

## Research



**Cite this article:** Laan A, Gil de Sagredo R, de Polavieja GG. 2017 Signatures of optimal control in pairs of schooling zebrafish. *Proc. R. Soc. B* **284**: 20170224. <http://dx.doi.org/10.1098/rspb.2017.0224>

Received: 3 February 2017

Accepted: 13 March 2017

**Subject Category:**

Behaviour

**Subject Areas:**

behaviour

**Keywords:**

collective behaviour, interaction rules, optimal control

**Authors for correspondence:**

Andress Laan

e-mail: [andres.laan@neuro.fchampalimaud.org](mailto:andres.laan@neuro.fchampalimaud.org)

Gonzalo G. de Polavieja

e-mail: [gonzalo.polavieja@neuro.fchampalimaud.org](mailto:gonzalo.polavieja@neuro.fchampalimaud.org)

fchampalimaud.org

Electronic supplementary material is available online at <https://dx.doi.org/10.5061/dryad.f0g32>.

# Signatures of optimal control in pairs of schooling zebrafish

Andress Laan, Raul Gil de Sagredo and Gonzalo G. de Polavieja

Champalimaud Neuroscience Programme, Champalimaud Center for the Unknown, Lisbon, Portugal

GGdP, 0000-0001-5359-3426

Animals moving in groups coordinate their motion to remain cohesive. A large amount of data and analysis of movement coordination has been obtained in several species, but we are lacking theoretical frameworks that can derive the form of coordination rules. Here, we examine whether optimal control theory can predict the rules underlying social interactions from first principles. We find that a control rule which is designed to minimize the time it would take a pair of schooling fish to form a cohesively moving unit correctly predicts the characteristics of social interactions in fish. Our methodology explains why social attraction is negatively modulated by self-motion velocity and positively modulated by partner motion velocity, and how the biomechanics of fish swimming can shape the form of social forces. Crucially, the values of all parameters in our model can be estimated from independent experiments that need not relate to measurement of social interactions. We test our theory by showing a good match with experimentally observed social interaction rules in zebrafish. In addition to providing a theoretical rationale for observed decision rules, we suggest that this framework opens new questions about tuning problems and learnability of collective behaviours.

## 1. Introduction

Autonomous systems in general, and animals in particular, often benefit from coordinating their behaviour with other similar agents [1,2]. In suitable conditions, highly complex and coordinated multi-agent behavioural displays may emerge through local social interactions. Many decades of research has uncovered the functional rationale underlying the various forms of collective behaviours, which can range from increased protection from predation and risk dilution [3] to optimal information use and accelerated learning [4].

An improved understanding of the functional roles of collective behaviour has opened the way to theoretical analyses that derive the rules of collective behaviour from first principles. One example originates from the study of decision-making in collectives. While early models were mostly phenomenological [5] (but see Condorcet theorem [6] and information cascades models [7]), more recently, Bayesian estimation theory has provided a gateway for predicting biologically observed decision rules through a treatment grounded in probability theory [8,9].

In the study of collective motion, phenomenological models still dominate over normative models. Many heuristic and data-driven models have provided highly successful descriptions of varied systems like fish schooling, pedestrian traffic flows, avian aerial displays and ant trail construction [1]. When attempts have been made to theoretically predict the finer details of social interactions through, for example, evolutionary computing [10], the underlying functional form of animal interactions has been based on heuristic assumptions or measurements [11–14] rather than theoretical considerations (but see [15]).

While phenomenological models use mathematical equations to fully specify agent behaviour, these equations are arrived at primarily through the aid of intuition or by fitting equations of motion to data. For new examples, the modeller is given many opportunities to change the form of the equations any time an

application shifts to a new species or domain. Normative models are different. Rather than specifying the governing equations directly, the theorist specifies an underlying logical framework from which equations can be derived for particular examples. When successful, normative models usually achieve greater generality, a better grounding of the models in experimental measurements and allow explicit connections to be made between mechanistic and functional explanations. An *ab initio* prediction of rules governing social interactions has remained an unsolved challenge.

Here we demonstrate that control theory makes predictions about interaction rules for pairs of schooling fish which are qualitatively consistent with experimentally measured social forces in zebrafish. In our control theory framework, we set out to solve the following problem. The sensory system of a fish has access to its own velocity, the velocity of its partner fish and their mutual distance. The locomotor system, on the other hand, can specify the acceleration produced by the muscles of the fish. What relationship between the sensory inputs of the fish and its locomotor output would enable the schooling fish to maximize the cohesion of the group?

We developed a control-theoretic treatment of fish schooling and we compared theoretical predictions against experiments in an effort to test the applicability of our framework on the domain of collective behaviour. While our treatment is grounded in the known biology of the fish locomotor system, we do make the simplifying assumption that fish move in a quasi-one-dimensional environment. This is done because in our set-up, zebrafish spend 80% of their time swimming parallel to the walls along approximately one-dimensional tracks and our access to two-dimensional data is thus more limited. In addition, to the best of our knowledge, the biomechanical constraints of turning in fish are less well understood, which has limited our ability to create a two-dimensional theory of turning.

## 2. Material and methods

### (a) Experimental set-up

Ten-month-old zebrafish from the five-dimensional strain were kept in a tank with 40 individuals (equal numbers of male and female) in a 10 L:14 D cycle from which pairs of animals were randomly sampled for schooling experiments. We recorded 13 pairs of zebrafish swimming in a  $35 \times 35$  cm tank with water depth of 5 cm for 30 min. The fish were recorded at 30 fps and the resulting videos were tracked using IDTRACKER [16] giving as output 26 individual trajectories. The trajectories were first smoothed with a 170 ms moving average filter (or 5 out of 30 frames per second). The  $x$ -components of velocity and acceleration at time  $t_i$  were calculated as  $v_x(t_i) = x(t_i) - x(t_{i-1})$  and  $a_x(t_i) = v_x(t_{i+1}) - v_x(t_i)$ , respectively, where  $t_i = i/30$ , with  $i$  the frame number and  $t_i$  the acquisition time of that frame in seconds.

### (b) Kinematics

For analysis of viscous drag, the local maxima and minima were extracted for each time series of the velocity waveform. The descending phases of the velocity trajectory were localized and each local descending region was fitted with an equation  $v_{\max,i} e^{-\alpha t}$ , where  $v_{\max,i}$  was the amplitude of velocity at the  $i$ th local maxima. The parameter  $\alpha$  was shared for all local descending regions and all fish. The value of  $\alpha$  was chosen such that the mean absolute deviation between the fit and the observed trajectories was minimized. For quadratic drag,

the same procedure was used except for the equation which was written as  $v_{t,i} = v_{\max,i}/(1 + v_{\max,i}\alpha t)$ .

### (c) Predictability of acceleration

The amount of unexplained linear variation was calculated as  $U = \sum_{t,i} |v_{t,i} - v_{\text{obs},t,i}|$ , where  $v_{t,i}$  was the predicted velocity at time  $t$  of descent episode  $i$  and  $v_{\text{obs},t,i}$  was the observed velocity. The total linear variation was calculated as  $T = \sum_{t,i} |v_{\text{med},i} - v_{\text{obs},t,i}|$ , where  $v_{\text{med},i}$  was the median velocity of the  $i$ th decay waveform. The total fraction of explained linear variation is  $f = (T - U)/T$ .

In order to analyse the predictability of acceleration using social information, we used an artificial neural network with eight input units representing the  $x$ - and  $y$ -coordinates and the  $x$ - and  $y$ -components of the velocity of the focal and partner fish at time  $t$ . All input signals were normalized to unit variance and had zero mean. The output signal was the  $x$ - and  $y$ -components of the acceleration at time  $t + 1$ . We used one hidden layer with 40 neurons and fitted the net individually for each fish after splitting its data randomly into a training set (70% of data), validation set (15%) and test set (15%). Nets were fitted with Levenberg–Marquard optimization using MATLAB neural networks toolbox. For predicting acceleration at time  $t + 1$  from its value at time  $t$ , ordinary linear regression was used.

### (d) Empirical forcemaps

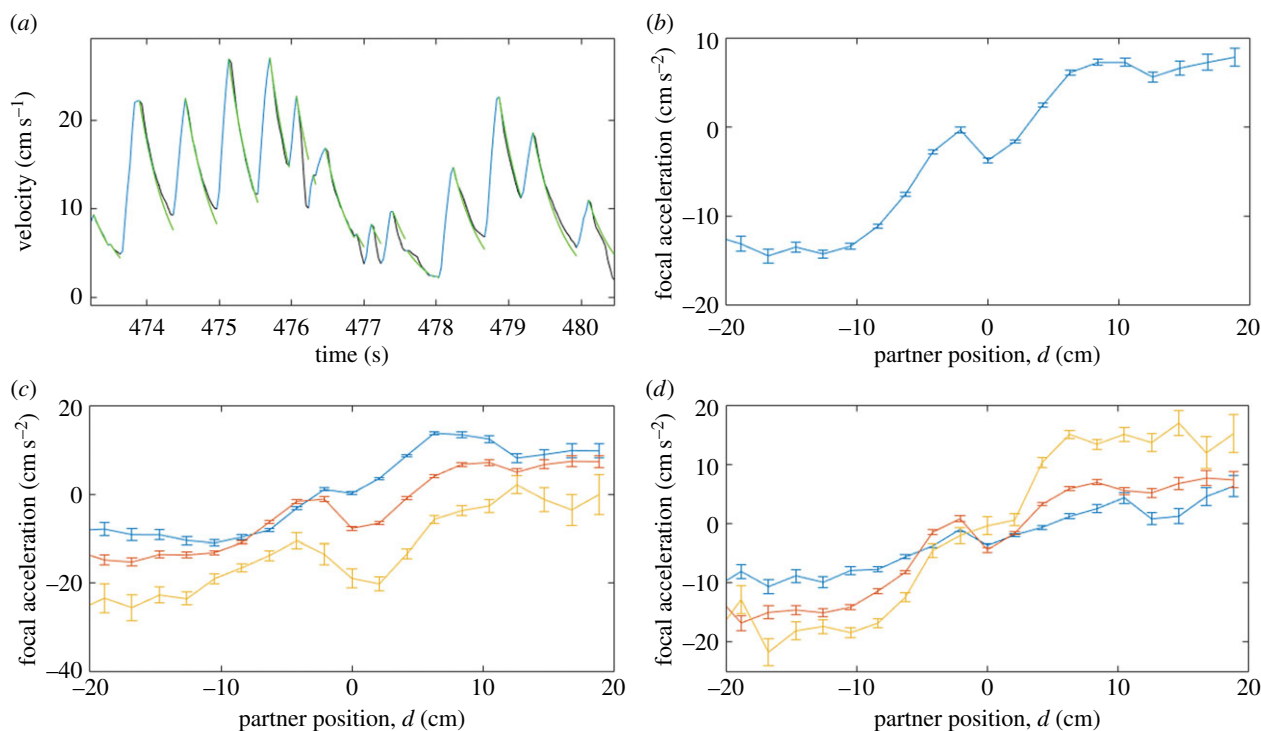
The forcemaps were extracted by analysing how the component of acceleration of the focal fish that was parallel to focal fish velocity depends the position of the other fish in a focal fish centric coordinate system (procedure as in [11]). The distance of the other fish was also measured using the position vector component that was parallel to the velocity of the focal fish. The velocity of the other fish was similarly projected onto the velocity of the focal fish and the amplitude of the projection was used for analysis.

The social forcemaps might potentially suffer from contaminating data from periods of non-social interaction. In order to correct for this potential bias, we analysed the expected number of instances in which a temporally shuffled dataset would generate observations within the variable range of our forcemaps. We found this number to be less than 25% of total observations in our experimental forcemaps. We corrected the amplitude of social forces by multiplying the observed forces by the number  $1/(1 - k)$ , where  $k$  is the total number of observations found in a shuffled forcemap divided by the total number of observations in the non-shuffled forcemap.

## 3. Results

### (a) Characteristics of zebrafish swimming

A control-theoretic treatment of fish locomotion begins with a characterization of the motor output of the fish. We first obtained a description of the forces at play in zebrafish swimming in one dimension. As in many fish species, zebrafish swimming is a highly structured temporal process [17]. Active periods of near-maximal acceleration alternate with periods of passive gliding where deceleration occurs due to viscous drag. This process generates a saw-tooth-like evolution of the velocity profile (figure 1a). We used the deceleration phase of the locomotor cycle in zebrafish to characterize the drag dynamics. We found that simple exponential decay accounted for 62% of the linear variations in speed. This performance measure was calculated across all decay waveforms for all fish using only a single free parameter shared across all



**Figure 1.** Analysis of zebrafish swimming. (a) Example of magnitude of velocity over time for one fish showing acceleration (blue) and deceleration phases (black). Also shown are exponential decay fits to decelerating phases (green). (b) One-dimensional social force as a function of the relative distance between focal fish and partner fish. (c) Modulation of social forces by focal fish velocity. Focal fish velocity ranges: 0–10.5 (blue), 10.5–21 (red) and 21–32.5 (yellow). (d) Modulation of social forces by partner fish velocity. Partner fish velocity ranges: 0–10.5 (blue), 10.5–21 (red) and 21–32.5 (yellow).  $n = 13$ . Bars are s.e.m.

traces, thus characterizing the average viscous drag on the body of a zebrafish (see ‘Kinematics’). This is consistent with viscous drag forces proportional to velocity, with acceleration

$$a_{\text{dec}} = -\alpha v, \quad (3.1)$$

and  $\alpha = 2.7 \text{ s}^{-1}$  in our case. A drag force quadratic in velocity gives a worse fit accounting for 50% of variation. The rising phase of the velocity profile is well described by an acceleration towards maximal velocity  $v_{\text{max}}$  as

$$a_{\text{acc}} = \alpha(v_{\text{max}} - v), \quad (3.2)$$

with  $v_{\text{max}} = 27 \text{ cm s}^{-1}$ . This value is consistent with the observed velocities, with 95% of them below this value. The few instances of observed velocities above  $v_{\text{max}} = 27 \text{ cm s}^{-1}$  might be accounted for by rare periods of anaerobic acceleration.

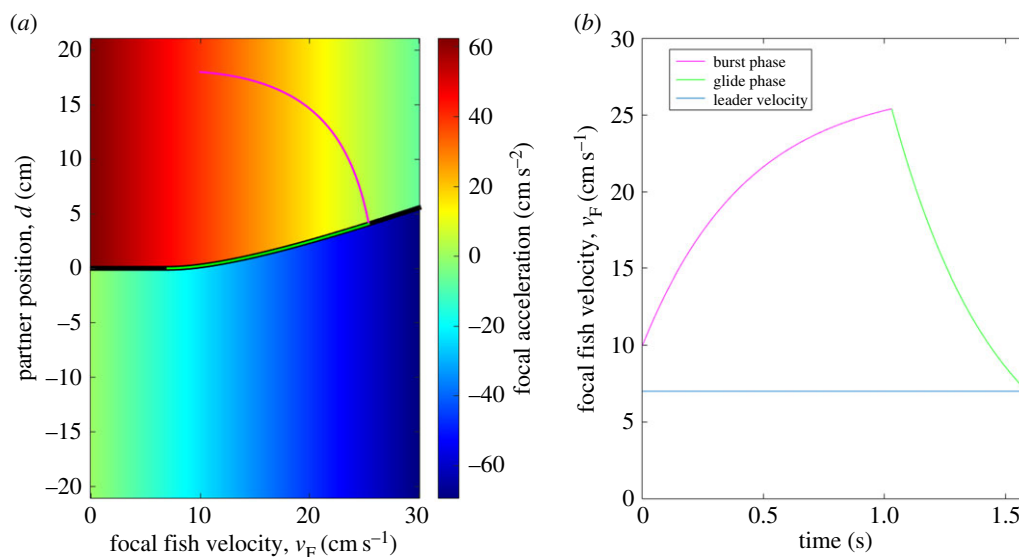
In experiments using pairs of zebrafish, active swimming dynamics still strongly shapes swimming patterns. The acceleration of a fish at time  $t$  predicts  $45 \pm 0.76\%$  of the linear variation in its acceleration at time  $t + 1$  (see ‘Predictability of acceleration’). By contrast, a neural network provided with spatial information and the velocity vectors of both the focal fish and its partner fish at time  $t$ , which characterizes the social information available to a fish, can explain only  $6.5 \pm 0.3\%$  of acceleration at time  $t + 1$  (see ‘Predictability of acceleration’). Our results contrast with a common view of social interactions, which concentrates on the influence of social parameters such as the relative inter-individual position, heading and velocity as determinants of future movements. In our experiments, the non-social part of the locomotor cycle is a better predictor of future acceleration values than social information even in social settings. The strong influence of non-social processes on behaviour is consistent with previous reports [18–20].

To extract the characteristics of social interactions, we used the forcemap approach [10,11]. It consists of computing the dependency of the mean social force (focal fish mean acceleration) on inter-animal distance and velocity of focal and partner. In our set-up, zebrafish spend 80% of their time swimming parallel to the walls along approximately one-dimensional tracks. We restricted our analysis to periods where zebrafish velocity vectors were highly parallel to each other, with differences in heading angles less than  $45^\circ$ , the distance between fish in the direction perpendicular to focal fish velocity vector less than 7 cm and the distance of fish in the parallel direction less than 20 cm (a total of 30% of all data were thus used for analysis).

From these approximately one-dimensional tracks, we measured how acceleration of a focal fish depends on its distance to its partner, the speed of the partner and the speed of the focal fish. A partner fish located in front (behind) of the focal fish generates a strong positive (negative) acceleration response in the focal fish (figure 1b). The velocity of the focal fish modulates the acceleration of the focal fish in a negative fashion: increasing focal fish velocity causes increased deceleration (figure 1c). This is consistent with passive viscous forces both reducing the maximal acceleration attainable to a fast-moving fish as well as increased viscous deceleration experienced during the gliding phase. By contrast, increasing partner fish velocity appears to promote increased acceleration of the focal fish, particularly when the partner fish is located in front of the focal individual (figure 1d).

### (b) Social interactions from optimal controller

A theoretical treatment of group swimming must also specify the objective of the fish. This objective must be expressible in terms of observed variables and should capture the biological



**Figure 2.** State space depiction of the optimal control rule. (a) Colourmap depicting how the acceleration of the focal fish changes as a function of the velocity of the focal fish and the inter-agent distance (position). For this figure, the velocity of the partner fish is held constant at  $7 \text{ cm s}^{-1}$ . The black line depicts the decision boundary. Above the boundary, fish engage in burst acceleration, below the boundary in gliding. An example trajectory of a focal fish is depicted in the state space with burst (purple) and gliding (green) phases. (b) Time-evolution of the velocity profile of the focal fish for the state space trajectory depicted in (a). Blue horizontal line: velocity of partner fish.

intuition that schooling fish are trying to move in a cohesive manner. We define a pair of fish as moving with maximal cohesion if both individuals are moving with identical velocities and are positioned as close to each other as physically possible. We next derived an approximately optimal controller which tries to maintain maximal cohesion in order to compare whether such a controller might explain the observed characteristics of zebrafish social interactions.

Our controller takes as input the relative position of the partner fish, as well as velocities of both the focal and partner fish, and it outputs a dynamic decision which specifies whether the focal fish should glide or accelerate. The controller is designed to optimize the time it takes to bring two fish into a maximally cohesive state from any initial state. The minimum time objective was chosen partly because it might minimize predation risk by allowing a dispersed group of fish to arrive at a maximally cohesive and presumably safest group configuration in the minimum amount of time. A second reason behind our choice of cost function is mathematical. Minimum time controllers are easy to optimize. It remains possible that other cost functions such as minimization of inter-fish distance over time will also produce behaviours consistent with data. The exploration of alternative cost functions is left as future work.

The Pontryagin minimum principle specifies the bang–bang controller as the optimal controller for our scenario [17,21] (while this is a standard result, we have also added the sketch of a proof for our particular problem to the electronic supplementary material). A bang–bang controller prescribes that if an agent should decide to accelerate, it should do so with maximal intensity. Deceleration decisions should likewise use only passive gliding to attain maximal possible deceleration.

The decision-making of a zebrafish is divided into a simple classification problem: accelerate maximally or glide. From these arguments, we obtain that the amplitude of deceleration should be equal to  $a_F = a_{\text{dec}} = -\alpha v_F$  and the amplitude of

acceleration should be given by  $a_F = a_{\text{acc}} = \alpha(v_{\text{max}} - v_F)$ , where  $v_F$  is the velocity of the focal individual and  $v_{\text{max}}$  is the maximal possible speed obtainable through aerobic swimming.

The decision to glide or accelerate can then be formalized as a three-dimensional decision problem with inputs  $v_F$  (focal fish velocity),  $v_P$  (partner fish velocity) and  $d = x_{\text{partner}} - x_{\text{focal}}$  (inter-animal distance). We derive the location of the decision boundary from the following standard argument: if a zebrafish is located on the decision boundary and the partner fish is expected to maintain a constant velocity  $v_P$  (see ‘Discussion’ about possible caveats associated with this assumption), then passive gliding will eventually bring the focal fish to a maximally cohesive state (same point in space and same velocity). In the electronic supplementary material, we derive the equation describing this decision surface as

$$T = d - \frac{1}{\alpha} \left[ v_F - v_P - v_P \ln \frac{v_F}{v_P} \right] = 0. \quad (3.3)$$

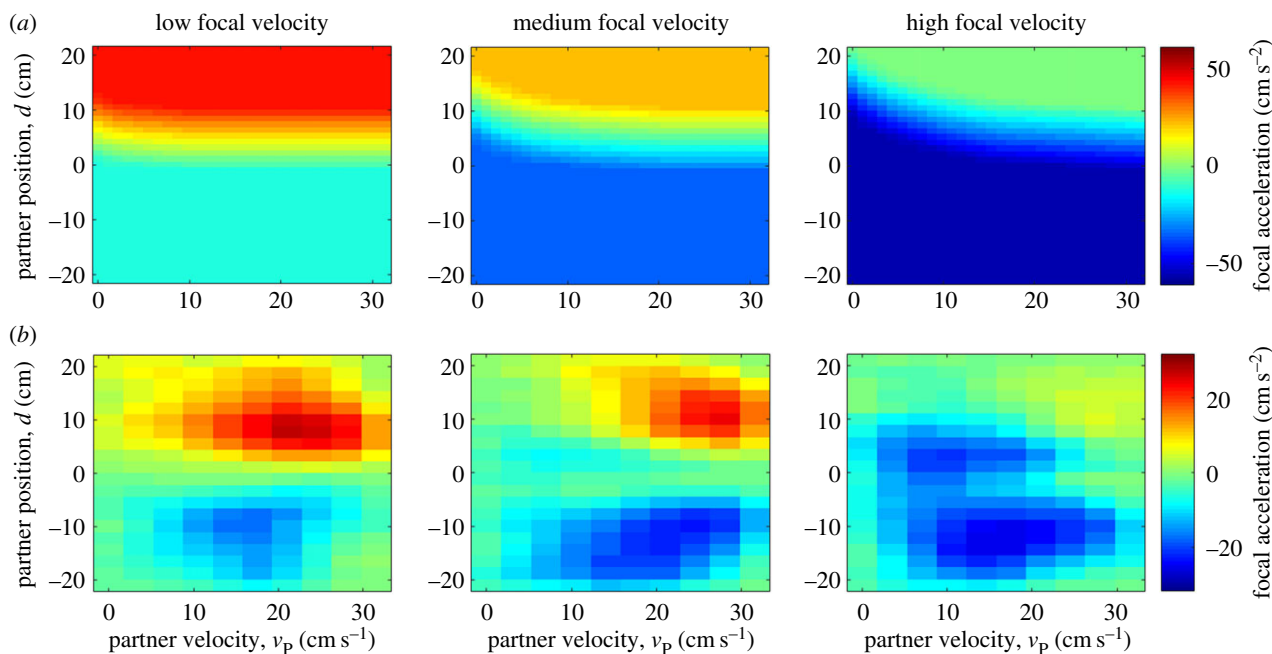
The resulting decision rule then corresponds to fish accelerating,  $a_F = a_{\text{acc}}$ , at one side of the decision surface,  $T > 0$ , or decelerating,  $a_F = a_{\text{dec}}$ , at the other side of the decision surface,  $T < 0$ . Additionally, when it is the focal fish that is in front of the partner,  $d < 0$ , it glides until  $d > 0$ . In summary, the rule is completely characterized by the following equations describing how the focal agent should accelerate:

$$\text{if } d < 0 \Rightarrow a_F = a_{\text{dec}} \quad (3.4)$$

$$\text{else if } T > 0 \Rightarrow a_F = a_{\text{acc}} \quad (3.5)$$

$$\text{and else } a_F = a_{\text{dec}}. \quad (3.6)$$

Our decision rule is fundamentally a three-dimensional process, but it can be visualized as a series of two-dimensional sections. In figure 2a, we depict one such cross section, where we have fixed the value of the partner fish velocity in order to illustrate how changes in focal fish velocity and the inter-animal distance influence acceleration



**Figure 3.** Comparison of optimal controller and zebrafish social forces. (a) Theoretical predictions of three-dimensional decision maps. On each colourmap, we plot for a restricted range of values of focal fish velocity how acceleration of zebrafish varies as a function of partner fish velocity and the relative distance between focal and partner fish. Velocity ranges from left to right were 0–10.5, 10.5–21 and 21–32.5 (cm s<sup>-1</sup>). For the calculations, we have also added noise to the sensory system of the focal fish which causes an imprecise estimate of the location of the partner fish (see the electronic supplementary material, figure S1 for noiseless predictions). (b) Same as (a) but for zebrafish experiments ( $n = 13$  pairs).

decisions. The decision boundary is visualized as a black line, which separates the acceleration zone from the deceleration zone. As can be seen from the figure, increasing focal fish velocity increases deceleration, as is apparent from the progressively deeper blue tones as we move from the left to the right on the graph.

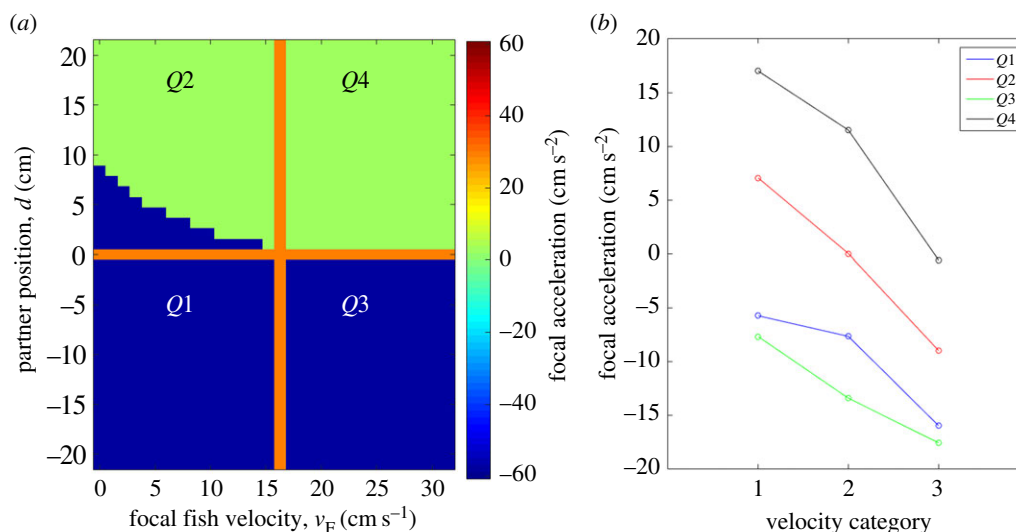
One advantage of our two-dimensional visualization is that focal fish behaviour can be visualized as trajectories in the state space graph. In figure 2*b*, we have depicted one such example trajectory in the state space. During the burst phase (purple), the focal individual increases its velocity and decreases its distance relative to the partner fish until it reaches the decision boundary. After reaching the decision boundary, the focal individual switches to gliding mode (green) and moves along the decision boundary until its position and velocity match with its partner. The resulting evolution of the velocity of the focal fish is depicted in figure 2*b*, with a saw-tooth-like velocity waveform from the switch from bursting to gliding.

We present an alternative depiction of the structure of our optimal controller, where we fix to a limited range the velocity values of the focal fish in order to see the effects of varying partner fish velocity (figure 3*a*). This was modelled with sensory noise added in the estimation of interfish distance (see the electronic supplementary material, methods 5.3 for details of noise modelling and comparison with noise-free predictions). We also provide analogous plots of the experimentally measured zebrafish interactions (figure 3*b*). The optimal controller has three features that are qualitatively consistent with the measured decision maps. First, the amplitude of the deceleration and acceleration zones are negatively modulated by the focal zebrafish velocity. Second, large values of partner fish velocities activate acceleration behaviours when the partner fish is located in front of the focal

fish. Third, the transition zone between the acceleration and deceleration zones has a negative slope in the  $d-v_{\text{partner}}$  axis.

Overall, both the biological controller and the optimal controller have a simple intuitive structure. If the partner fish is located in front of the focal fish and moving faster than the focal fish, then the focal fish should accelerate. This behaviour leads to an increase in its speed and will allow it to catch up with the partner fish. The acceleration should continue until the focal fish reaches a sufficient speed and position such that a switch to passive gliding is expected to bring it into a maximally cohesive state with its partner. If the speed of the focal fish is much greater than the speed of the partner fish, then the focal fish should always glide. Gliding causes the passive drag of water to bring its velocity closer to the velocity of its partner and the inter-individual distance will simultaneously shrink as well, since the greater speed of the focal fish relative to the partner will cause the focal fish to approach the partner. Finally, if the focal fish is located in front of the partner fish, then due to the blind angle, a precise estimation of the position and velocity of the partner fish is not possible and the best decision is to glide passively until the partner fish overtakes the position of the focal fish.

This qualitative consistency can be seen in figure 4. We have divided the forcemaps into four quadrants (figure 4*a*) and plotted how the average acceleration in each quadrant varies with focal fish velocity (figure 4*b*). Consistent with theory, acceleration in all four quadrants is negatively modulated by focal fish velocity. The relative ranking of the intensity of acceleration in all quadrants is also qualitatively consistent with theory. Quadrants Q1 and Q3 have the lowest values of acceleration, because they correspond to situations where the partner is located behind the focal fish and the focal fish is always decelerating. Quadrants Q2 and Q4 partially overlap with acceleration regions so they show



**Figure 4.** Analysis of quadrants of experimental control surface. The two-dimensional forcemaps were divided into four quadrants and the average acceleration was computed for each for the three groupings of focal fish velocity ( $n = 13$  pairs): group 1, 0–10.5; group 2, 10.5–21; group 3, 21–32.5 (cm s<sup>-1</sup>). Consistent with theory, all quadrants are negatively modulated by focal fish velocity. The relative ranking of the acceleration amplitudes is also largely consistent with theory, which predicts that  $Q4 > Q2 > Q3 \approx Q1$ .

higher acceleration values. Because quadrant Q4 corresponds to higher values of the partner fish velocity, it has higher average acceleration values than quadrant Q2.

While the controller map is qualitatively consistent with observed zebrafish behaviour, there are quantitative differences. Notably, the amplitudes of the acceleration and deceleration zones are approximately 40% weaker than predicted by theory and the transitions between acceleration and deceleration zones are less abrupt. The sources of this discrepancy are further addressed in the next section, which details a simulation approach to study contributions of non-social behaviour to the empirical forcemaps.

One prominent feature in both the optimal controller and biologically measured controllers is velocity-dependent modulation of social interactions. In the electronic supplementary material, we show mathematically that velocity-dependent modulation of social interactions is a necessary characteristic of fish social interactions given other features known about fish swimming control systems such as the relative dominance of frontal acceleration zones over caudal deceleration zones as well as the use of egocentric representations of decision rules. Control theoretic analysis is thus more generally able to successfully explain many qualitative features of social interactions in fish collectives.

### (c) Closing the loop through simulations

We next attempted to close the loop by simulating fish following the optimal control rule to see if we could find a match between the most relevant summary statistics of fish locomotion in experiments and simulation.

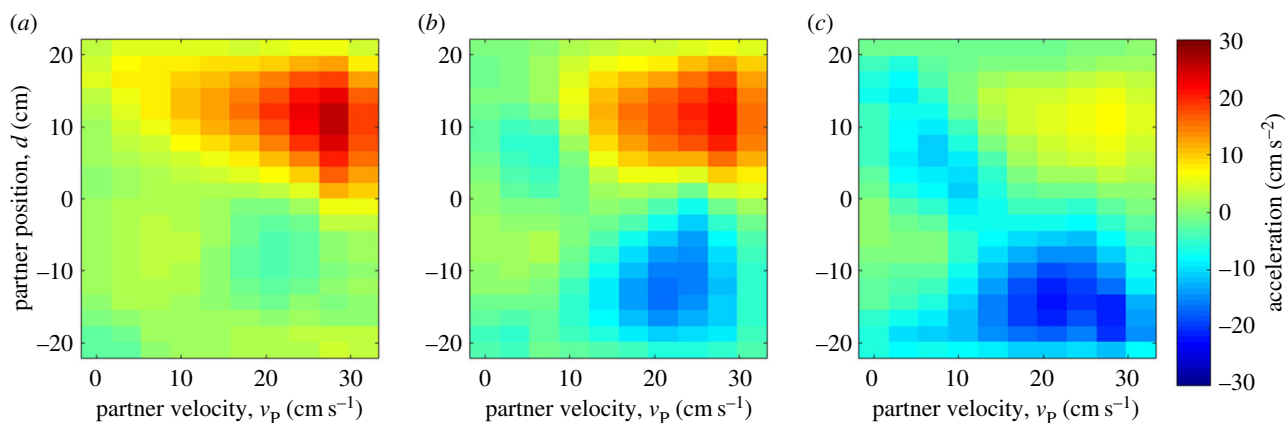
Closing the loop required making some assumptions about the non-social behaviour of fish. There is a leader fish, who moves non-socially (it does not respond to the movements of the follower) in a burst and glide swimming pattern. We tested for leadership in zebrafish by examining velocity vector cross-correlations, which showed that zebrafish pairs exhibit both asymmetric and alternating leadership patterns (electronic supplementary material, figure S2). The follower

fish obeyed our control rule except for non-social periods where it exhibited burst and glide behaviour (see the electronic supplementary material). The probability of the follower being in a non-social state was 0.2 and the duration of a non-social episode was 15 s on average (similar to the 5–15 s duration of shoal cohesion oscillations in zebrafish [18]).

To calculate the forcemaps, we summed the occupancies and accelerations of both the leader and the follower fish in each bin, and then calculated the averaged forcemaps by dividing the net acceleration with the net occupancy. This is important because zebrafish leadership can switch unpredictably during the recording session and our experimental forcemaps contain contributions from both the leader and the follower fish. We subsequently extracted empirical forcemaps from simulated trajectories and found them to have a good qualitative match with experimental measurements (figure 5; cf. figure 3).

We note the following key features: even though our control rule generated peak acceleration values of 60 cm s<sup>-2</sup>, the simulated forcemaps have peak acceleration amplitudes of 25 cm s<sup>-2</sup>, which is much closer to the experimentally measured peak values of 26 cm s<sup>-2</sup>. It appears that non-social behaviour dilutes the contribution of the control rule both in simulation and in experiments. A second feature of the simulations which is qualitatively consistent with the experimental forcemaps is the non-uniform appearance of acceleration and deceleration zones. For example, in the medium- and low-velocity maps in both experiments and simulations, regions of peak acceleration are concentrated in the lower part of quadrant 4. Such peaks are absent from the theoretical approximations and again suggest the importance of the influence of non-social periods of activity on the maps. Our simulations also closely reproduced key summary statistics of fish velocity (simulations:  $13.5 \pm 5.8$  cm s<sup>-1</sup> and  $12 \pm 6.5$  cm s<sup>-1</sup>).

Even in our simulations a certain mismatch remains with experiments. First, the simulated deceleration zones tend to have smaller amplitudes. This may be caused by an overly simplistic model of non-social fish behaviour, which lacks known features like homebase behaviour [22], sporadic



**Figure 5.** Forcemaps extracted from simulated trajectories. (a–c) Same as for figure 3 but for simulated fish trajectories (see main text for details) instead of ideal controller. Velocity ranges were (a) 0–10.5, (b) 10.5–21 and (c) 21–32.5 ( $\text{cm s}^{-1}$ ).

aggression [16], and switching between fearful and more pacific internal states [23]. Second, the median interfish distance in our simulations was 4.6 cm while in experiments this measure was 5.6 cm. This is easily explained because in experiments fish often spend extended periods exploring the arena. Overall, our simulations reproduced the most salient experimental features of zebrafish schooling.

#### 4. Discussion

Early phenomenological descriptions of fish schooling used self-propelled particle models where agents kept a constant velocity [1,13]. Experimental measurements of schooling recently found that changes in the magnitude of velocity were integral to the process of schooling [11,12]. The same studies also first described velocity-dependent modulation of social forces although the functional reason behind this modulation was not discussed. Our theoretical framework provides a strong rationale behind this phenomena. Velocity-dependent modulation of interactions is a necessary precondition for stable schooling in fish species where acceleration and deceleration zone amplitudes have a rostral-caudal asymmetry and the frontal acceleration zone has a stronger amplitude than the caudal deceleration zone. A more precise coordination of velocity-dependent forces can even be used to construct a near-optimal controller, and signatures of such a controller appear present in the experimental data for zebrafish.

The results of our study suggest many new avenues of experimental research. One exciting approach concerns inter-species comparison of schooling. Different fish species exhibit different locomotor dynamics and it will be important to test the effects of altered locomotion on the characteristics of forcemaps. The characteristics of forcemaps are also influenced by the behaviour of fish during non-social episodes. It would be interesting to apply our approach to fish species where leadership is less dynamic and more asymmetric than in zebrafish. In such species, where leaders and followers are easier to separate, the leader and follower maps might look substantially different [24]. Comparing the predictions of theory with experiments in these species represents a further important test of our theory.

The presence of signatures of optimal control in simple situations where pairs of zebrafish interact in a quasi-one-dimensional geometry raises multiple questions. One is

how the framework extends to two- and three-dimensional scenarios and to larger group sizes. Another is how fish estimate the variables needed for real-time computation of the decision boundary from their retinal image. A third is how these variables are re-estimated during development, as changes in fish body shape modify viscous drag.

We believe that a joint solution to these problems can be found through the methodology of deep multi-agent reinforcement learning [25,26]. Multi-agent reinforcement learning can learn to solve complex control problems in situations where analytical methods of traditional control theory are no longer tractable. Extensions of our model to multi-agent three-dimensional scenarios would be one such example situation, although this extension will also require modifications to our cost function, which is currently only stated for the case of pairwise interactions. Deep reinforcement learning also provides a solution to the sensory estimation problem, because it can learn useful policies directly from high-dimensional data such as retinal activity patterns. As deep reinforcement learning is fundamentally an adaptive method which responds to changes in the environment or the body of the zebrafish, it also provides a natural solution to the problem of tuning the decision boundary [15,27] over developmental time scales.

Multi-agent deep reinforcement learning is not only conceptually attractive [28], but also inherently biologically plausible, because the neural apparatus needed to implement its computations is present in vertebrate brains [29]. One potential challenge to the use of deep learning in schooling is the large hunger for data often needed to train such systems. Our preliminary computational experiments indicate that this is not a problem for the case of schooling in pairs, because an artificial neural network composed of just 200 neurons is capable of learning competent schooling policies using data gathered in under 30 min of real-time schooling behaviour.

The present work has outlined how control theory can provide a simple mathematical description of zebrafish schooling in pairs. We believe that the presence of simple signatures of optimal control in relatively simple scenarios is not accidental but a sign of a more complex controller, which (among many other problems) is also able to find near-optimal ways to coordinate collective behaviour in simple scenarios. An important future extension of this work would explore whether methods such as deep reinforcement learning could provide a unified description of the diversity of social behaviours seen in different species and in different environments.

**Ethics.** Experiments were approved by the Champalimau Ethics Committee (Champalimau Foundation, Portugal).

**Data accessibility.** Data are available from Dryad Digital Repository [30].

**Author's contributions.** A.L. and G.G.d.P. conceived the idea, A.L. and G.G.d.P. decided methodology, R.G.d.S. and G.G.d.P. designed the experimental set-up, R.G.d.S. obtained video recordings, A.L. wrote analysis software, A.L. and G.G.d.P. analysed and modelled data, and A.L. and G.G.d.P. wrote the manuscript.

**Competing interests.** Authors declare no competing interests exist.

**Funding.** We acknowledge funding from the Champalimau Foundation (to G.G.d.P.) and from Fundação para a Ciência e Tecnologia PTDC/NEU-SCC/0948/2014 (to G.G.d.P.).

**Acknowledgments.** We thank Robert Hinz for discussions, Mattia G. Bergomi for independent computational experiments with reinforcement learning and Ana Catarina Certal for animal care (Champalimau fish facility). Francisco Heras and Mattia G. Bergomi helped with a critical reading of the final manuscript. We are indebted to the anonymous reviewers for suggestions on how to improve the readability of the manuscript and to include a simulation study in our results.

## References

- Vicsek T, Zafeiris A. 2012 Collective motion. *Phys. Rep.* **517**, 71–140. (doi:10.1016/j.physrep.2012.03.004)
- Lopez U, Gautrais J, Couzin ID, Theraulaz G. 2012 From behavioural analyses to models of collective motion in fish schools. *Interface Focus* **2**, 20120033. (doi:10.1098/rsfs.2012.0033)
- Roberts G. 1996 Why individual vigilance declines as group size increases. *Anim. Behav.* **51**, 1077–1086. (doi:10.1006/anbe.1996.0109)
- Laland KN, Atton N, Webster MM. 2011 From fish to fashion: experimental and theoretical insights into the evolution of culture. *Phil. Trans. R. Soc. B* **366**, 958–968. (doi:10.1098/rstb.2010.0328)
- Sumpter DJT. 2010 *Collective animal behavior*. Princeton, NJ: Princeton University Press.
- Boland PJ. 1989 Majority systems and the condorcet jury theorem. *The Statistician* **38**, 181–189. (doi:10.2307/2348873)
- Banerjee AV. 1992 A simple model of herd behavior. *Q. J. Econ.* **107**, 797–817. (doi:10.2307/2118364)
- Pérez-Escudero A, de Polavieja GG. 2011 Collective animal behavior from Bayesian estimation and probability matching. *PLoS Comput. Biol.* **7**, e1002282. (doi:10.1371/journal.pcbi.1002282)
- Arganda S, Pérez-Escudero A, de Polavieja GG. 2012 A common rule for decision making in animal collectives across species. *Proc. Natl Acad. Sci. USA* **109**, 20 508–20 513. (doi:10.1073/pnas.1210664109)
- Hein AM, Rosenthal SB, Hagstrom GI, Berdahl A, Torney CJ, Couzin ID. 2015 The evolution of distributed sensing and collective computation in animal populations. *eLife* **4**, e10955. (doi:10.7554/eLife.10955)
- Katz Y, Tunstrom K, Ioannou CC, Huepe C, Couzin ID. 2011 Inferring the structure and dynamics of interactions in schooling fish. *Proc. Natl Acad. Sci. USA* **108**, 18 720–18 725. (doi:10.1073/pnas.1107583108)
- Herbert-Read JE, Perna A, Mann RP, Schaerf TM, Sumpter DJT, Ward AJW. 2011 Inferring the rules of interaction of shoaling fish. *Proc. Natl Acad. Sci. USA* **108**, 18 726–18 731. (doi:10.1073/pnas.1109355108)
- Couzin ID, Krausew J, James R, Ruxtony GD, Franksz NR. 2002 Collective memory and spatial sorting in animal groups. *J. Theor. Biol.* **218**, 1–11. (doi:10.1006/jtbi.2002.3065)
- Gautrais J, Ginelli F, Fournier R, Blanco S, Soria M, Chaté H, Theraulaz G. 2012 Deciphering interactions in moving animal groups. *PLoS Comput. Biol.* **8**, e1002678. (doi:10.1371/journal.pcbi.1002678)
- Bialek WS. 2012 *Biophysics: searching for principles*. Princeton, NJ: Princeton University Press.
- Pérez-Escudero A, Vicente-Page J, Hinz RC, Arganda S, de Polavieja GG. 2014 idTracker: tracking individuals in a group by automatic identification of unmarked animals. *Nat. Methods* **11**, 743–748. (doi:10.1038/nmeth.2994)
- Weihls D. 1974 Energetic advantages of burst swimming of fish. *J. Theoret. Biol.* **48**, 215–229. (doi:10.1016/0022-5193(74)90192-1)
- Miller NY, Gerlai R. 2008 Oscillations in shoal cohesion in zebrafish (*Danio rerio*). *Behav. Brain Res.* **193**, 148–151. (doi:10.1016/j.bbr.2008.05.004)
- Ginelli F, Peruani F, Pillot MH, Chaté H, Theraulaz G, Bon R. 2015 Intermittent collective dynamics emerge from conflicting imperatives in sheep herds. *Proc. Natl Acad. Sci. USA* **112**, 12 729–12 734. (doi:10.1073/pnas.1503749112)
- Hinz RC, de Polavieja GG. 2017 Ontogeny of collective behavior reveals a simple attraction rule. *Proc. Natl Acad. Sci. USA* **114**, 2295–2300. (doi:10.1073/pnas.1616926114)
- Kirk DE. 2004 *Optimal control theory: an introduction*. New York, NY: Dover Publications.
- Stewart A, Cachat J, Wong K, Gaikwad S, Gilder T, DiLeo J, Chang K, Utterback E, Kalueff AV. 2010 Homebase behavior of zebrafish in novelty-based paradigms. *Behav. Process.* **85**, 198–203. (doi:10.1016/j.beproc.2010.07.009)
- Miller N, Gerlai R. 2012 From schooling to shoaling: patterns of collective motion in Zebrafish (*Danio rerio*). *PLoS ONE*. **7**, e48865. (doi:10.1371/journal.pone.0048865)
- Schaerf TM, Herbert-Read JE, Myerscough MR, Sumpter DJT, Ward AJW. 2016 Identifying differences in the rules of interaction between individuals in moving animal groups. *arXiv* 1601.08202.
- Lillicrap TP, Hunt JJ, Pritzel A, Heess N, Erez T, Tassa Y, Silver D, Wierstra D. 2015 Continuous control with deep reinforcement learning. *arXiv* 1509.02971.
- Tampuu A, Matisen T, Kodelja D, Kuzovkin I, Korjus K, Aru J, Aru J, Vicente R. 2015 Multiagent cooperation and competition with deep reinforcement learning. *arXiv* 1511.08779.
- Seung HS. 1996 How the brain keeps the eyes still. *Proc. Natl Acad. Sci. USA* **93**, 13 339–13 344. (doi:10.1073/pnas.93.23.13339)
- Biro D, Sasaki T, Portugal SJ. 2016 Bringing a time–depth perspective to collective animal behaviour. *Trends Ecol. Evol.* **31**, 550–562. (doi:10.1016/j.tree.2016.03.018)
- Schultz W, Dayan P, Montague PR. 1997 A neural substrate of prediction and reward. *Science (New York, NY)* **275**, 1593–1599. (doi:10.1126/science.275.5306.1593)
- Laan A, Gil de Sagredo R, de Polavieja GG. Data from: Signatures of optimal control in pairs of schooling zebrafish. Dryad Digital Repository. (doi:10.5061/dryad.f0g32)