N*°1 and comply with the European legislation for animal welfare.

Experimental procedures and data collection

Hemigrammus rhodostomus (rummy-nose tetras, Fig 12A) were purchased from Amazonie Labège (http://www.amazonie.com) in Toulouse, France. Fish were kept in 150 L aquariums on a 12:12 hour, dark:light photoperiod, at 27.7°C (±0.5°C) and were fed *ad libitum* with fish flakes. The average body length of the fish used in these experiments was 31 mm (± 2.5 mm). The experimental tank (120 × 120 cm) was made of glass and was set on top of a box to isolate fish from vibrations. The setup was placed in a chamber made by four opaque white curtains surrounded by four LED light panels to provide an isotropic lighting. A ring-shaped corridor was set inside the experimental tank 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 28.1°C (±0.7°C) (Fig 12B). The corridor was made of a vertical circular outer wall of radius 35 cm and a circular inner wall with a conic shape of radius 25 cm at the bottom, so that the effective width of the corridor available to fish for swimming ranges from 10 cm at the bottom to 12 cm at the surface. The conic shape was chosen to avoid the occlusion on videos of fish swimming too close to the inner wall. Fish were randomly sampled from their breeding tank for a trial and were used at most in only one experiment per day. Groups of 2 or 5 fish were introduced in the experimental tank and acclimatized to their new environment for a period of 10 minutes. Their behavior was then recorded for one hour by a Sony HandyCam HD camera filming from above the setup at 50 images per second in HDTV resolution (1920x1080p). We performed 10 trials for each group size of 2 and 5 fish.

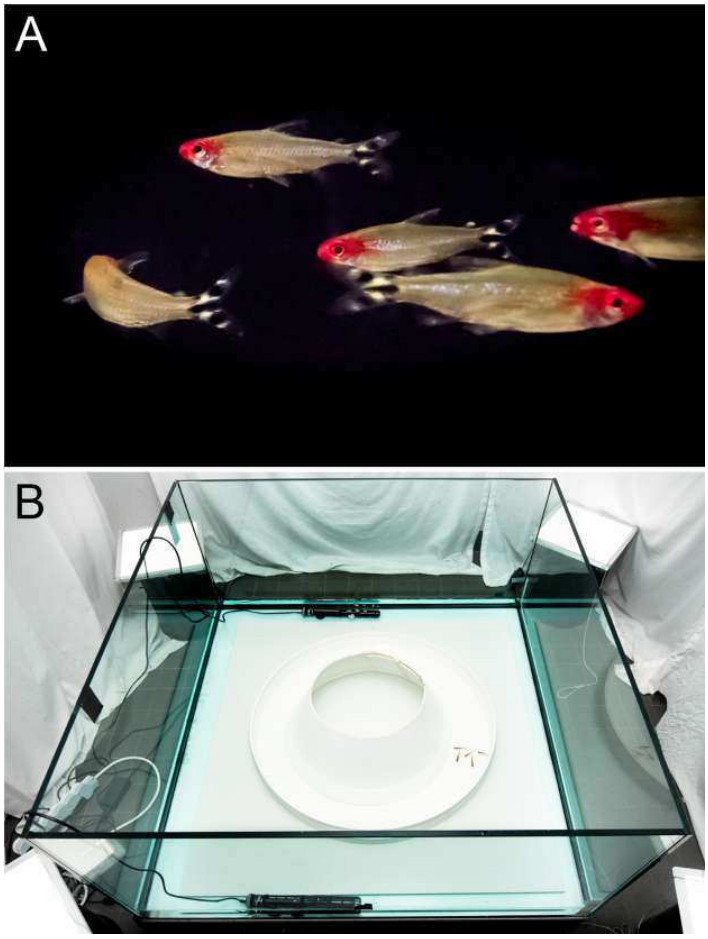


Fig 12. Fish and experimental setup. (A) A spontaneous U-turn initiated by a single fish in a group of five *Hemigrammus rhodostomus* fish. (B) Experimental ring-shaped tank, ©David Villa ScienceImage/CBI/CNRS, Toulouse.

https://doi.org/10.1371/journal.pcbi.1005822.g012

Data extraction and pre-processing

The positions of each fish on each frame were tracked with idTracker 2.1 [10]. Fish were sometimes misidentified by the tracking software, for instance when two fish were swimming too close to each other for a long period of time. In those cases, the missing positions were corrected manually. All sequences with 50 consecutive missing positions or less were interpolated. Larger sequences of missing values were checked by eye to determine whether interpolating was reasonable or not; if not, namely the trajectory doesn't look like a straight line, then merging positions with closest neighbors were considered. Time series of positions were converted from pixels into meters. The origin of the coordinate system was set to the center of the ring-shaped tank. Body orientation of fish were measured using the first axis of a principal component analysis of the fish shapes detected by idTracker 2.1.

Detection and quantification of collective U-turns

Since the experiments were performed in an annular setup, the direction of rotation can be converted into a binary value: clockwise or anti-clockwise. We choose the anti-clockwise direction as the positive values for angular position. Before a U-turn event, all fish move in the same direction, say clockwise. Then, one fish, not necessarily the one located at the front of the group, changes its direction of motion to anti-clockwise direction. After a short transient, the other fish of the group display the same direction change, from clockwise to anti-clockwise. We defined the whole process of changing direction as a collective U-turn (see examples in Fig 1 and in S8 Fig). After data extraction and pre-processing, we found 1111 and 475 collective U-turns in groups of 2 and 5 fish, respectively. The duration distribution of collective U-turns in groups of 2 fish is shown in S3 Fig while the results for groups of 5 fish are shown in S4 Fig. Most of the collective U-turns last between 1 and 3 seconds, while the individual turning time usually lasts between 0.4 and 1 second.

The procedure used to define an individual U-turn for a fish F_i is as follows: we first determine the time $t_{m,i}$ at which the sign of the angle of incidence of fish F_i changes sign (from negative to positive or vice versa). Then, starting from $t_{m,i}$, we reverse time step by step until the first time at which the absolute value of the angle of incidence is higher than a threshold $\bar{\theta}_{s,i}$ is reached. We denote this time by $t_{s,i}$. Similarly, we start again from $t_{m,i}$ and go forward step by step until the first time at which the absolute value of the angle of incidence is higher than a second threshold $\bar{\theta}_{e,i}$ is reached. We denote this time by $t_{e,i}$. To determine the values of the thresholds $\bar{\theta}_{s,i}$ and $\bar{\theta}_{e,i}$, we first compute the moving average of the angle of incidence over a period of 50 time steps (1s in real time), before and after the middle point $t_{m,i}$, with a window of 5 time steps (0.1s in real time), respectively. Then we set the threshold values as the maximum values of the absolute moving average. Doubling the length of the period of time over which the average is computed, or doubling the width of the window, do not affect the results. Finally, the time at which the collective U-turn starts (resp. ends) is defined by $\min\{t_{s,i}\}_{i=1}^{N}$ (resp. $\max\{t_{e,i}\}_{i=1}^{N}$).

Position rank in a group

The relative position of a fish F_i in a group of N fish is calculated by projecting the vector position of the fish $\vec{u}_i$ on the average group velocity vector $\vec{z} = (1/N)\sum_{i=1}^{N}\vec{v}_i$. This allows us to define a group centroid in the direction of $\vec{z}$, with respect to which the fish are ranked: the first fish in the group is the fish whose projection on $\vec{z}$ is the most advanced one in the direction of motion of the group (given by $\vec{z}$), the second fish in the group is the second most advanced, and so on. Relative distance between fish are not taken into account when establishing the rank.

Optimal setting parameters for influential neighbors identification

Four parameters are used to identify influential neighbors: the time-delay τ, the window size w, the correlation threshold $C_{\min}$ above which individuals are supposed to be interacting, and the threshold ε for selecting more than one influential fish.

The time delay must be specified along the whole trajectory of the focal fish: it is thus a series of values $\{\tau_k^*\}_{k=0}^{M}$, where M is the number of time-steps or frames in the individual U-turn. The parameters $C_{\min}$, ε and w are in turn given for all time and for all fish by means of a sensitivity analysis described in the next section.

Assume by now that the three values $C_{\min}$, ε and w are known, and denote by F_i the focal fish and by F_j one of its neighbors. Then, the series of time-delays $\{\tau_k^*\}_{k=0}^{M_i}$ is built recursively as follows (actually only w is required to extract the time delays).

Denote by $\Gamma_i(t_k)$ the highest value of the pairwise directional correlation C_{ij} of the velocity of fish F_i at time t_k with the velocity of F_j at each time-step in the range of the previous $(\tau_{k-1}^* + 1)$ time-steps $R_k = [0, \tau_{k-1}^* + 1]$:

$$\Gamma_i(t_k, w) = \max_{\tau_r \in R_k}\{C_{ij}(t_k, \tau_r, w)\}. \tag{6}$$

Then, the time-delays τ_k^*, $k = 1, \ldots, M_i$, are determined by the smallest value of the time-delay $\tau_r \in R_k$ where $\Gamma_i(t_k, w)$ reaches its maximum. For t_1, the maximum correlation is reached at $C_{ij}(t_1, \tau_1^*, w)$, for some time-delay $\tau_1^* \in R_1 = [0, \tau_0^* + 1]$. We set $\tau_0^* = 50$ for the initial value of the recurrence. For the rest of time-delays τ_k^*, $k = 2, \ldots, M_i$, the size of R_k is based on the assumption that if, at some time t, F_i copies the behavior that F_j displayed at a previous time $t - \tau$, then, after time t, F_i will not copy the behavior that F_j displayed at any time earlier than $t - \tau$.

Time-delays obtained with more complicated and time consuming procedures such as the time-ordered technique developed in [26] or through the similarity analysis based on Fréchet distances [25] would in principle produce similar values.

Fig 13B shows the distribution of time-delays obtained with this procedure in groups of two fish. The distribution is clearly bimodal with a first peak when $\tau = 0$ and a second one around $\tau = 0.4$ s. Considering a reaction time threshold of 50-100 ms for a fish to integrate information and reach a decision [42], we cannot attribute small values of time-delays to situations where the behavioral decision of the focal fish has been influenced by its neighbors. This is confirmed by the analysis of the spatial distribution of the extracted time-delays (Fig 13A), where we show that the lowest average values of τ are found mostly when the neighbor was behind the focal fish, in a zone with the lowest perception [15], while the highest values of $\tau >$ 0.4 s are found when the neighbor is located in front of the focal fish. This has lead us to consider in our analyzes only situations where $\tau > \tau_R = 0.04$ s.

Parameter selection

Although the time-delays $\{\tau_k^*\}_{k=0}^{M}$ are determined once w is known, they also strongly depend on $C_{\min}$ and ε, as the value of these three parameters must be fixed at the same time. This is done by means of a sensitivity analysis in which we have tested the following 40 combinations of parameter values: $w \in \{0, 1, 2, 3, 4\}$, $\varepsilon = \{3, 5\}$, and $C_{\min} \in \{0.995, 0.99, 0.95, 0.5\}$.

Each combination $(C_{\min}, \varepsilon, w)$ gives rise to four histograms like those depicted in Fig 7. These histograms constitute the solution of our method of analysis, and can be characterized by a vector $\vec{S}(C_{\min}, \varepsilon, w)$ in 19 dimensions: (*i*) the 5 proportions of the number of influential neighbors in groups of 5 fish, (*ii*) the 4 proportions of their distance rank, (*iii*) the 5 proportions of their position rank, and (*iv*) the 5 proportions of their turning rank. This allows us to determine how similar are the results arising from two combinations $(C_{\min}, \varepsilon, w)$ and $(C'_{\min}, \varepsilon', w')$, by computing the cosine similarity of the two vectors $\vec{S}(C_{\min}, \varepsilon, w)$ and $\vec{S}'(C'_{\min}, \varepsilon', w')$.

The cosine similarity of two vectors $\vec{a}$ and $\vec{b}$, denoted $\cos_{\text{sim}}(\vec{a}, \vec{b})$, is the cosine of the angle between these two vectors. Thus, two colinear vectors are such that $\cos_{\text{sim}}(\vec{a}, \vec{b}) = \pm 1$ independently of their magnitude, while two perpendicular vectors are such that $\cos_{\text{sim}}(\vec{a}, \vec{b}) = 0$. In our case, the components of the vectors are positive, so $\cos_{\text{sim}}(\vec{S}, \vec{S}') \geq 0$ for all $(C_{\min}, \varepsilon, w)$

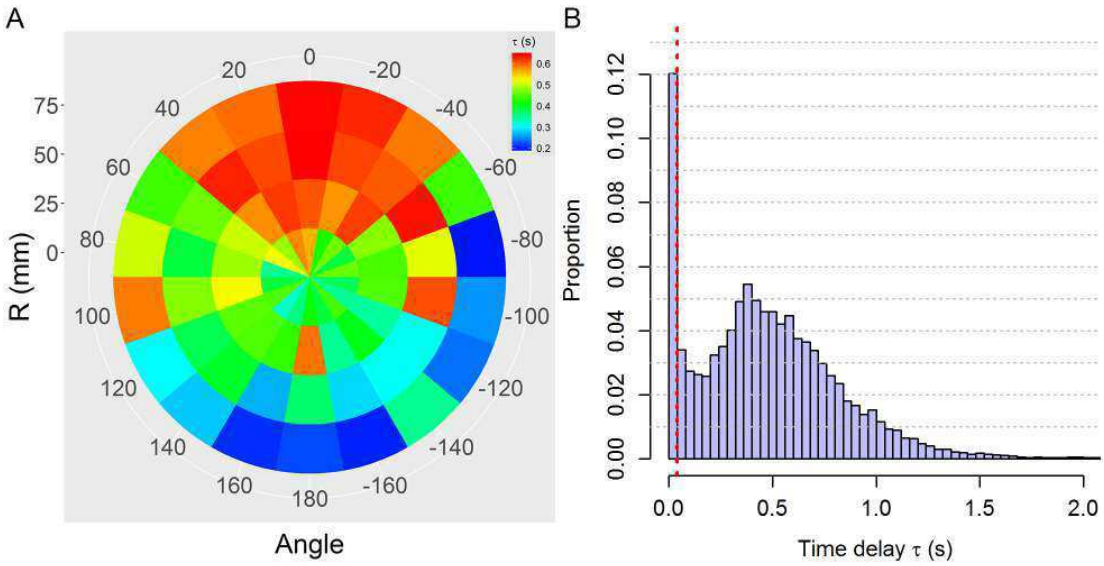


Fig 13. Distribution of time-delay τ. (A) Spatial distribution of time-delays obtained by selecting the maximum of the pairwise correlation between the focal fish and its neighbor. The color of each bin represent the mean value of all the cases in that bin. An angle of 0˚ degree corresponds to when the influential neighbor is in front of the focal individual. (B) Spatially integrated distribution of time-delays. The data here are the same as those used in panel A. The dotted line corresponds to the reaction time threshold $\tau_R = 0.04$s.

https://doi.org/10.1371/journal.pcbi.1005822.g013

and $(C'_{\min}, \varepsilon', w')$. Moreover, as the components are proportions, colinearity implies identity, both in direction and magnitude. Thus, $\cos_{\text{sim}}(\vec{S}, \vec{S}') = 1$ means that both results are identical, while $\cos_{\text{sim}}(\vec{S}, \vec{S}') = 0$ means that they differ as much as possible.

S5 Fig shows the cosine similarity matrix for the 40 combinations we have tested. Note that the matrix is symmetric with respect to the diagonal, where $\cos_{\text{sim}}(\vec{S}, \vec{S}) = 1$. Except for $C_{\min} = 0.5$, all similarity values are in the thin range [0.96, 1], showing that all combinations yield practically the same results. The higher dissimilarity is found in the white-yellow lines, where one of the combinations is $(C_{\min}, \varepsilon, w) = (0.5, 3, 2)$.

The selection of parameter values is thus done as follows.

We choose $w = 2$, which corresponds to the higher dissimilarity regions. The selected time window size is sufficiently large so that the jagged nature of the movement data is smoothed out but not too large so that the actual turns gets washed out from the data.

Using $\varepsilon = 3$ or $\varepsilon = 5$ yields very similar results and we have arbitrarily chosen $\varepsilon = 3$.

The selection of $C_{\min}$ is done by a specific procedure, which consists in calculating the number of data points that remain available for our analysis for each value of $C_{\min}$. S6 and S7 Figs exhaustively demonstrate that the larger $C_{\min}$ is, the less data points remain available, and *vice versa*. We might be prone to choose a sufficiently small $C_{\min}$ in order to get the maximum number of data points. However, according to our definition of influential neighbor, $C_{\min}$ should be sufficiently large to select only the real influential neighbors. We have thus chosen the highest value which provides a sufficiently large number of data points, that is, the largest value before the fall of the number of data points in S11 Fig, $C_{\min} = 0.95$. This value preserves 61% (23830) and 76% (69703) of data points for $N = 2$ and $N = 5$ respectively.

Null model of collective U-turns

We want to design artificial collective U-turns in groups of 5 fish where all fish perform an individual U-turn at more or less the same place and more or less the same time, and in the same direction (clockwise or anti-clockwise). Fish must coincide in time and space to constitute a "group", but individual U-turns must happen in an absolutely independent way. Correlations at hand in this paper are thus reduced to a minimum, while preserving the general aspect of a group of fish changing direction.

Our experimental data provide us with $5 \times 475 = 2375$ trajectories of individual fish, which we have conveniently normalized and combined to build 1000 groups of 5 fish changing direction in the same spatiotemporal interval. This is done as follows.

The whole trajectory of a fish F_i during a U-turn takes place in an interval of time $[t_{s,i}, t_{e,i}]$, where $t_{s,i}$ is the instant at which the individual U-turn of fish F_i starts, and $t_{e,i}$ is the time at which the individual U-turn ends. See the paragraph above Eq (1). The trajectory of fish F_i in radial coordinates is given by

$$\{(\rho_i(t_k), \psi_i(t_k))\}_{k=1}^{N_i}, \tag{7}$$

where $\rho_i(t_k)$ is the radius (distance of the fish from the center of the tank), $\psi_i(t_k)$ the already defined angle position (computed anticlockwise as positive), and N_i is the number of time-steps t_k in the trajectory.

Denote by T_i the instant at which fish F_i effectively turns, *i.e.*, F_i is perpendicular to the wall: $\sin(\theta_{wi}(T_i)) = 0$. In well defined individual U-turns as the ones we are using in our data, this happens only once per U-turn. Accordingly, $(\rho_i(T_i), \psi_i(T_i))$ denotes the fish position at time T_i.

Although we would like to have absolutely uncorrelated fish, it would not make sense to use groups of trajectories that do not reproduce a consistent U-turn, *e.g.*, if one fish makes its U-turn much later than another, or on the other side of the tank. We thus try to decorrelate fish trajectories as much as possible, while preserving at the same time the typical spatiotemporal shape of real collective U-turns.

The decorrelation of all individual U-turns is done with the following two steps:

- Spatial rotation: For all individual fish F_i in all U-turns, we rotate its trajectory an angle $-\psi_i(T_i) + \pi/2 + \xi_i$, where ξ_i is a random number in $[-\pi/12, \pi/12]$ sampled uniformly, so that the new location of fish F_i at the time T_i when it performs its individual U-turn is in the upper part of the tank around $\pi/2$, in $[5\pi/12, 7\pi/12]$.
- Time shift: For all individual fish F_i in all U-turns, we shift the time scale a value $-T_i + \zeta_i$, where ζ_i is a random number sampled uniformly in $[-1, 1]$ s, so that F_i makes its individual U-turn at around time 0, in $[-1, 1]$ seconds.

The artificial collective U-turn is thus built as follows:

1. Select randomly 5 real collective U-turns, and, from each collective U-turn, select randomly one trajectory. Rotate and time-shift trajectories according to the process described above.
2. Select randomly one of the 5 fish as the fish of reference F_{ref} for building the artificial U-turn. If necessary, mirror the trajectories of other fish so that all fish move in the same direction as F_{ref} with respect to the center of the tank, *i.e.*, clockwise or anti-clockwise.

Then, the fish of reference of the artificial U-turn will make its individual U-turn at time $\zeta_{\text{ref}} \in [-1, 1]$ s and position $(\rho_{\text{ref}}(T_{\text{ref}}), \pi/2 + \xi_{\text{ref}})$. The other four fish F_j will make their individual U-turn at time $\zeta_j \in [-1, 1]$ s and position $(\rho_j(T_j), \pi/2 + \xi_j)$ respectively, for $j = 1, \ldots, 5, j \neq$ ref.

We have depicted in S9 Fig a set of artificial U-turns for comparison with the real U-turns shown in S8 Fig. Note that in these figures the time-scale has been shifted again so that collective U-turns start at $t = 0$ s.

Supporting information

S1 Video. Sample video of an U-turn event in a group of 5 fish. Original video of an U-turn event, corresponding to Fig 4 and S2 Video.
(AVI)

S2 Video. Sample video of an U-turn dynamic in a group of 5 fish. Video showing the velocities of fish and interaction dynamics in the group, corresponding to Fig 4 and S1 Video.
(AVI)

S1 Fig. Directional correlation $H_{ij}(t, \tau)$ between fish F_i and F_j. For $i = 2, \ldots, 5$ (rows) and $j = 1, \ldots, 5, j \neq i$ (columns), *e.g.*, first row is for fish F_2: (A) $H_{21}(t, \tau)$, (B) $H_{23}(t, \tau)$, (C) $H_{24}(t, \tau)$ and (D) $H_{25}(t, \tau)$.
(TIF)

S2 Fig. Different values of τ^* for different subsets of the same data set computed with the method of Nagy *et al.* [23]. Consider the dataset of U-turns of 2 fish composed by U-turn number 1 to U-turn number 36, coming from the same experiment, and divide it in two subsets S_A and S_B containing respectively the U-turns [1,. . .,18] and the U-turns [19,. . .,36]. (A) Average directional correlation C_{ij} with respect to time-delay τ for the U-turns from dataset S_A. Red star and dashed blue vertical line denotes $\tau^* = 0.96$. (B) C_{ij} for the U-turns from dataset S_B. Red star: $\tau^* = 0.32$. (C) C_{ij} for all the U-turn in data set $S_A \cup S_B$. Red star: $\tau^* = 0.80$. The method of Nagy *et al.* is based on the assumption that the pairwise interaction between two individuals in a group has a constant time-delay τ^*. However, Panels A and B provide different values of τ^* for different data sets, showing that the method of Nagy *et al.* is not suitable for studing our data, and that the method we introduce here, which is based on the detection of dynamic time-delays, has potential for a broader range of applications.
(TIF)

S3 Fig. Distribution of the average duration (in seconds) of (A) individual and (B) collective U-turns in groups of 2 fish. Collective U-turns last around twice the duration of individual U-turns.
(TIF)

S4 Fig. Distribution of the average duration (in seconds) of (A) individual and (B) collective U-turns in groups of 5 fish. Collective U-turns last almost four times the duration of individual U-turns.
(TIF)

S5 Fig. Parameter comparison matrix. Matrix of 40 × 40 square cells, where each cell corresponds to the similarity value SV arising from the comparison of the two parameter combinations shown in the corresponding horizontal and vertical axes. We considered 40 parameter combinations, thus the size of the matrix. The similarity value SV is represented by the color of the cell, where the brightest red color corresponds to SV = 1 and the white color to SV = 0.92. For instance, the top-left cell displays a similarity value of SV = 0.95, showing how similar the results are when comparing the two combinations $\{\varepsilon = 5, C_{\min} = 0.995, w = 0\}$ (horizontal axis) and $\{\varepsilon = 3, C_{\min} = 0.5, w = 4\}$ (vertical axis). Cells along the diagonal correspond to the comparison of two identical parameter combinations and therefore SV = 1 there.
(TIF)

S6 Fig. Available data for different values of the average directional correlation threshold C_{min} in the case of $N = 2$ fish. Small panels: (there are 10, one per experiment) Number of data points available from the respective experiment for each value of C_{min} in [0.5, 1]. The values of C_{min} are denoted by small circles. Three specific values are shown by arrows: 0.6, 0.95 and 0.995. The value highlighted in red corresponds to the value we chose and is denoted by a star instead of a circle. Each vertical line corresponds to the fish that is taken as being the focal fish: F_1 (red) and F_2 (cyan). For instance, selecting $C_{min} = 0.6$ in the upper-left small panel, 700 data points will be available for both fish. For $C_{min} = 0.95$, around 450 points will be available for both fish.
Leftmost higher panel: Total number of data points available from all fish from all the experiments (summary of the 10 small panels, *i.e.*, there is only one –pink– line). Vertical axis: ratio between the available number of data points for C_{min} and the number of data points available for $C_{min} = 0.5$. Total data points available from all the experiments (for $C_{min} = 0.5$): 39381; data points available for $C_{min} = 0.95$: 23830.
(TIF)

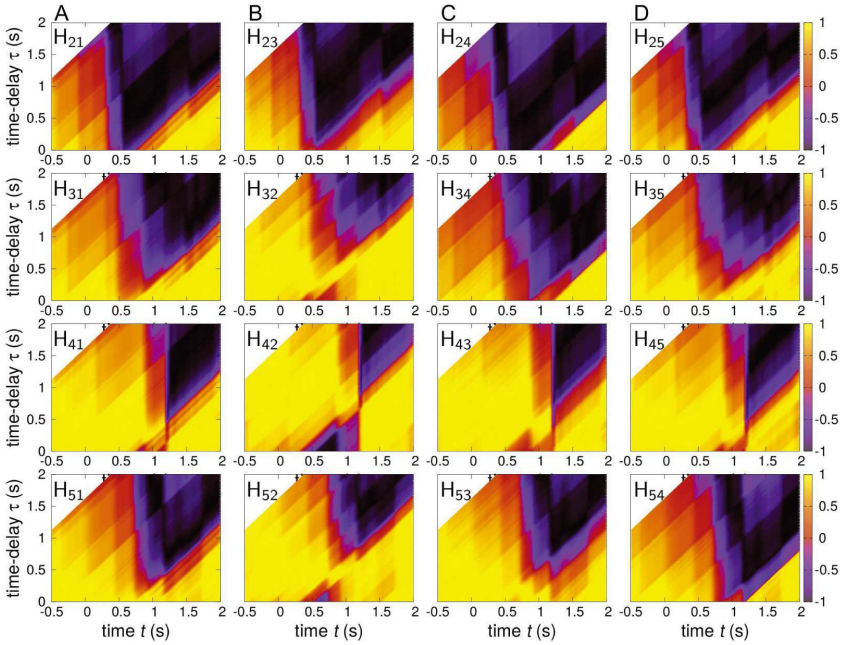
S7 Fig. Available data for different values of the average directional correlation threshold C_{min}, in the case of $N = 5$ fish. Small panels: (there are 10, one per experiment) Number of data points available from the respective experiment for each value of C_{min} in [0.5, 1]. The values of C_{min} are denoted by small circles. Three specific values are shown by arrows: 0.6, 0.95 and 0.995. The value highlighted in red corresponds to the value we chose and is denoted by a star instead of a circle. Each vertical line corresponds to the fish that is taken as being the focal fish: F_1 (red), F_2 (yellow), F_3 (green), F_4 (blue) and F_5 (magenta). For instance, selecting $C_{min} = 0.6$ in the third small panel of the upper row, 55 data points will be available for each one of the 5 fish. For $C_{min} = 0.95$, around 75 points will be available for each fish.
Leftmost higher panel: Total number of data points available from all fish from all the experiments (summary of the 10 small panels, *i.e.*, there is only one –pink– line). Vertical axis: ratio between the available number of data points for C_{min} and the number of data points available for $C_{min} = 0.5$. Total data points available from all the experiments (for $C_{min} = 0.5$): 91827; data points available for $C_{min} = 0.95$: 69703.
(TIF)

S8 Fig. Collective U-turns observed in experiments with $N = 5$ fish.
(TIF)

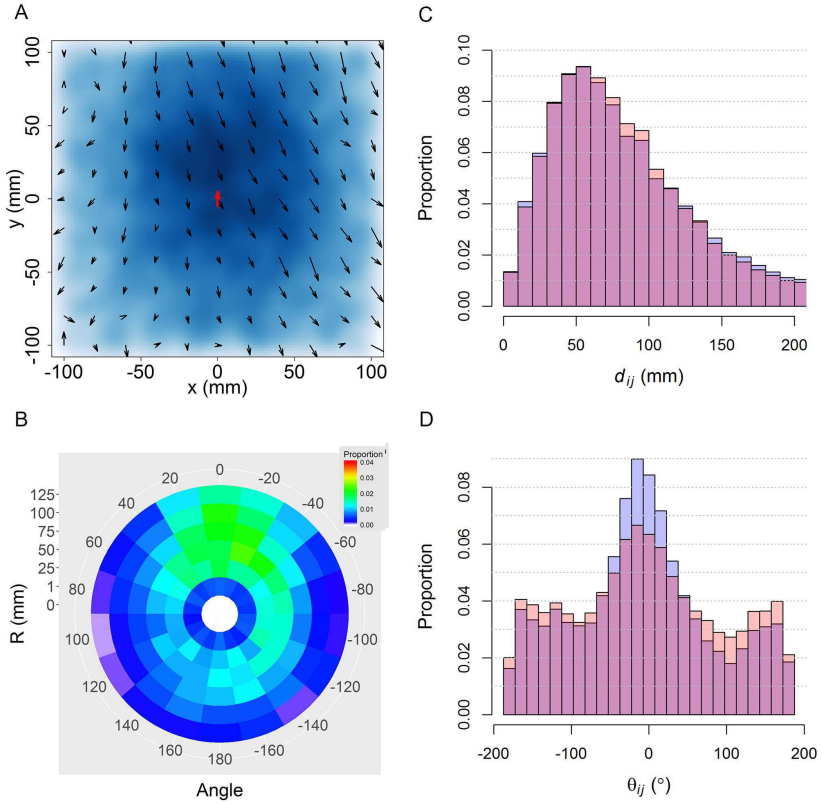
S9 Fig. Artificial collective U-turns obtained with the null model.
(TIF)

S10 Fig. Homogeneous (isotropic) spatial distribution of "influential neighbors" in collective artificial U-turns. (A) Density map of "influential neighbors" location (blue) and their average relative velocity field (arrows) with respect to the focal fish (red arrow). (B) Average spatial distribution.
(TIF)

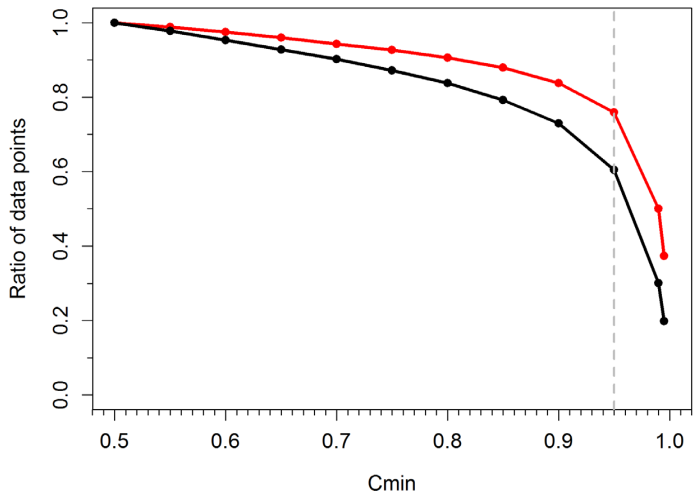
S11 Fig. Number of available data points for different values of C_{min}. Solid black line: Remaining data points for each value of C_{min} for $N = 2$ according to the leftmost panel in S6 Fig. Red line: same thing, for $N = 5$, according to S7 Fig. Dashed line: highest number of available data points before the sharp fall of the black curve at $C_{min} = 0.95$.
(TIF)


A
B
C
D
H21
H23
H24
H25
H31
H32
H34
H35
H41
H42
H43
H45
H51
H52
H53
H54
time-delay τ (s)
time t (s)

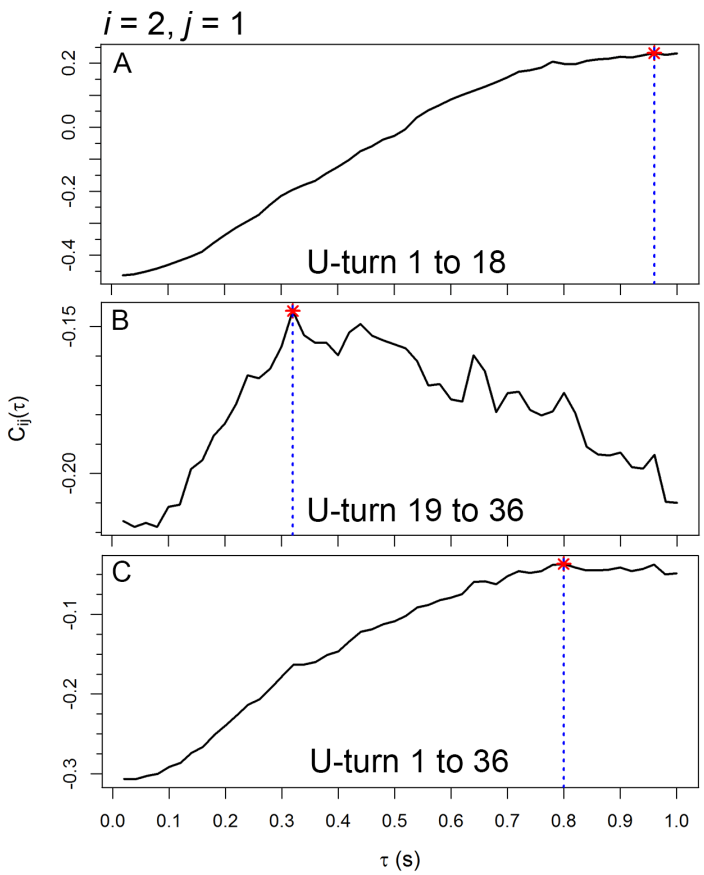
S1 Fig.


A
B
C
D
y (mm)
x (mm)
Proportion
d_{ij} (mm)
R (mm)
Angle
θ_{ij} (°)

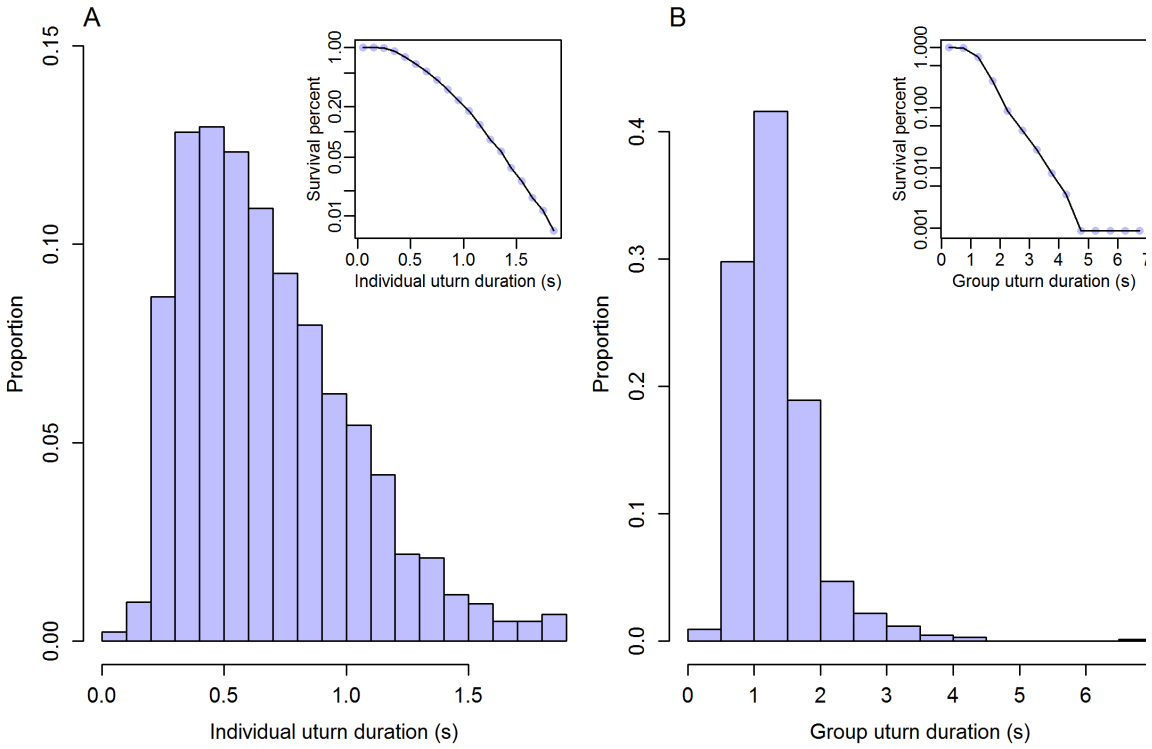
S2 Fig.


Ratio of data points
0.0
0.2
0.4
0.6
0.8
1.0
0.5
0.6
0.7
0.8
0.9
1.0
Cmin

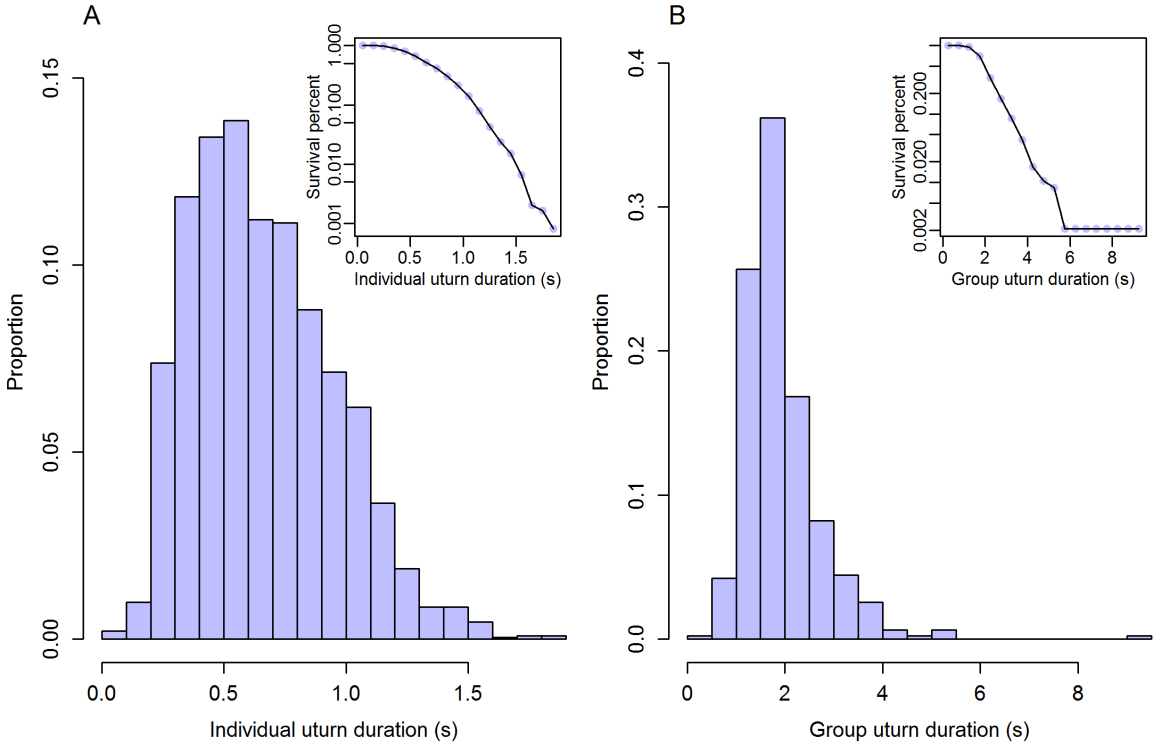
S3 Fig.


i = 2, j = 1
A
U-turn 1 to 18
B
U-turn 19 to 36
C
U-turn 1 to 36
Cij(τ)
τ (s)

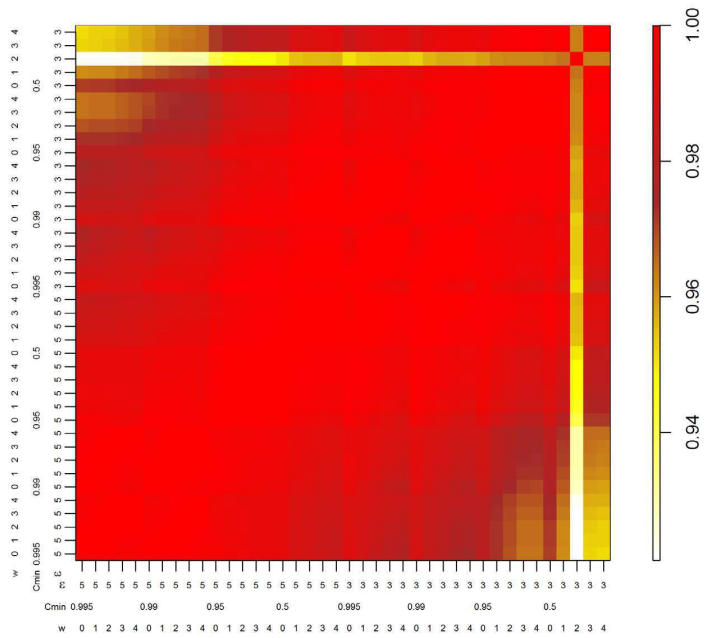
S4 Fig.


A
Proportion
0.00
0.05
0.10
0.15
0.0
0.5
1.0
1.5
Individual uturn duration (s)
Survival percent
0.01
0.05
0.20
1.00
0.0
0.5
1.0
1.5
Individual uturn duration (s)
B
Proportion
0.0
0.1
0.2
0.3
0.4
0
1
2
3
4
5
6
Group uturn duration (s)
Survival percent
0.001
0.010
0.100
1.000
0
1
2
3
4
5
6
Group uturn duration (s)

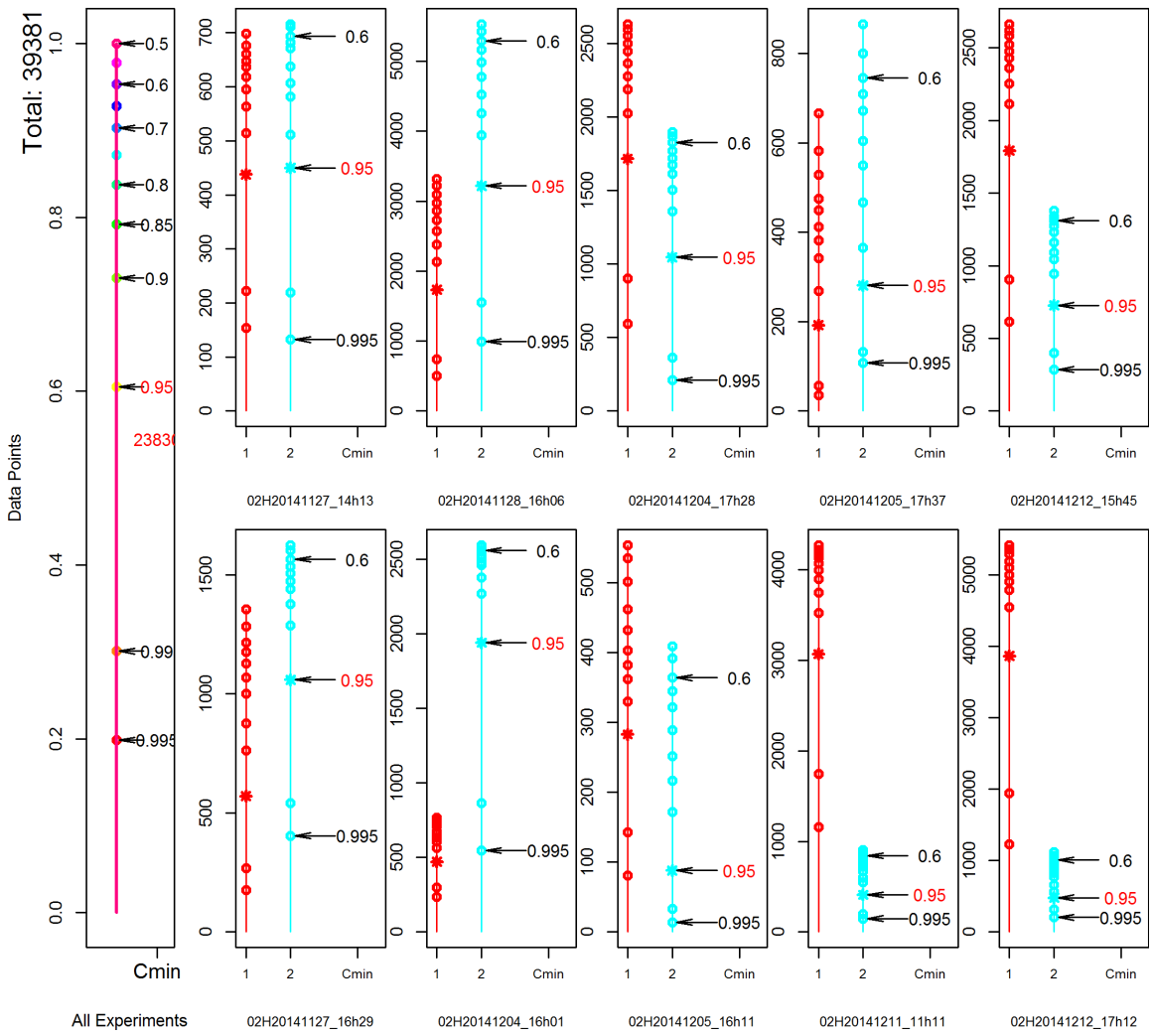
S5 Fig.


A
Proportion
0.00
0.05
0.10
0.15
0.0
0.5
1.0
1.5
Individual uturn duration (s)
Survival percent
0.001
0.010
0.100
1.000
0.0
0.5
1.0
1.5
Individual uturn duration (s)
B
Proportion
0.0
0.1
0.2
0.3
0.4
0
2
4
6
8
Group uturn duration (s)
Survival percent
0.002
0.020
0.200
0
2
4
6
8
Group uturn duration (s)

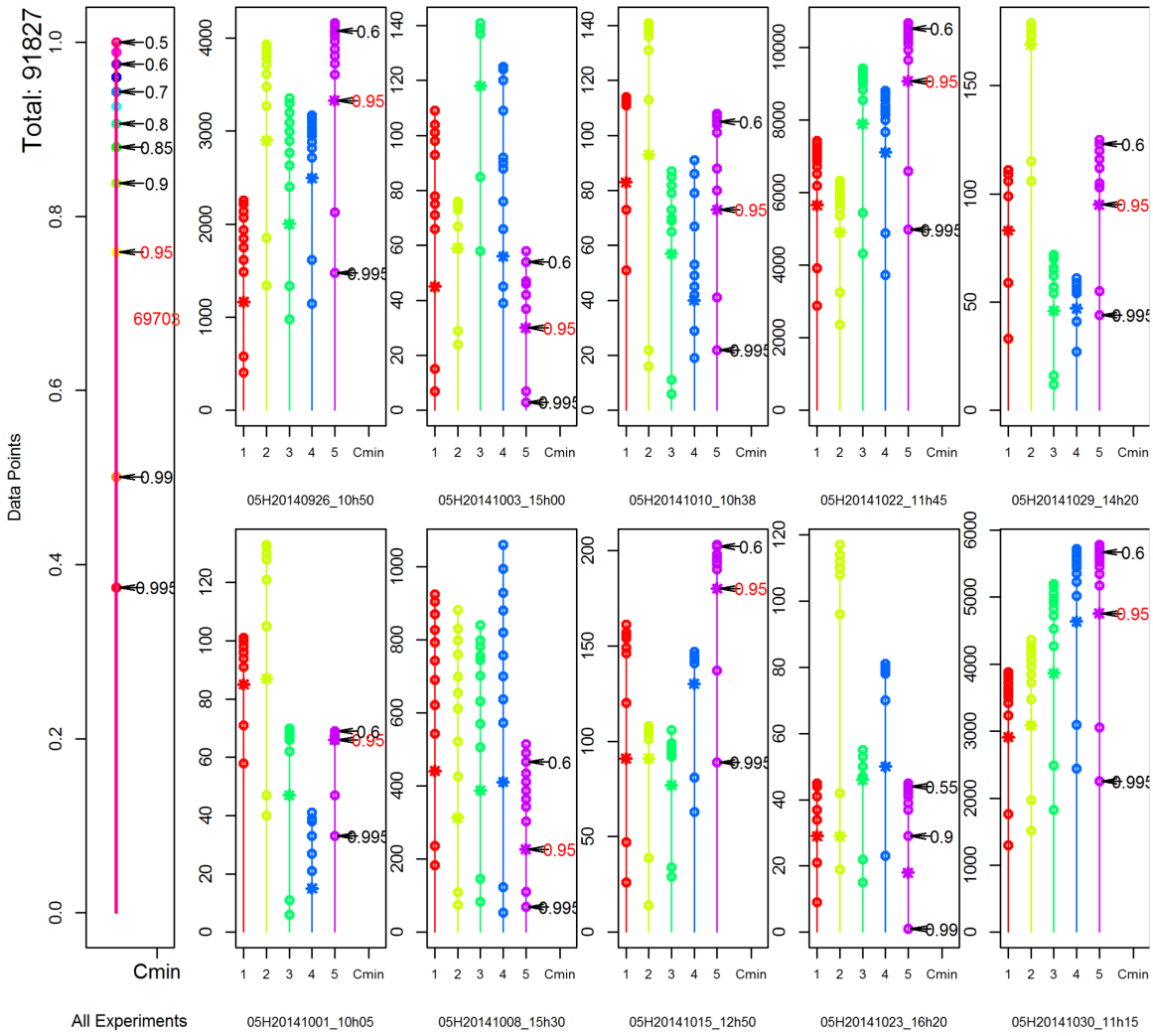
S6 Fig.


1.00
0.98
0.96
0.94
Cmin 0.995
0.99
0.95
0.5
0.995
0.99
0.95
0.5
w 0 1 2 3 4

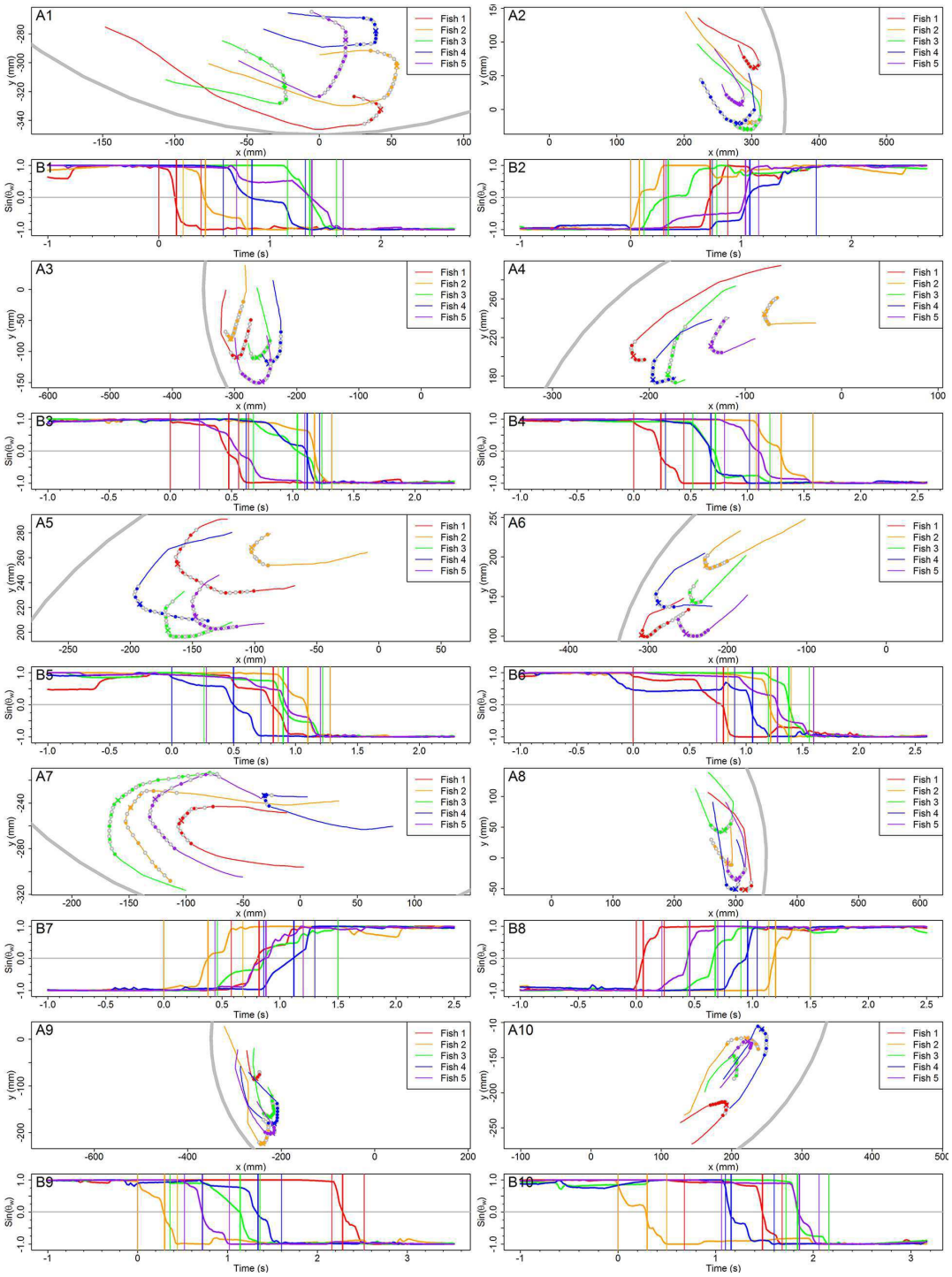
S7 Fig.


Total: 39381
Data Points
Cmin
All Experiments
02H20141127_14h13
02H20141128_16h06
02H20141204_17h28
02H20141205_17h37
02H20141212_15h45
02H20141127_16h29
02H20141204_16h01
02H20141205_16h11
02H20141211_11h11
02H20141212_17h12

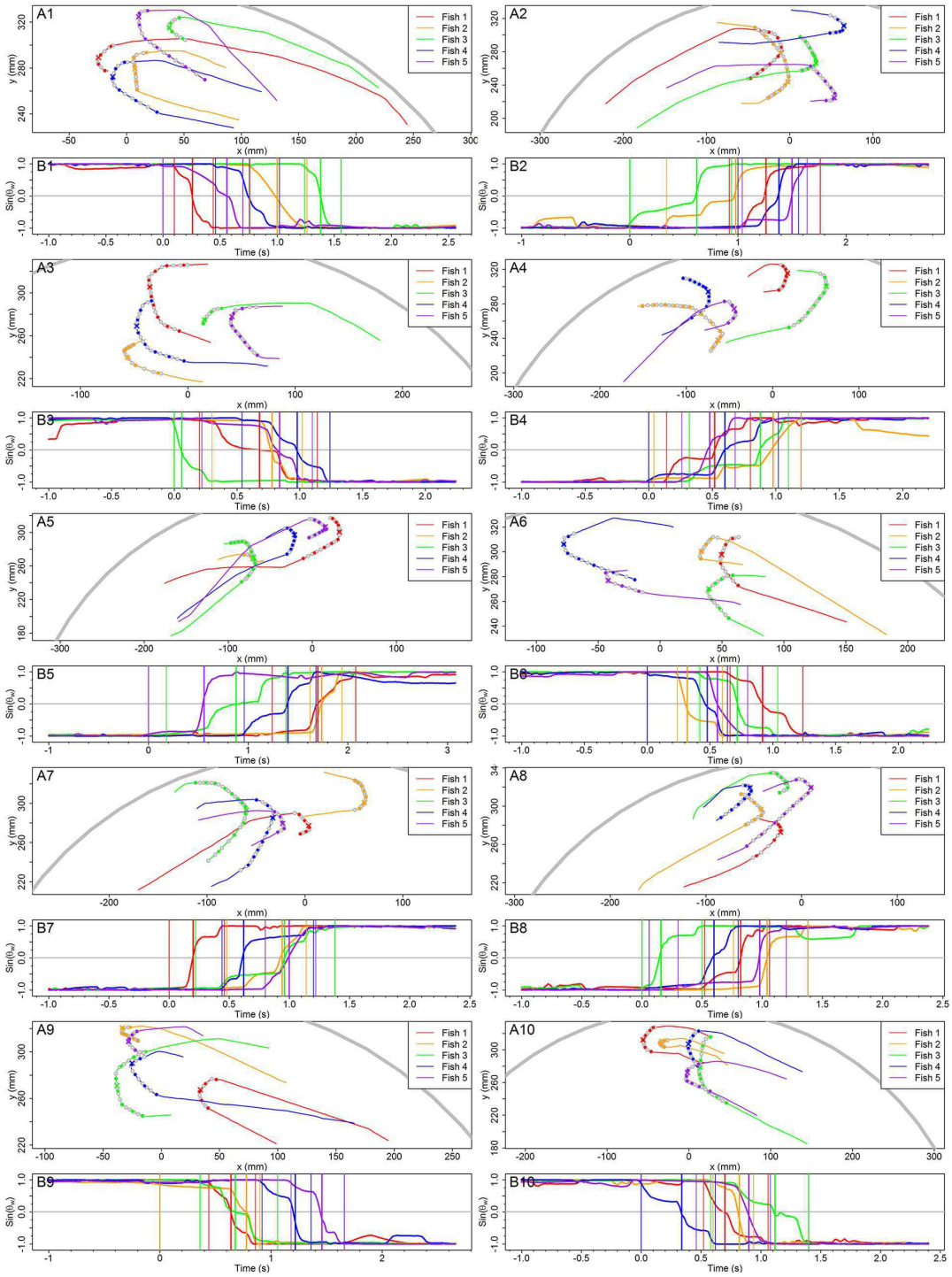
S8 Fig.


Total: 91827
Data Points
Cmin
All Experiments
05H20140926_10h50
05H20141003_15h00
05H20141010_10h38
05H20141022_11h45
05H20141029_14h20
05H20141001_10h05
05H20141008_15h30
05H20141015_12h50
05H20141023_16h20
05H20141030_11h15

S9 Fig.


A1
A2
A3
A4
A5
A6
A7
A8
A9
A10
B1
B2
B3
B4
B5
B6
B7
B8
B9
B10
Fish 1
Fish 2
Fish 3
Fish 4
Fish 5
x (mm)
y (mm)
Time (s)

S10 Fig.


A1
A2
A3
A4
A5
A6
A7
A8
A9
A10
B1
B2
B3
B4
B5
B6
B7
B8
B9
B10
Fish 1
Fish 2
Fish 3
Fish 4
Fish 5
x (mm)
y (mm)
Time (s)

S11 Fig.

Acknowledgments

We would like to acknowledge Schloss Dagstuhl Leibniz Zentrum für Informatik in Germany and all the participants at the seminar 16022 on Geometric and Graph-based Approaches to Collective Motion organized in 2016 where many ideas developed in this paper have been discussed. In particular we would like to thank Martin Beye, Anael Engel, Marc van Kreveld, Frank Staals and Goce Trajcevski. We also thank Pierre Tichit and Gérard latil for technical assistance.

Author Contributions

Conceptualization: Guy Theraulaz.

Data curation: Valentin Lecheval.

Formal analysis: Li Jiang, Luca Giuggioli, Ramón Escobedo.

Funding acquisition: Zhangang Han, Guy Theraulaz.

Investigation: Li Jiang, Ramón Escobedo, Valentin Lecheval.

Methodology: Luca Giuggioli, Andrea Perna, Ramón Escobedo, Clément Sire, Guy Theraulaz.

Project administration: Zhangang Han, Guy Theraulaz.

Software: Li Jiang.

Supervision: Zhangang Han, Guy Theraulaz.

Validation: Luca Giuggioli, Andrea Perna, Ramón Escobedo, Clément Sire, Guy Theraulaz.

Visualization: Li Jiang, Ramón Escobedo.

Writing – original draft: Li Jiang, Luca Giuggioli, Andrea Perna, Ramón Escobedo, Guy Theraulaz.

Writing – review & editing: Li Jiang, Luca Giuggioli, Andrea Perna, Ramón Escobedo, Guy Theraulaz.

References

1. Vicsek T, Zafeiris A. Collective motion. Physics Reports. 2012; 517(3):71–140. https://doi.org/10.1016/j.physrep.2012.03.004
2. Aoki I. A simulation study on the schoolng mechanism in fish. Bull Japan Soc Sci Fish. 1982; 48. https://doi.org/10.2331/suisan.48.1081
3. Reynolds CW. Flocks, herds and schools: A distributed behavioral model. ACM SIGGRAPH computer graphics. 1987; 21(4):25–34. https://doi.org/10.1145/37402.37406
4. Huth A, Wissel C. The simulation of the movement of fish schools. Journal of theoretical biology. 1992; 156(3):365–385. https://doi.org/10.1016/S0022-5193(05)80681-2
5. Niwa HS. Self-organizing dynamic model of fish schooling. Journal of theoretical Biology. 1994; 171(2): 123–136. https://doi.org/10.1006/jtbi.1994.1218
6. Vicsek T, Czirók A, Ben-Jacob E, Cohen I, Shochet O. Novel type of phase transition in a system of self-driven particles. Physical review letters. 1995; 75(6):1226. https://doi.org/10.1103/PhysRevLett.75.1226 PMID: 10060237
7. Couzin ID, Krause J, James R, Ruxton GD, Franks NR. Collective memory and spatial sorting in animal groups. Journal of theoretical biology. 2002; 218(1):1–11. https://doi.org/10.1006/jtbi.2002.3065 PMID: 12297066
8. Straw AD, Branson K, Neumann TR, Dickinson MH. Multi-camera real-time three-dimensional tracking of multiple flying animals. Journal of The Royal Society Interface. 2010; p. rsif20100230. http://rsif.royalsocietypublishing.org/content/early/2010/07/13/rsif.2010.0230

9. Wu HS, Zhao Q, Zou D, Chen YQ. Automated 3D trajectory measuring of large numbers of moving particles. Optics express. 2011; 19(8):7646–7663. https://doi.org/10.1364/OE.19.007646 PMID: 21503074

10. Pérez-Escudero A, Vicente-Page J, Hinz RC, Arganda S, de Polavieja GG. idTracker: tracking individuals in a group by automatic identification of unmarked animals. Nature methods. 2014; 11(7):743–748. https://doi.org/10.1038/nmeth.2994 PMID: 24880877

11. Katz Y, Tunstrøm K, Ioannou CC, Huepe C, Couzin ID. Inferring the structure and dynamics of interactions in schooling fish. Proceedings of the National Academy of Sciences. 2011; 108(46):18720–18725. https://doi.org/10.1073/pnas.1107583108

12. Herbert-Read JE, Perna A, Mann RP, Schaerf TM, Sumpter DJ, Ward AJ. Inferring the rules of interaction of shoaling fish. Proceedings of the National Academy of Sciences. 2011; 108(46):18726–18731. https://doi.org/10.1073/pnas.1109355108

13. Gautrais J, Ginelli F, Fournier R, Blanco S, Soria M, Chaté H, et al. Deciphering interactions in moving animal groups. PLoS Comput Biol. 2012; 8(9):e1002678. https://doi.org/10.1371/journal.pcbi.1002678 PMID: 23028277

14. Mwaffo V, Anderson RP, Butail S, Porfiri M. A jump persistent turning walker to model zebrafish locomotion. Journal of The Royal Society Interface. 2014; 12(102):20140884. https://doi.org/10.1098/rsif.2014.0884

15. Calovi DS, Litchinko A, Lecheval V, López U, Escudero AP, Chaté H, Sire C, Theraulaz G. Disentangling and modeling interactions in fish with burst-and-coast swimming. https://arxiv.org/abs/1703.03801

16. Lopez U, Gautrais J, Couzin ID, Theraulaz G. From behavioural analyses to models of collective motion in fish schools. Interface focus. 2012; p. rsfs20120033. https://doi.org/10.1098/rsfs.2012.0033 PMID: 24312723

17. Ginelli F, Chaté H. Relevance of metric-free interactions in flocking phenomena. Physical Review Letters. 2010; 105(16):168103. https://doi.org/10.1103/PhysRevLett.105.168103 PMID: 21231019

18. Lemasson B, Anderson J, Goodwin R. Collective motion in animal groups from a neurobiological perspective: the adaptive benefits of dynamic sensory loads and selective attention. Journal of theoretical biology. 2009; 261(4):501–510. https://doi.org/10.1016/j.jtbi.2009.08.013 PMID: 19699212

19. Lemasson BH, Anderson JJ, Goodwin RA. Motion-guided attention promotes adaptive communications during social navigation. Proceedings of the Royal Society of London B: Biological Sciences. 2013; 280(1754):20122003. https://doi.org/10.1098/rspb.2012.2003

20. Collignon B, Séguret A, Halloy J. A stochastic vision-based model inspired by zebrafish collective behaviour in heterogeneous environments. Royal Society open science. 2016; 3(1):150473. https://doi.org/10.1098/rsos.150473 PMID: 26909173

21. Pomeroy H, Heppner F. Laboratory determination of startle reaction time of the starling (Sturnus vulgaris). Animal Behaviour. 1977; 25:720–725. https://doi.org/10.1016/0003-3472(77)90121-X

22. Bleckmann H. Reaction time and stimulus frequency in prey localization in the surface-feeding fish Aplocheilus lineatus. Journal of comparative physiology. 1980; 140(2):163–172. https://doi.org/10.1007/BF00606308

23. Nagy M, Akos Z, Biro D, Vicsek T. Hierarchical group dynamics in pigeon flocks. Nature. 2010; 464(7290):890–893. https://doi.org/10.1038/nature08891 PMID: 20376149

24. Nagy M, Vásárhelyi G, Pettit B, Roberts-Mariani I, Vicsek T, Biro D. Context-dependent hierarchies in pigeons. Proceedings of the National Academy of Sciences. 2013; 110(32):13049–13054. https://doi.org/10.1073/pnas.1305552110

25. Konzack M, McKetterick T, Ophelders T, Buchin M, Giuggioli L, Long J, et al. Visual analytics of delays and interaction in movement data. International Journal of Geographical Information Science. 2016; p.1–26. http://www.tandfonline.com/doi/full/10.1080/13658816.2016.1199806

26. Giuggioli L, McKetterick TJ, Holderied M. Delayed response and biosonar perception explain movement coordination in trawling bats. PLoS Comput Biol. 2015; 11(3):e1004089. https://doi.org/10.1371/journal.pcbi.1004089 PMID: 25811627

27. Agrillo C, Dadda M, Serena G, et al. Do fish count? Spontaneous discrimination of quantity in female mosquitofish[J]. Animal cognition. 2008; 11(3):495–503. https://doi.org/10.1007/s10071-008-0140-9 PMID: 18247068

28. Lecheval V, Jiang L, Tichit P, Sire C, Hemelrijk C, Theraulaz G. Domino-like propagation of collective U-turns in fish schools. Submitted. 0; 0(0):0. https://doi.org/10.1101/138628

29. Herbert-Read JE, Krause S, Morrell LJ, Schaerf TM, Krause J, Ward AJW. The role of individuality in collective group movement Proc. R. Soc. B. 2013; 280(1752):20122564. https://doi.org/10.1098/rspb.2012.2564 PMID: 23222452

30. McKetterick TJ, Giuggioli L. Exact dynamics of stochastic linear delayed systems: Application to spatio-temporal coordination of comoving agents Phys. Rev. E. 2014; 90(4):042135. https://doi.org/10.1103/PhysRevE.90.042135

31. Virágh C, Vásárhelyi G, Tarcai N, Szörényi T, Somorjai G, Nepusz T, Vicsek T. Flocking algorithm for autonomous flying robots Bioinspiration & biomimetics. 2014; 9(2):025012. https://doi.org/10.1088/1748-3182/9/2/025012

32. Codling E, Pitchford J, Simpson S. Group navigation and the "many-wrongs principle" in models of animal movement. Ecology. 2007; 88(7):1864–1870. https://doi.org/10.1890/06-0854.1 PMID: 17645033

33. Ward AJ, Herbert-Read JE, Sumpter DJ, Krause J. Fast and accurate decisions through collective vigilance in fish shoals. Proceedings of the National Academy of Sciences. 2011; 108(6):2312–2315. https://doi.org/10.1073/pnas.1007102108

34. Berdahl A, Torney CJ, Ioannou CC, Faria JJ, Couzin ID. Emergent sensing of complex environments by mobile animal groups. Science. 2013; 339(6119):574–576. https://doi.org/10.1126/science.1225883 PMID: 23372013

35. Strandburg-Peshkin A, Twomey CR, Bode NW, Kao AB, Katz Y, Ioannou CC, et al. Visual sensory networks and effective information transfer in animal groups. Current Biology. 2013; 23(17):R709–R711. https://doi.org/10.1016/j.cub.2013.07.059 PMID: 24028946

36. Arganda S, Pérez-Escudero A, de Polavieja GG. A common rule for decision making in animal collectives across species. Proceedings of the National Academy of Sciences. 2012; 109(50):20508–20513. https://doi.org/10.1073/pnas.1210664109

37. Bode NW, Franks DW, Wood AJ. Limited interactions in flocks: relating model simulations to empirical data. Journal of The Royal Society Interface. 2010; p. rsif20100397. http://rsif.royalsocietypublishing.org/content/early/2010/09/04/rsif.2010.0397.full

38. Camperi M, Cavagna A, Giardina I, Parisi G, Silvestri E. Spatially balanced topological interaction grants optimal cohesion in flocking models. Interface focus. 2012; 2(6):715–725. https://doi.org/10.1098/rsfs.2012.0026 PMID: 24312725

39. Ballerini M, Cabibbo N, Candelier R, Cavagna A, Cisbani E, et al. Interaction ruling animal collective behavior depends on topological rather than metric distance: evidence from a field study. Proceedings of the National Academy of Sciences. 2008; 105(4):1232–1237. https://doi.org/10.1073/pnas.0711437105

40. Ingle D. Two visual mechanisms underlying the behavior of fish. Psychologische Forschung. 1967; 31(1):44–51. https://doi.org/10.1007/BF00422385 PMID: 5605116

41. Guthrie D. Role of vision in fish behaviour. In: The behaviour of Teleost fishes. Springer; 1986. p. 75–113.

42. Eaton RC, Bombardieri RA, Meyer DL. The Mauthner-initiated startle response in teleost fish. Journal of Experimental Biology. 1977; 66(1):65–81. PMID: 870603

AUTEUR : Valentin Lecheval

TITRE : Experimental analysis and modelling of the behavioural interactions underlying the coordination of collective motion and the propagation of information in fish schools

DIRECTEUR DE THÈSE : Dr. Guy Theraulaz

CO-DIRECTRICE DE THÈSE : Prof. Dr. Charlotte K. Hemelrijk

SOUTENANCE : Le 5 décembre 2017, Université de Groningue, Groningue, Pays-Bas.

RÉSUMÉ

Les bancs de poissons sont des entités pouvant regrouper plusieurs milliers d'individus qui se déplacent de façon synchronisée, dans un environnement sujet à de multiples perturbations, qu'elles soient endogènes (e.g. le départ soudain d'un congénère) ou exogènes (e.g. l'attaque d'un prédateur). La coordination de ces bancs de poissons, décentralisée, n'est pas encore totalement comprise. Si les mécanismes sous-jacents aux interactions sociales proposés dans des travaux précédents reproduisent qualitativement les structures collectives observées dans la nature, la quantification de ces interactions et l'accord quantitatif entre ces mesures individuelles et les motifs collectifs sont encore rares dans les recherches récentes et forment l'objet principal de cette thèse.

L'approche de ce travail repose sur une étroite combinaison entre les méthodes expérimentales et de modélisation dans l'objectif de découvrir les liens entre les comportements individuels et les structures observées à l'échelle collective. Nous avons caractérisé et quantifié les interactions et mécanismes à l'origine, d'abord, de la coordination des individus dans les bancs de poissons et, ensuite, de la propagation d'information, quand le groupe subit une perturbation endogène ou exogène. Ces travaux, tous réalisés en étudiant la même espèce de poisson d'eau douce, le nez-rouge (*Hemigrammus rhodostomus*), ont mobilisé une diversité de méthodes expérimentales, d'analyses statistique et de modélisation, à l'interface de l'éthologie, de la physique statistique et des sciences computationelles.

INTITULÉ ET ADRESSE DES LABORATOIRES

Centre de Recherches sur la Cognition Animale, UMR 5169, CNRS - UPS, Bâtiment IVR3, 118 route de Narbonne, 31062 Toulouse, France.
Groningen Institute for Evolutionary Life Sciences, University of Groningen, P.O. Box 11103 9700 CC Groningen, Pays-Bas