

Emergence of in-line swimming patterns in zebrafish pairs

Maurizio Porfiri^{1,2*}, Mert Karakaya¹, Raghu Ram Sattanapalle¹ and Sean D. Peterson³

¹Mechanical and Aerospace Engineering Department, New York University Tandon School of Engineering, 6 MetroTech Center, Brooklyn, NY, 11201, USA

²Biomedical Engineering Department, New York University Tandon School of Engineering, 6 MetroTech Center, Brooklyn, NY, 11201, USA

³Mechanical and Mechatronics Engineering, University of Waterloo, 200 University Ave W, Waterloo, Ontario, N2L 3G1, Canada

*Corresponding author. E-mail: mporfiri@nyu.edu

Received: 8 November 2020; **Revised:** XX Month 2020; **Accepted:** XX Month 2021

Keywords: collective behavior, swimming/flying, vortex interactions

Abstract

Mathematical models promise new insight into the mechanisms underlying the emergence of collective behavior in fish. Here, we establish a mathematical model to examine collective behavior of zebrafish, a popular animal species in preclinical research. The model accounts for social and hydrodynamic interactions between individuals, along with the burst-and-coast swimming style of zebrafish. Each fish is described as a system of coupled stochastic differential equations, governing the time-evolution of their speed and turn rate. Model parameters are calibrated using experimental observations of zebrafish pairs swimming in a shallow water tank. The model successfully captures the main features of the collective response of the animals, by predicting their preference to swim in-line, with one fish leading and the other trailing. During in-line swimming, the animals share the same orientation and keep a distance from each other, due to hydrodynamic repulsion. Hydrodynamic interaction is also responsible for an increase in the speed of the pair swimming in-line. By linearizing the equations of motion, we demonstrate local stability of in-line swimming to small perturbations for a wide range of model parameters. This structural property is unique to this schooling pattern, whereby the same analysis applying to a side-by-side configuration would yield instability for any parameter choice. Mathematically-backed results presented herein support the application of dynamical systems theory to unveil the inner workings of fish collective behavior.

Impact Statement

Fish collective behavior is an open area of research that continues to attract the interest of a broad scientific community and the curiosity of the general public. How and why do fish align their bodies, synchronize their motion, and swim close to each other? And how do they choose one pattern over another? In search for some answers, we establish an experimentally-validated mathematical model for the collective behavior of zebrafish, a popular species in laboratory research. The model accounts for social and hydrodynamic interactions between animals, and it incorporates key features of the burst-and-coast swimming style of zebrafish. In agreement with experimental observations, the model predicts a strong preference of zebrafish to swim in-line, with one fish leading the other. We mathematically demonstrate that the emergence of this collective pattern is related to its local stability, such that zebrafish interactions will dynamically compensate for any small perturbation to in-line swimming. Interestingly, the stability of in-line swimming is modulated by the distance between the animals, due to short-range hydrodynamic repulsion. Just as the proposed model is expected to find application in preclinical research on zebrafish, the presented analytical tools create an important connection between stability and collective behavior.

1. Introduction

The majority of fish species live part of their lives in a group; this affords several advantages, including enhanced ability to escape from predators, searching for food, and finding correct migratory routes (Larsson, 2012). Common to life in a group is swimming together in a school, where fish synchronize their motion in a crystallized formation of tightly-packed individuals (Miller and Gerlai, 2011). The mechanisms underlying the formation of these schooling patterns and their potential hydrodynamic implications have long been the subject of intense debate, which is yet to be resolved (Weihs, 1973; Partridge and Pitcher, 1979; Ashraf et al., 2017).

While experimental research is the cornerstone against which new hypotheses must be tested, mathematical modeling offers a promising approach for detailing the inner workings of collective behavior. Since the seminal work of Aoki (1982), a number of modeling approaches have been proposed to study collective behavior (Vicsek and Zafeiris, 2012). Reaching beyond agent-based models implementing behavioral rules in discrete time, several authors have proposed modeling approaches based on stochastic differential equations, including Butail et al. (2016); Calovi et al. (2014); Calovi et al. (2015); Zienkiewicz et al. (2015a); Zienkiewicz et al. (2015b).

Particularly promising is the recent model by Filella et al. (2018), which considers, for the first time, the presence of hydrodynamic interactions between swimming fish. These interactions are modelled by associating each fish with a vortex dipole, which encapsulates both self-propulsion and the velocity field generated by a swimming fish. The velocity field generated by one fish induces advection and rotation on other animals in the vicinity, and vice versa, thereby providing a mechanism for bidirectional hydrodynamic interactions between fish. While paving the way toward the mathematical integration of social and hydrodynamic interactions, the study was not grounded in experimental observations and all of the claims were based on numerical simulations.

Here, we seek to address these limitations by tailoring the model of Filella et al. (2018) to real data on pairs of zebrafish, a species of choice in the study of collective behavior in laboratory settings (Orger and de Polavieja, 2017; Kalueff et al., 2013). To achieve this goal, we include realistic turn rate dynamics that capture the burst-and-coast swimming style of zebrafish (Kalueff et al., 2013). We calibrate the model on real data from Zienkiewicz et al. (2015a) and demonstrate its predictive power on multiple measures of collective behavior. Going beyond numerical simulations, we offer analytical insight into the model by elucidating the stability of two different schooling patterns: in-line and side-by-side swimming. Both patterns have been experimentally and numerically studied in the technical literature on collective behavior, but their stability has never been analytically investigated (De Bie et al., 2020; Kato et al., 2004; Laan et al., 2017; Perna et al., 2014).

The rest of the paper is organized as follows: §2 introduces the modelling framework; §3 offers analytical insight into the stability of schooling patterns; §4 discusses calibration of model parameters on experimental results; §5 presents and discusses the results; and §6 summarizes the main findings of the work and identifies avenues of future research.

2. Modeling collective behavior of a zebrafish pair

We model zebrafish as self-propelled bodies swimming in an unbounded two-dimensional plane. At time t , fish f ($f \in \{1, 2\}$) is identified by the position of its centroid $\vec{r}_f(t)$ and swimming direction $\hat{v}_f(t)$ with respect to a global reference frame (Fig. 1(a)). Working with a Cartesian coordinate system with unit vectors $\hat{i}$ and $\hat{j}$ for X and Y , respectively, we write $\vec{r}_f(t) = x_f(t)\hat{i} + y_f(t)\hat{j}$ and $\hat{v}_f(t) = \cos \theta_f(t)\hat{i} + \sin \theta_f(t)\hat{j}$ (Fig. 1(a)).

Each fish is assumed to have a constant self-propulsion speed v_f . When swimming in a background flow, the velocity of the animal in the global reference frame is the superposition of self-propulsion

and advection. In particular, swimming in a pair causes the speed of fish f to dynamically change with respect to the fixed coordinate system due to the flow created by the other individual in the pair. Specifically, the velocity of the focal fish f is given by (Filella et al., 2018)

$$\dot{\vec{r}}_f(t) = v_f \hat{v}_f(t) + \vec{U}_f(\vec{r}_f(t)), \quad (1)$$

where a superimposed dot means a time derivative and $\vec{U}_f(\vec{r}_f(t))$ is the advection velocity experienced by fish f at time t due to the presence of the other fish in the pair.

As a first approximation to the fluid flow generated by the other fish in the pair, we follow the work by Filella et al. (2018); Gazzola et al. (2016); Tchieu et al. (2012), wherein vorticity shedding due to swimming is neglected and the fluid is modeled as incompressible, inviscid, and irrotational. From an application perspective, the most restrictive of these assumptions are those of irrotationality and no vorticity shedding. Irrotationality will be violated in any sort of shear flow, including flow near solid boundaries, thereby challenging the use of the model to examine wall-hugging behaviors that could be associated with stress (Kalueff et al., 2013). Furthermore, neglecting vorticity shedding precludes the possibility to examine hydrodynamic interactions when the fish are in close proximity, interacting with each other's wake (Weihs, 1973).

Within potential flow theory, fish f in the pair is represented as a dipole in the far field, wherein the vortices comprising the dipole are positioned orthogonal to the swimming direction at a distance $r_{0,f}$. At position $\vec{r}_f(t)$, the other fish, denoted as $\check{f}$, creates an advective field given by (Filella et al., 2018)

$$\vec{U}_f(\vec{r}_f(t)) = \frac{r_{0,\check{f}}^2 v_{\check{f}}}{\rho(t)^2} \left[\sin \theta_{\check{f},f}(t) \hat{e}_{\check{f}}^\theta(t) + \cos \theta_{\check{f},f}(t) \hat{e}_{\check{f}}^\rho(t) \right], \quad (2)$$

where $\rho(t) = |\vec{r}_{\check{f}}(t) - \vec{r}_f(t)|$ is the distance between the two fish. In (2), the unit vectors $\hat{e}_{\check{f}}^\rho(t)$ and $\hat{e}_{\check{f}}^\theta(t)$ define a local polar coordinate system centered at fish $\check{f}$, in which $\hat{e}_{\check{f}}^\rho(t)$ points toward the other fish in the pair and forms a right-handed coordinate system with $\hat{e}_{\check{f}}^\theta(t)$ and $\hat{k} = \hat{i} \times \hat{j}$. The viewing angle $\theta_{\check{f},f}(t)$ is the angle between the swimming direction of fish j and $\hat{e}_{\check{f}}^\rho(t)$ (Fig. 1(a)). The distance $r_{0,f}$

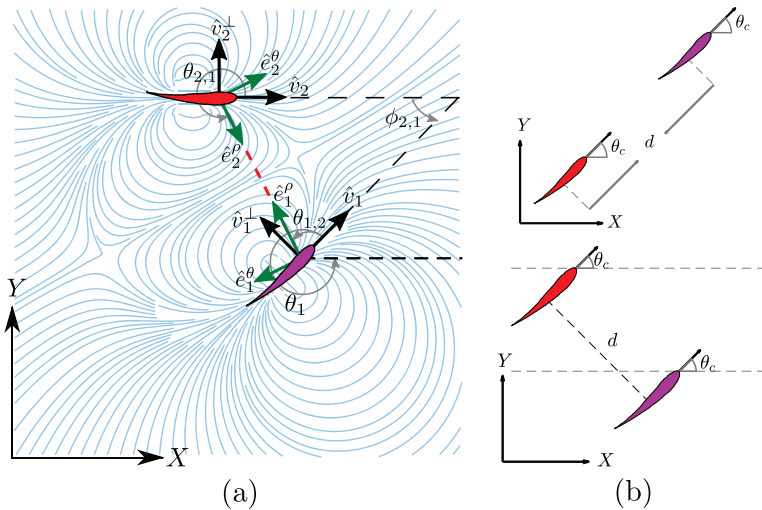


Figure 1. Schematic of an interacting zebrafish pair: (a) generic configuration with streamlines overlaid and key variables of fish 1 identified, and (b) in-line and side-by-side schooling patterns, where θ_c is the common orientation of the fish and d is their inter-individual distance.

should be on the order of the amplitude of the tail beating and its relationship with the speed of the animal provides the circulation of each vortex in the dipole, $\Gamma_f = 2\pi r_{0,f} v_{0,f}$.

For the analysis, it is more convenient to express the velocity field (2) in the global Cartesian coordinate system. From Fig. 1(a), we deduce

$$\hat{e}_2^\rho(t) = -\hat{e}_1^\rho(t) = \frac{\Delta x(t)\hat{i} + \Delta y(t)\hat{j}}{\rho(t)}, \quad (3a)$$

$$\hat{e}_2^\theta(t) = -\hat{e}_1^\theta(t) = \frac{-\Delta y(t)\hat{i} + \Delta x(t)\hat{j}}{\rho(t)}, \quad (3b)$$

where $\Delta x(t) = x_1(t) - x_2(t)$ and $\Delta y(t) = y_1(t) - y_2(t)$, such that $\rho(t) = \sqrt{\Delta x(t)^2 + \Delta y(t)^2}$. Similarly, we can express the sines and cosines of the viewing angle in (2) in terms of individual orientations as

$$\cos \theta_{\check{f},f}(t) = \hat{v}_{\check{f}}(t) \cdot \hat{e}_{\check{f}}^\rho(t) = \pm \frac{\Delta x(t) \cos \theta_{\check{f}}(t) + \Delta y(t) \sin \theta_{\check{f}}(t)}{\rho(t)}, \quad (4a)$$

$$\sin \theta_{\check{f},f}(t) = (\hat{v}_{\check{f}}(t) \times \hat{e}_{\check{f}}^\rho(t)) \cdot \hat{k} = \pm \frac{\Delta y(t) \cos \theta_{\check{f}}(t) - \Delta x(t) \sin \theta_{\check{f}}(t)}{\rho(t)}, \quad (4b)$$

where the plus (minus) sign is for f equal to 1 (2).

By replacing (3) and (4) into (2), we establish

$$\vec{U}_f(\vec{r}_f(t)) = \frac{r_{0,f}^2 v_{0,f}}{\rho(t)^4} \left[((\Delta x(t)^2 - \Delta y(t)^2) \cos \theta_{\check{f}}(t) + 2\Delta x(t)\Delta y(t) \sin \theta_{\check{f}}(t))\hat{i} + ((\Delta y(t)^2 - \Delta x(t)^2) \sin \theta_{\check{f}}(t) + 2\Delta x(t)\Delta y(t) \cos \theta_{\check{f}}(t))\hat{j} \right]. \quad (5)$$

To complete the model of the fish pair, we must prescribe governing equations for the orientation $\theta_f(t)$, $f \in \{1, 2\}$. Different from Filella et al. (2018), we employ a dynamic model that captures inherent delays in zebrafish to respond to visual or hydrodynamic input. Based on earlier models of social interaction, such as those by Calovi et al. (2014); Calovi et al. (2015), we propose

$$\dot{\theta}_f(t) = \omega_f(t), \quad (6a)$$

$$d\omega_f(t) = -\eta_f(\omega_f(t) - \omega_f^\star(t) - \Omega_f(t))dt + dN_f(t). \quad (6b)$$

Here, η_f is the mean reversion rate that defines the characteristic time-scale for fish to respond to external stimuli, thereby condensing neurological processes in decision-making and physical delays in reaction to external stimuli into a single damping coefficient per unit moment of inertia (in a kinematic model like in Filella et al. (2018), $\eta_f \rightarrow \infty$); $\omega_f^\star(t)$ is the imposed turn rate from social interactions; $\Omega_f(t)$ is the imposed turn rate due to hydrodynamic interactions; and $N_f(t)$ is added noise in the form of a zero-mean stationary stochastic process that captures non-deterministic forcing from the surrounding. All these contributions are linearly superimposed in the present model.

The response function $\omega_f^\star(t)$ accounts for social interactions, which are primarily due to visual cues and result in the tendency of the animals to align their bodies during swimming (schooling) and swim in the vicinity of others (shoaling), see, for example, Miller and Gerlai (2011). We write the response function for fish f as originally proposed by Calovi et al. (2014) and Calovi et al. (2015)

$$\omega_f^\star(t) = (1 + \cos \theta_{f,\check{f}}(t))(k_{v,f} v_f \sin \phi_{f,\check{f}}(t) + k_{p,f} \rho \sin \theta_{f,\check{f}}(t)), \quad (7)$$

where $k_{v,f}$ and $k_{p,f}$ are gain parameters for alignment and attraction, respectively (Fig. 1(a)), summarizing multisensory cues that underlie social response. The response function is such that the contribution to the imposed turn rate from the alignment rule is maximized when $\phi_{f,\check{f}} = \pm\pi/2$, which corresponds to fish swimming orthogonally. On the other hand, the contribution from the attraction rule is maximized when $\theta_{f,\check{f}} = \pm\pi/2$, which corresponds to the animals swimming in parallel. In this

model, alignment is taken to be independent of the relative distance between the animals, while attraction increases the further the animals are from one another. These approximations are expected to be valid when fish swim within about ten body lengths; otherwise, it may be appropriate to include a decay function that reduces the strength of the interaction for large distances (Zienkiewicz et al., 2015b). The prefactor $(1 + \cos \theta_{f,\check{f}}(t))$ accounts for the cone of vision of zebrafish (Pita et al., 2015), which preferentially biases interactions toward animals that are swimming in front rather than those in the back. As discussed by Calovi et al. (2015), this prefactor “introduces a strong asymmetry between the force exerted by f on $\check{f}$ and the one exerted by $\check{f}$ on f , and hence breaks the (Newtonian) action–reaction principle which is most familiar in the context of purely physical force, such as gravitation.”

Within the potential flow framework, the turn rate imposed by hydrodynamic interactions $\Omega_f(t)$ arises from differences in the velocity experienced by each vortex comprising the dipole of fish f . Thus, the gradient of streamwise velocity in the direction orthogonal to $\hat{v}_f$ induces the hydrodynamic turning

$$\Omega_f(t) = \hat{v}_f(t) \cdot \nabla \vec{U}_f(\vec{r}_f(t)) \hat{v}_f^\perp(t), \quad (8)$$

where $\hat{v}_f^\perp(t) = \hat{k} \times \hat{v}_f(t)$ is the unit vector orthogonal to $\hat{v}_f(t)$ forming a local right-handed coordinate system and ∇ is the gradient operator in the global coordinate system. By utilizing the expression of the velocity $\vec{U}_f(\vec{r}_f(t))$ in (2), along with the Cartesian representation of $\hat{v}_f^\perp(t)$ and $\hat{v}_f$, we establish

$$\Omega_f(t) = \pm \frac{2r_{0,\check{f}}^2 v_{0,\check{f}}}{\rho(t)^6} ((-3\Delta x(t)^2 \Delta y(t) + \Delta y(t)^3) \cos(2\theta_f(t) + \theta_{\check{f}}(t)) + (\Delta x(t)^3 - 3\Delta x(t) \Delta y(t)^2) \sin(2\theta_f(t) + \theta_{\check{f}}(t))) \quad (9)$$

where the plus (minus) sign is for f equal to 1 (2).

Added noise in (6b) captures the uncertainty in the behavior of the animal, which may not necessarily balance between social and hydrodynamic interactions. Zebrafish locomotion is characterized by a unique burst-and-coast swimming style in which “fish move forward (burst) in a single motion and glide (coast) to a slow speed, or stop from which they burst forward again” (Kalueff et al., 2013). This swimming style contributes to a rich repertoire of locomotory patterns, including sudden U- and C-turns that alternate with instances of straight swimming (Kalueff et al., 2013). As a result, uncertainty to turn rate dynamics cannot be captured by Gaussian white noise (that is, increments of a Wiener process), as traditionally advocated for other swimming animals (Gautrais et al., 2009).

This issue has been addressed by Mwaffo et al. (2015) by modeling added noise as the sum of two terms: i) a Wiener process $\sigma_f W(t)$, where $W(t)$ is a standard Wiener process and σ_f is a positive constant¹, and ii) a compound Poisson process $J_f(t) = \sum_{k=1}^{P_f(t)} Z_{f,k}$, where $P_f(t) - P_f(s)$ (for $s \leq t$) is a Poisson random variable with parameter $\lambda_f(t - s)$, and $Z_{f,k} \sim \mathcal{N}(0, \gamma_f^2)$ ². The Wiener process represents the baseline uncertainty with animal locomotion, associated with free will in the decision-making process. The compound Poisson process allows to capture jumps in the turn rate dynamics, associated with C- and U-turns. Overall, added noise is modelled through three independent parameters: σ_f , setting a baseline activity for the animal; λ_f , quantifying the frequency of the jumps; and γ_f , measuring the strength of the jumps.

In light of the choice of the form of added noise, (6b) can be viewed as a jump reverting mean diffusion process (Tankov, 2003) in which the turn rate relaxes toward a time-varying mean, given by the sum of the turn rates imposed by social and hydrodynamic interactions in (7) and (8), respectively. The original model by Filella et al. (2018) has neither the feature of reversion nor the presence of jumps. Hence, at every time, the turn rate of each fish matches the sum of the turn rates imposed by social and hydrodynamic interactions, subject to added Gaussian white noise³.

¹For a standard Wiener process, $dW(t) \sim \mathcal{N}(0, dt)$, where $\mathcal{N}(\mu, \sigma^2)$ is a normal distribution with mean μ and variance σ^2 , so that multiplying by σ_f scales the variance of the normal distribution by σ_f^2 .

²The increment in the compound Poisson process can be written as $dJ_f(t) = Z dP_f(t)$, where $Z \sim \mathcal{N}(0, \gamma_f^2)$ and $dP_f(t) \sim \mathcal{B}(\lambda_f dt)$ ($\mathcal{B}(q)$ being a Bernoulli distribution with parameter q).

³In the model by Filella et al. (2018), $\omega_f(t) = \omega_f^*(t) + \Omega_f(t) + \sigma_f \frac{dW(t)}{dt}$, where σ_f has units rad s^{-1} rather than $\text{rad s}^{-3/2}$ as in our study.

3. Analytical insight into the model

Predicting the behavior of a zebrafish pair requires the time integration of (1) and (6). While the general mathematical treatment of this system may not be feasible, important insight can be gathered by examining its linearized dynamics about salient configurations that have previously been documented (De Bie et al., 2020; Kato et al., 2004; Laan et al., 2017; Perna et al., 2014). To further simplify the problem, we assume the two fish to be identical, such that $r_{0,1} = r_{0,2} = r_0$, $v_{0,1} = v_{0,2} = v_0$, $k_{p,1} = k_{p,2} = k_p$, $k_{v,1} = k_{v,2} = k_v$, $\eta_1 = \eta_2 = \eta$, $\sigma_1 = \sigma_2 = \sigma$, $\lambda_1 = \lambda_2 = \lambda$, and $\gamma_1 = \gamma_2 = \gamma$.

Following standard practice in the literature on synchronization of coupled dynamical systems (Pikovsky et al., 2003), we decompose the dynamics of the fish pair into an average and an error dynamics. The former captures the evolution of the center of mass of the pair and of the average orientation, whereas the latter describes changes in relative distance and orientation. Hence, we write

$$x_f(t) = \bar{x}(t) \pm \frac{1}{2}\Delta x(t), \quad y_f(t) = \bar{y}(t) \pm \frac{1}{2}\Delta y(t), \quad \theta_f(t) = \bar{\theta}(t) \pm \frac{1}{2}\Delta\theta(t), \quad \omega_f(t) = \bar{\omega}(t) \pm \frac{1}{2}\Delta\omega(t), \quad (10)$$

where the plus (minus) sign is for f equal to 1 (2), a superimposed bar indicates the average dynamics, and the symbol Δ is consistently used to label relative dynamics.

By using this coordinate transformation, we consider the eight-dimensional state variable $z(t) = [\bar{x}(t), \bar{y}(t), \bar{\theta}(t), \bar{\omega}(t), \Delta x(t), \Delta y(t), \Delta\theta(t), \Delta\omega(t)]^T$, where T indicates matrix transposition. The time-evolution of the state is given by

$$\dot{z}(t) = F(z(t), u(t)) \quad (11)$$

where $u(t) \in \mathbb{R}^2$ encapsulates added noise to the turn rate dynamics of the animals and the nonlinear function $F : \mathbb{R}^8 \times \mathbb{R}^2 \rightarrow \mathbb{R}^8$ captures the first order dynamics of the turn rate in (6b) along with the coupling between the state variables of the two fish through the advection velocity (2), social interactions (7), and hydrodynamic interactions (8).

The class of nominal solutions about which we perform a linearization is part of the schooling patterns identified by Filella et al. (2018), wherein fish swim along the same, constant direction, that is, $\Delta\theta = 0$ and $\bar{\theta}(t) = \theta_c$ for some θ_c . Fish will follow a straight line ($\Delta\omega = 0$ and $\bar{\omega}(t) = 0$) and rigidly translate one with respect to each other in two possible patterns, as shown in Fig. 1(b). In the first case of in-line (sometimes called “tandem”) swimming, one fish follows the other, such that $\Delta x(t) = d \cos \theta_c$ and $\Delta y(t) = d \sin \theta_c$ where d is the constant distance between the fish. In the second case of side-by-side swimming, the fish swim parallel to each other, such that $\Delta x(t) = -d \sin \theta_c$ and $\Delta y(t) = d \cos \theta_c$. Without loss of generality, we set $\theta_c = 0$, such that the X-axis of the Cartesian coordinate system is aligned with the schooling direction.

3.1. In-line swimming

It is easy to verify that for any choice of d , in-line swimming constitutes a nominal solution of (11) in the absence of added noise. Specifically, both the imposed turn rates due to social and hydrodynamic interactions in (7) and (8) are identically zero, such that the orientation of each fish remains constant ($\hat{v}_f(t) = \hat{i}$). With respect to the positional coordinates, we observe that $\vec{U}_f(\vec{r}_f(t)) = \frac{r_0^2 v_0}{d^2} \hat{i}$ for $f \in \{1, 2\}$, such that each fish will be subject to the same advective velocity, thereby preserving the relative configuration. The center of mass of the pair will move at a constant speed given by $v_0 \left(1 + \frac{r_0^2}{d^2}\right)$, which yields the increase in speed due to hydrodynamic interaction observed through simulations on large groups in Filella et al. (2018).

The homogeneous, linearized dynamics of the system is informative of the local stability of the in-line configuration. By taking the gradient of the nonlinear vector field in (11) and evaluating it at the nominal solution, we obtain the following state matrix describing the evolution of perturbations with

respect to the nominal, in-line solution:

$$\dot{\tilde{x}}^\delta(t) = -\frac{2r_0^2 v_0}{d^3} \Delta x^\delta(t), \quad (12a)$$

$$\dot{\tilde{y}}^\delta(t) = v_0 \left(1 - \frac{r_0^2}{d^2} \right) \bar{\theta}^\delta(t) + \frac{2r_0^2 v_0}{d^3} \Delta y^\delta(t), \quad (12b)$$

$$\dot{\bar{\theta}}^\delta(t) = \bar{\omega}^\delta(t), \quad (12c)$$

$$\dot{\bar{\omega}}^\delta(t) = -k_p \eta d \bar{\theta}^\delta(t) - \eta \bar{\omega}^\delta(t) + k_p \eta \Delta y^\delta(t) + \left(\frac{\eta r_0^2 v_0}{d^3} + \frac{k_p \eta d}{2} + k_v \eta v_0 \right) \Delta \theta^\delta(t), \quad (12d)$$

$$\dot{\Delta x}^\delta(t) = 0, \quad (12e)$$

$$\dot{\Delta y}^\delta(t) = v_0 \left(1 - \frac{r_0^2}{d^2} \right) \Delta \theta^\delta(t), \quad (12f)$$

$$\dot{\Delta \theta}^\delta(t) = \Delta \omega^\delta(t), \quad (12g)$$

$$\dot{\Delta \omega}^\delta(t) = \left(\frac{12\eta r_0^2 v_0}{d^3} + 2k_p \eta d \right) \bar{\theta}^\delta(t) - \left(2k_p \eta + \frac{12\eta r_0^2 v_0}{d^4} \right) \Delta y^\delta(t) - (k_p \eta d + 2k_v \eta v_0) \Delta \theta^\delta(t) - \eta \Delta \omega^\delta(t). \quad (12h)$$

Here and henceforth, a superscript δ identifies perturbations with respect to the nominal solution.

Equation (12) demonstrates several interesting features of the model. Predictably, perturbations on the position of the center of mass do not enter the dynamics of any of the other state variables. Second, the relative distance along the schooling direction remains constant, independent of any other state variable. Third, the perturbation on the relative distance $\Delta y^\delta(t)$ depends only on the relative rotation $\Delta \theta^\delta(t)$. Fourth, the response of the fish pair is asymmetric, due to the presence of social interactions which are sensitive to the relative orientation of the animals (leading versus trailing fish)⁴.

To further clarify the structure of the system, we perform a coordinate transformation that accounts for the specific nature of in-line swimming. Specifically, we define $\bar{\theta}_\star^\delta(t) = \bar{\theta}^\delta(t) - \frac{\Delta y^\delta(t)}{d}$ to filter out rigid-body motion in the definition of the average angle; in fact, in-line swimming would not be affected if both the fish moved along the Y -axis while rotating their body to preserve straight swimming. Consistently, we define $\bar{\omega}_\star^\delta(t) = \dot{\bar{\theta}}_\star^\delta(t)$. By using these variables in (12) and ordering the components of the state vector as $z^\delta(t) = [\bar{x}^\delta(t), \bar{y}^\delta(t), \Delta x^\delta(t), \Delta y^\delta(t), \bar{\theta}_\star^\delta(t), \bar{\omega}_\star^\delta(t), \Delta \theta^\delta(t), \Delta \omega^\delta(t)]^T$, we can write the state matrix of the system in the following block-triangular form:

$$A_{\text{in-line}} = \left[\begin{array}{cccc|cccc} 0 & 0 & -\frac{2r_0^2 v_0}{d^3} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{v_0}{d} \left(1 + \frac{r_0^2}{d^2} \right) & v_0 \left(1 - \frac{r_0^2}{d^2} \right) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & v_0 \left(1 + \frac{r_0^2}{d^2} \right) & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -k_p \eta d & -\eta \left(\frac{k_p \eta d}{2} + k_v \eta v_0 - \frac{\eta v_0}{d} \right) - \frac{v_0}{d} \left(1 + \frac{r_0^2}{d^2} \right) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & \frac{12\eta v_0 r_0^2}{d^3} + 2k_p \eta d & 0 & -k_p \eta d - 2k_v \eta v_0 & -\eta \end{array} \right] \quad (13)$$

The stability of in-line swimming to external perturbations is ascertained by examining the spectrum of the four-by-four block on the bottom diagonal, $A_{\text{in-line},22}$. In-line swimming is asymptotically stable if and only this sub-matrix is Hurwitz. If $A_{\text{in-line},22}$ is Hurwitz, the orientations converge to a common

⁴Asymmetry is evident in the turn rate dynamics by combining (12d) and (12h) to obtain governing equations for $\omega_1^\delta(t)$ and $\omega_2^\delta(t)$.

value ($\Delta\theta^\delta(t), \Delta\omega^\delta(t) \rightarrow 0$), such that the two fish swim along a straight line ($\bar{\theta}_\star^\delta(t), \bar{\omega}_\star^\delta(t) \rightarrow 0$). The other blocks have no role on stability, whereby the location of the center of mass and the relative distance $\Delta x^\delta(t)$ are inconsequential with respect to in-line swimming, and $\Delta y^\delta(t)$ converges to a constant value that depends on the common orientation attained by the fish pair.

While a precise evaluation of the stability of in-line swimming requires the calculation of the spectrum of $A_{\text{in-line},22}$, important insight can be gathered from the application of the Routh-Hurwitz criterion. Briefly, application of the criterion requires the construction of an ancillary table of coefficients, computed from the coefficients of the characteristic polynomial (Meinsma, 1995). By counting the number of sign changes in the first columns of this table, it is possible to exactly count the number of eigenvalues with positive real part. In general, it is difficult to tell the sign of these coefficients, but several interesting limit cases can be explored. For example, for d sufficiently large, we discover that there are no sign changes, demonstrating that in-line swimming is asymptotically stable for sufficiently large separation distances. On the other hand, for small values of d , there are two sign changes, indicating the presence of two eigenvalues with positive real part.

Another interesting case is the limit of large values of η , which is reminiscent of the setup by Filella et al. (2018); in this case, $(d^4 - 3d^2r_0^2)k_p - 6dr_0^2v_0k_v + 6r_0^2v_0 > 0$ is the necessary and sufficient condition for the matrix to be Hurwitz. For any choice of d , this condition suggests that large values of k_v will hinder stability of the system; on the other hand, the effect of k_p is modulated by d , whereby for $d < \sqrt{3}r_0$, large gains are again detrimental to stability, while for $d > \sqrt{3}r_0$, they could guarantee stability. Should hydrodynamic interactions be dismissed by setting $r_0 = 0$, we would always obtain asymptotic stability.

3.2. Side-by-side swimming

The study of side-by-side swimming follows closely the previous analysis. However, there are some key differences between these schooling patterns. Unlike in-line swimming, side-by-side swimming requires a specific distance between the animals, which is obtained by balancing the repulsion from hydrodynamics with attraction due to social interactions, leading to $d = \sqrt[4]{\frac{2r_0^2v_0}{k_p}}$. While maintaining this distance, the advective velocity is $\vec{U}_f(\vec{r}_f(t)) = -\frac{r_0^2v_0}{d^2}\hat{i}$ for $f \in \{1, 2\}$, thereby slowing down the motion of the fish that will swim at a reduced speed of $v_0 \left(1 - \frac{r_0^2}{d^2}\right)$.

Similar to in-line swimming, we can derive the equations of motion for a perturbation about this nominal solution, that is,

$$A_{\text{side-by-side}} = \left[\begin{array}{cccc|cccc} 0 & 0 & 0 & 0 & 0 & \frac{2r_0^2v_0}{d^3} & 0 & 0 \\ 0 & 0 & v_0 \left(1 + \frac{r_0^2}{d^2}\right) & 0 & \frac{2r_0^2v_0}{d^3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & k_p\eta d & -\eta & k_p\eta & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & v_0 \left(1 - \frac{r_0^2}{d^2}\right) & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -\left(2k_p\eta + \frac{12\eta r_0^2v_0}{d^4}\right) & k_p\eta d - 2k_v\eta v_0 & -\eta \end{array} \right], \quad (14)$$

with state variables ordered as $z^\delta(t) = [\bar{x}^\delta(t), \bar{y}^\delta(t), \bar{\theta}^\delta(t), \bar{\omega}^\delta(t), \Delta x^\delta(t), \Delta y^\delta(t), \Delta\theta^\delta(t), \Delta\omega^\delta(t)]^T$.

This system is already in a block-triangular form, in which the block on the top of the diagonal, $A_{\text{side-by-side},11}$, has a block-triangular structure, with two zero eigenvalues corresponding to the motion of the center of mass. The average rotation of the pair has a nontrivial dynamics, which is characterized

by two real eigenvalues $\frac{\eta}{2} \left(-1 \pm \sqrt{1 + \frac{4k_p d}{\eta}} \right)$. One of these eigenvalues is always positive, indicating that perturbations will trigger the exponential growth of the average orientation away from zero. The growth of the average orientation is not accompanied by any change in the relative distance along the X -axis (fifth row of the matrix) that could maintain the side-by-side pattern. As a result, side-by-side swimming is always unstable. The remaining three non-zero eigenvalues of the block on the bottom of the diagonal, $A_{\text{side-by-side}, 22}$, shape the dynamics of the relative orientation and relative distance along the Y -direction, which, however, bear no effect on the stability of the pattern.

4. Experimental data and model calibration

We calibrated the model using experimental data from a previous study (Zienkiewicz et al., 2015a); the same dataset was also used in Butail et al. (2016) and is included as Supplementary Material. The dataset contains 17 different pairs of adult zebrafish of average body length (BL) of 30 mm swimming in a circular arena of 0.9 m diameter and water depth of 0.1 m. The shallow depth encouraged the fish to swim in a quasi-two dimensional plane. Fish motion was filmed at 30 frames s^{-1} for 20 min, and a multi-target tracking system (Ladu et al., 2014) was used to track the centroids of each individual. The first 10 min were treated as a habituation phase, in which the animals explored the novel tank and acclimatize to the environment.

The last 10 min of videos were used for the analysis. Each pair of experimental trajectory data consists of positions $\vec{r}_f(t)$ and orientations $\theta_f(t)$ in the instantaneous direction of motion of the two animals sampled at a time-step $\Delta t = 1/30$ s. We filtered the data in two steps. Time windows where the two fish were closer than half a BL from each other were removed to avoid diving instances, which would violate the assumption of planar swimming. To mitigate the effect on un-modelled hydrodynamic interactions with the walls, we also removed instances when either fish swam closer than two BLs to the wall. Overall, these pre-processing steps curtailed approximately half of the dataset.

This same dataset was also used for validating the predictions of the model, by evaluating salient metrics of collective behavior that emerge within the pair. Specifically, we considered the average polarization, which measures the relative orientation of the pair: $\text{Pol} = \frac{1}{T} \int_0^T \frac{\|\hat{v}_1(t) + \hat{v}_2(t)\|}{2} dt$, where T is the duration of the observation (Calovi et al., 2014). A polarization value of one indicates that the swimming directions of both fish are always aligned during the observation window, whereas a value of zero identifies that the animals are swimming in opposite directions. While polarization helps quantify schooling, it does not distinguish between in-line and side-by-side swimming, which we capture through an alignment index defined as $\text{AI} = \frac{1}{T} \int_0^T \frac{(\hat{v}_1(t) + \hat{v}_2(t)) \cdot (\vec{r}_1(t) - \vec{r}_2(t))}{2\rho(t)} dt$. A value of 1 (−1) indicates in-line swimming, with the first (second) fish is swimming in front of the second (first) fish. A zero value corresponds to the pair swimming side-by-side. The average inter-individual distance between the pair of fish, $\text{Dist} = \frac{1}{T} \int_0^T |\rho(t)| dt$, is used to quantify their shoaling tendency, that is, their preference to stay close to one another. Finally, we measured the average speed of the center of mass $\text{Speed} = \frac{1}{T} \int_0^T \frac{\|\vec{r}_1(t) + \vec{r}_2(t)\|}{2} dt$ to further detail the effect of interactions on locomotion.

For each fish, we identified eight model parameters (characteristic length, r_0 ; speed, v_0 ; baseline activity, σ_f ; reversion rate, η_f ; jump frequency, λ_f ; jump intensity, γ_f ; and attraction and alignment gains, $k_{p,f}$ and $k_{v,f}$ respectively) from the equations of motion (1) and (6). We performed the identification in two sequential steps. First, we calibrated characteristic length, r_0 , and speed, v_0 , by minimizing the sum-square-error cost function (SSECF) calculated from the experimentally measured speed and the speed in (1), defined as $\text{SSECF} = \sum_{k \in \text{Data}_f} \left(\|\vec{r}_f(t_k)\| - \|v_f \hat{v}_f(t_k) + \vec{U}_f(\vec{r}_f(t_k))\| \right)^2$, where “Data_{*f*}” identifies the subset of indices that remain after the removal presented above for fish *f*. The optimization was solved by constraining each parameter within a physically plausible range. Specifically, based on prior data on fish swimming in isolation and the typical body length of a zebrafish

(Ladu et al., 2014), we selected $v_f \in [0, 0.30] \text{ m s}^{-1}$, and $r_f \in [0, 30] \text{ mm}$. Calibrated parameters are presented in Fig. 2(a,b).

Second, we identified the remaining six parameters $\Theta_f = [\sigma_f, \eta_f, \lambda_f, \gamma_f, k_{p,f}, k_{v,f}]^T$ using a maximum likelihood approach on the turn rate dynamics in (6b). The maximum likelihood estimation was subject to further data pre-processing, whereby we removed instances when fish were swimming approximately in-line ($\text{AI} > 0.75$); in these instances the prefactor in the social interaction (7) would cause excessively small values of the response function, which would challenge robust identification.

Hence, we estimated model parameters by minimizing the sum of negative logarithms of the likelihood function $\mathcal{L}(\omega_f(t)|\omega_f(t - \Delta t), \Theta_f)$ over experimental time-series using FMINCON function in MATLAB (Butail et al., 2016). The likelihood function was constructed as the sum of two normal probability distributions associated with the two forms of added noise

$$\mathcal{L}(\omega_f(t)|\omega_f(t - \Delta t), \Theta_f) = (1 - \lambda_f \Delta t)H(\omega_f(t), \mu_f(\omega_f(t - \Delta t), \Delta t), \text{var}_f(\Delta t)) + \lambda_f \Delta t H(\omega_f(t), \mu_f(\omega_f(t - \Delta t), \Delta t), \text{var}_f(\Delta t) + \gamma_f), \quad (15a)$$

where

$$\mu_f(\omega_f(t - \Delta t), \Delta t) = \omega_f(t - \Delta t)e^{-\eta_f \Delta t} + (\omega_f^*(t - \Delta t) + \Omega_f(t - \Delta t))(1 - e^{-\eta_f \Delta t}), \quad (15b)$$

$$\text{var}_f(\Delta t) = \frac{\sigma_f^2}{2\eta_f}(1 - e^{-2\eta_f \Delta t}), \quad (15c)$$

and $H(\bullet, \mu, \sigma^2)$ is the probability mass function of $\mathcal{N}(\mu, \sigma^2)$.

Based on the social model of Butail et al. (2016), we constrained the parameter ranges for the search as follows. First, we set $\eta_f \in [1, 2] \text{ s}^{-1}$, $\sigma_f \in [1, 5] \text{ rad s}^{-3/2}$, and $\gamma_f \in [1, 5] \text{ rad s}^{-1}$. Then, we explored the range in jump frequency of $\lambda_f \in [0.9, 1.1]\lambda_{0,f}$, where $\lambda_{0,f}$ was an estimate of the frequency of extreme events in the time-series, defined as the ratio between the number of time-steps above three standard deviations from the mean of the turn rate and the total number of time-steps. With respect to interactions parameters, we explored the broad range $k_{p,f} \in [0, 100] \text{ rad m}^{-1} \text{ s}^{-1}$ and $k_{v,f} \in [0, 100] \text{ rad m}^{-1}$ to capture inter-individual variability in social behavior.

5. Results and discussion

By applying the calibration procedure explained above, we obtain the complete set of calibrated parameters in Fig. 2. The values of r_0 is one order of magnitude less than the animal body length ($r_0 = 3.1 \pm 3.2 \text{ mm}$), which corresponds to the typical tail beat amplitude of a cruising zebrafish (Gazzola et al., 2014). The self-propulsion speed ranges from 2BL to 6BL per second ($v_0 = 0.094 \pm 0.027 \text{ m s}^{-1}$), suggesting variability among animals with respect to their locomotion. Interestingly, these two variables are highly correlated⁵, such that larger distances between vortices comprising the dipole will be accompanied by larger speeds, thus implying larger values of the vortices' circulation, see table 1. This trend is analogous to the relationship between tail beat amplitude and the circulation of vortices shed during tail beating in real experiments by Mwaffo et al. (2017).

The baseline activity ($\sigma = 1.39 \pm 0.36 \text{ rad s}^{-3/2}$), mean reversion rate ($\eta = 1.73 \pm 0.33 \text{ s}^{-1}$), jump frequency ($\gamma = 2.52 \pm 0.73 \text{ s}^{-1}$), jump intensity ($\lambda = 0.64 \pm 0.12 \text{ s}^{-1}$), and gain parameters for attraction ($k_p = 3.92 \pm 4.18 \text{ rad m}^{-1} \text{ s}^{-1}$) and alignment ($k_v = 17.3 \pm 8.00 \text{ rad m}^{-1}$) are comparable with Butail et al. (2016), which does not include a model of hydrodynamic coupling. These parameter values are not all independent of each other (table 1). We determine a positive relationship between baseline activity and jump intensity, which suggests that animals exhibiting higher activity levels also display stronger bursts. We also record a dependency between the social interaction gains, which supports that social behavior equivalently reverberates through alignment and attraction tendencies. Finally, we discover

⁵Should one correct for multiple comparisons using a conservative Bonferroni correction, some significant correlations would become trends.

that self-propulsion speed is negatively correlated with the mean reversion rate, indicating that it is more difficult for faster animals to adjust their orientation in response to social and hydrodynamic cues.

Overall, the extent to which hydrodynamic interactions improve the quality of the fit can be ascertained using the Akaike information criterion (Akaike, 1998). Specifically, for the present model in (6b), this quantity is equal to 46.76 ± 32.45 . Should we compute the same for the model in Butail et al. (2016), which excluded hydrodynamic interactions, we would obtain a strikingly similar result of 46.74 ± 32.54 . Hence, accounting for hydrodynamic interactions does not lead to a tangible improvement on the fit of the turn rate, although it allows to capture speed-based interactions that would be otherwise neglected. Evidence in favor of the existence of speed-based interactions has been documented in a number of experimental studies on freshwater fish (see, for example, Herbert-Read et al. (2011); Katz et al. (2011)), but their connection to hydrodynamics has remained elusive.

To demonstrate the model predictive power, we perform a series of *in-silico* experiments, employing calibrated parameters from Fig. 2; a sample video is in the Supplementary Material. This analysis aims to ascertain whether the model predicts equivalent statistical outcomes to real experiments, thereby allowing for partial replacement of animal trials. We run the model 17 times to mirror experimental observations, using identical parameters for each fish in the pair (drawn from normal distributions approximating those in Fig. 2) and random initial conditions (orientations drawn uniformly in $[0, 2\pi]$ and distance within $4r_0$, similar to Tchieu et al. (2012)). For each metric of collective behavior, we perform *t*-tests comparisons between chance and either real or *in-silico* experiments. Chance is computed by randomly shuffling fish across the 17 experimental observations 10,000 times and taking the mean.

As shown in Fig. 3, the model is successful in anticipating a strong schooling tendency of the pair (experiments: $t_{(1,16)} = 20.67$, $p < 0.001$; *in-silico*: $t_{(1,16)} = 8.34$, $p < 0.001$). During schooling,

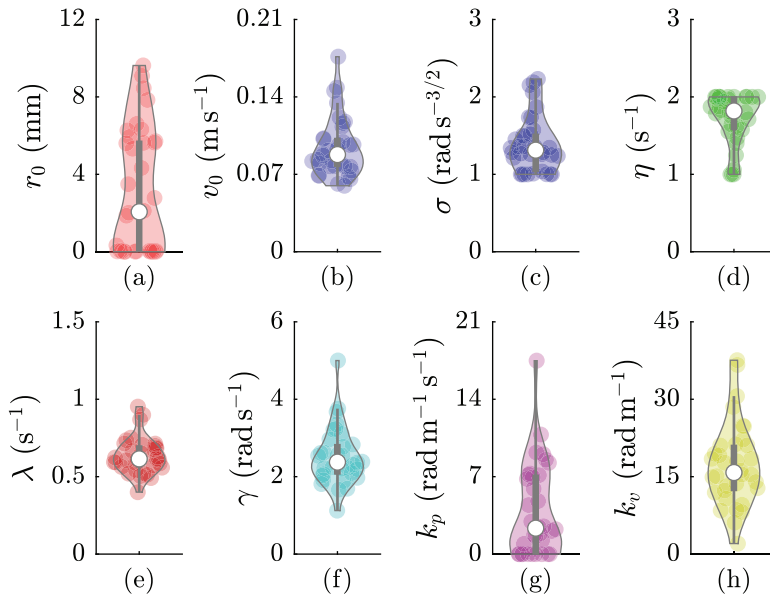


Figure 2. Calibrated model parameters: (a) characteristic length, (b) speed, (c) baseline activity, (d) mean reversion rate, (e) jump frequency, (f) jump intensity, (g) gain parameter for attraction, and (h) gain parameter for alignment. The colored area in each violin plot is the probability density and coloured circles are individual calibrations. Thick gray bars indicate first and third quartiles; thin gray bars identify minimum and maximum values; and white circles are the median. Calibrated parameters for each fish are reported in the Supplementary Material.

animals favor in-line swimming, which is accurately predicted by the model (experiments: $t_{(1,16)} = 14.74$, $p < 0.001$; *in-silico*: $t_{(1,16)} = 10.59$, $p < 0.001$). Although to a lesser quantitative extent, the model is also able to anticipate shoaling tendency of the pair (experiments: $t_{(1,16)} = -31.20$, $p < 0.001$; *in-silico*: $t_{(1,16)} = -2.65$, $p = 0.018$) and the increase in speed due to animal interactions (experiments: $t_{(1,16)} = 3.31$, $p = 0.004$; *in-silico*: $t_{(1,16)} = 3.05$, $p = 0.008$). Experimental subjects tend to swim within 2 to 3 BLs, which correspond to approximately one third of theoretical predictions, thereby leading to an underestimation of short-range interactions. The source of the discrepancy lies with the absence of social interactions in the speed dynamics, which could, however, be included in the model by following Zienkiewicz et al. (2015b).

Overall, Fig. 3 points to a strong tendency of live fish to school in an in-line pattern, which is accurately captured through *in-silico* experiments. Further insight can be garnered through the analysis of the eigenvalues of $A_{\text{in-line},22}$ in (13), which was derived from linearization of the proposed model. Specifically, we use calibrated values of r_0 , v_0 , and η from Fig. 2 to analyze the stability of in-line

	r_0	v_0	σ	η	λ	γ	k_p	k_v
r_0	—	×	×	×	×	×	×	×
v_0	0.416 (0.015)	—	×	×	×	×	×	×
σ	0.158 (0.373)	0.120 (0.500)	—	×	×	×	×	×
η	-0.174 (0.326)	-0.430 (0.011)	0.141 (0.427)	—	×	×	×	×
λ	0.092 (0.604)	-0.075 (0.672)	-0.001 (0.994)	-0.076 (0.668)	—	×	×	×
γ	0.328 (0.059)	0.095 (0.593)	0.797 (0.000)	0.010 (0.955)	0.223 (0.204)	—	×	×
k_p	0.179 (0.312)	-0.062 (0.728)	0.184 (0.298)	-0.014 (0.939)	-0.181 (0.305)	0.059 (0.742)	—	×
k_v	0.006 (0.974)	-0.248 (0.158)	0.026 (0.883)	0.090 (0.613)	0.014 (0.937)	-0.097 (0.586)	0.374 (0.030)	—

Table 1. Coefficient of determination (R) and associated p-value in parentheses between model variables. Significant correlations at a level of 0.05 are indicated in bold.

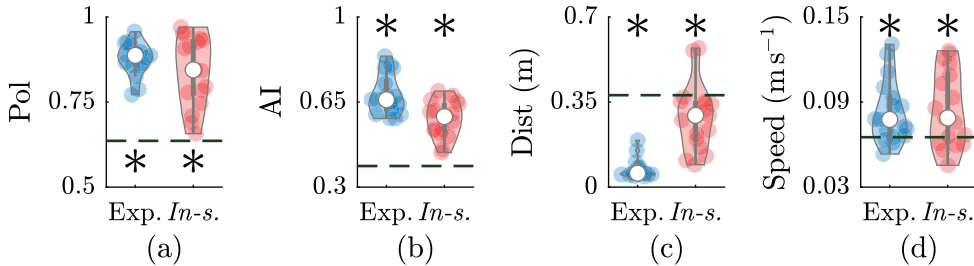


Figure 3. Predictive power of the proposed model for four metrics of collective behaviour: (a) polarization, (b) alignment index, (c) average inter-individual distance, and (d) average speed of the center of mass. Circles represent individual trials for real (Exp.) or *in-silico* (In-s.) experiments. Thick gray bars indicate first and third quartiles; thin gray bars identify minimum and maximum values; and white circles are the median. Dashed lines show chance levels; a star identifies a significant difference from chance in *t*-test comparison at a significance level of 0.05.

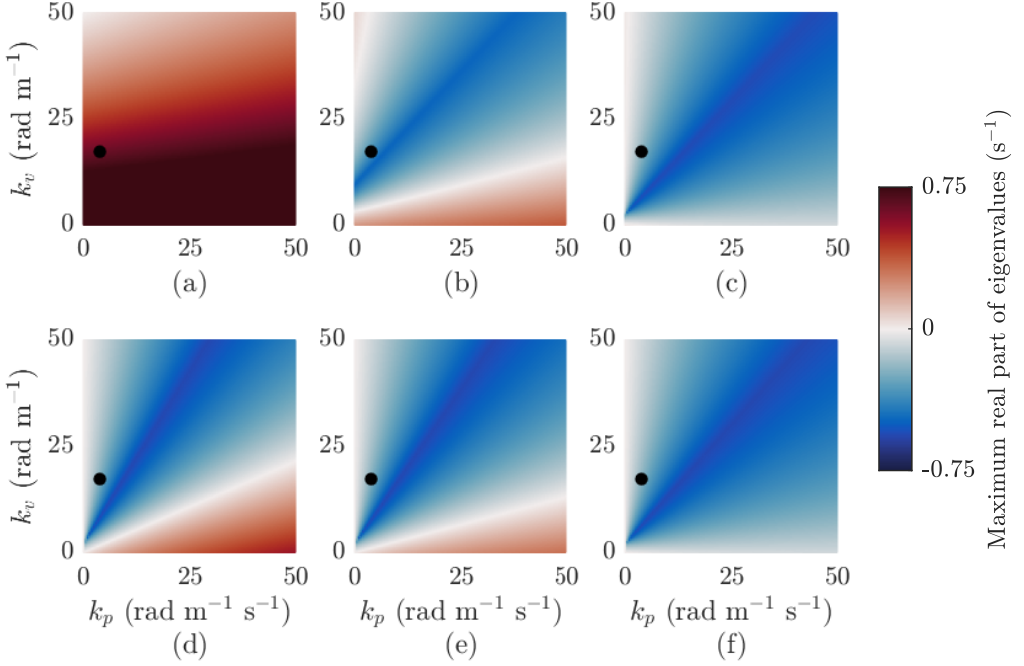


Figure 4. Stability analysis of in-line swimming as a function of social and hydrodynamic interactions, in the form of heat maps of the maximum real part of the eigenvalues of $A_{\text{in-line},22}$ in (13) for: (a,b,c) $r_0 = 3.1$ mm and (d,e,f) $r_0 = 0$ mm (no hydrodynamic interactions). For each of the two scenarios, we consider three values of inter-individual distances d : (a,d) 0.5 BL; (b,e) 1 BL; and (c,f) 2 BL. Other simulation parameters are mean values from Fig. 2: $\eta = 1.68 \text{ s}^{-1}$, and $v_0 = 0.094 \text{ m s}^{-1}$. White regions identify stability boundaries, and black points are mean values from Fig. 2.

swimming with respect to k_p , k_v , and d in Fig. 4. Numerical computation of the eigenvalues confirms the theoretical proposition from the Routh-Hurwitz criterion that small values of d hinder stability. For sufficiently large values of d , there is a wide range of combinations of the gains that exhibit stability. Using calibrated values of k_p and k_v , we always attain stability for any choice of d above approximately 2/3 BL, in agreement with observations on inter-individual distance in Fig. 3. Notably, for a given value of d , increasing either gain may hamper stability, in contrast to the kinematic model of Filella et al. (2018) that would suggest a beneficial role of large k_p for $d > \sqrt{3}r_0$. Sample videos of the effect of perturbations on in-line and side-by-side swimming are shown in the Supplementary Material.

For completeness, in Fig. 4, we also present results obtained by setting $r_0 = 0$, that is, dismissing hydrodynamic interactions without modifying any other model parameter. Hydrodynamic interaction effectively leads to a repulsion zone between the animals that would otherwise not be present. Repulsion has been widely documented in the literature on mathematical models of animal groups, but generally attributed to a social response (Vicsek and Zafeiris, 2012). For a given value of d , we also register a wider range of gain combinations that guarantee stability.

6. Conclusions

Here, we have presented and analyzed a mathematical model of collective behavior of zebrafish pairs. The model combines the jump persistent turning walker of Mwaffo et al. (2015) with the interaction

model by Filella et al. (2018) toward a realistic description of zebrafish locomotion and of social and hydrodynamic interactions. The model consists of a system of coupled nonlinear stochastic differential equations for the speed and turn rate of each fish. Model parameters were calibrated using real data on zebrafish pairs swimming in shallow water by Zienkiewicz et al. (2015a).

The calibrated model was successful in predicting all of the main features of the collective response of live animals. Not only did the animals exhibit strong shoaling and schooling tendencies by swimming in close physical proximity and aligning their bodies, but they also opted for an in-line pattern where one fish would lead and the other would follow. A contribution of this study was to demonstrate local stability of this specific collective pattern through a detailed analysis of the linearized dynamics. Above a critical value of the inter-individual distance, in-line swimming becomes stable for a wide parameter range. This is in stark contrast with side-by-side swimming, which was found to be always unstable.

Taken *in toto*, these results question some of the assumptions that are often used in modeling fish behavior (Vicsek and Zafeiris, 2012), explaining the mechanisms underlying schooling (Partridge and Pitcher, 1979; Weihs, 1973), and drawing inference of leader-follower relationship from experimental observations of their orientation (Krause et al., 2000). First, our theoretical results indicate that repulsion emerges from hydrodynamic interactions, rather than being a social rule pursued by individuals. Second, the long debate regarding the determinants of schooling patterns with some fish occupying frontal positions and other rear ones may reduce to a simple stability problem for the group. Neither reduced swimming costs nor foraging benefits are needed for fish pairs to opt for in-line swimming. Third, the inference of leadership traits from the frequency of instances in which an animal initiates a turn that is followed by others could be skewed by the stability of in-line swimming. Specifically, should the animals differ in their noise parameters, we would detect a false leader-follower association.

There are several directions for potential improvement of the model, including: i) the development of a dynamic model for the speed in the form of a stochastic different equation, similar to Zienkiewicz et al. (2015b); ii) the extension to hydrodynamic interactions with walls in the tank and background water currents toward an improved understanding of the transition from in-line to side-by-side swimming as a function of the flow speed (De Bie et al., 2020); and iii) the extension of individual differences in the stability analysis of in-line and side-by-side swimming that could help explain the modulatory role of fish body size on schooling and shoaling (Karakaya et al., 2020; Reeb, 2001).

Funding Statement. This research was supported by the National Science Foundation (CMMI 1505832 and CMMI 1901697).

Competing Interests. None

Data Availability Statement. Replication data can be found in the Dynamical System Laboratory GitHub: <https://github.com/dynamicalsystemslaboratory/InlineSwimming>.

Ethical Standards. Experiments analyzed herein were originally approved by the University Animal Welfare Committee of New York University under protocol number 13-1424.

Author Contributions. Conceptualization, M.P. and S.D.P.; Methodology, M.P., M.K., R.R.S., and S.D.P.; Software, M.P. and M.K.; Validation, M.P., M.K., and S.D.P.; Formal Analysis, M.P.; Investigation, M.P., M.K., R.R.S., and S.D.P.; Data Curation, M.K.; Writing – Original Draft, M.P. and S.D.P.; Writing – Review Editing Draft, M.P. and S.D.P.; Visualization, M.K.; Supervision, M.P. and S.D.P.; Project Administration, M.P.; Funding Acquisition, M.P. and S.D.P. All authors approved the final submitted draft.

Supplementary Material. Three videos of model predictions, tables of experimental trajectories of all zebrafish, and a table of calibrated model parameters have been provided with the submission.

References

- Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In *Selected Papers of Hirotugu Akaike*, pages 199–213. Springer.
- Aoki, I. (1982). A simulation study on the schooling mechanism in fish. *Bulletin of the Japanese Society of Scientific Fisheries (Japan)*, 48(8):1081–1088.
- Ashraf, I., Bradshaw, H., Ha, T.-T., Halloy, J., Godoy-Diana, R., and Thiria, B. (2017). Simple phalanx pattern leads to energy saving in cohesive fish schooling. *Proceedings of the National Academy of Sciences*, 114(36):9599–9604.

- Butail, S., Mwaffo, V., and Porfiri, M. (2016). Model-free information-theoretic approach to infer leadership in pairs of zebrafish. *Physical Review E*, 93(4):042411.
- 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.
- De Bie, J., Manes, C., and Kemp, P. S. (2020). Collective behaviour of fish in the presence and absence of flow. *Animal Behaviour*, 167:151–159.
- Fillella, A., Nadal, F., Sire, C., Kanso, E., and Eloy, C. (2018). Model of collective fish behavior with hydrodynamic interactions. *Physical Review Letters*, 120(19):198101.
- 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.
- Gazzola, M., Argentina, M., and Mahadevan, L. (2014). Scaling macroscopic aquatic locomotion. *Nature Physics*, 10(10):758–761.
- Gazzola, M., Tchieu, A. A., Alexeev, D., de Brauer, A., and Koumoutsakos, P. (2016). Learning to school in the presence of hydrodynamic interactions. *Journal of Fluid Mechanics*, 789:726–749.
- Herbert-Read, J. E., Perna, A., Mann, R. P., Schaefer, 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.
- Kalueff, A. V., Gebhardt, M., Stewart, A. M., Cachat, J. M., Brimmer, M., Chawla, J. S., Craddock, C., Kyzar, E. J., Roth, A., Landsman, S., et al. (2013). Towards a comprehensive catalog of zebrafish behavior 1.0 and beyond. *Zebrafish*, 10(1):70–86.
- Karakaya, M., Macri, S., and Porfiri, M. (2020). Behavioral teleporting of individual ethograms onto inanimate robots: experiments on social interactions in live zebrafish. *iScience*, 23(8):101418.
- Kato, S., Nakagawa, T., Ohkawa, M., Muramoto, K., Oyama, O., Watanabe, A., Nakashima, H., Nemoto, T., and Sugitani, K. (2004). A computer image processing system for quantification of zebrafish behavior. *Journal of Neuroscience Methods*, 134(1):1–7.
- Katz, Y., Tunström, 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.
- Krause, J., Hoare, D., Krause, S., Hemelrijk, C., and Rubenstein, D. (2000). Leadership in fish shoals. *Fish and Fisheries*, 1(1):82–89.
- 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 B: Biological Sciences*, 284(1852):20170224.
- Ladu, F., Butail, S., Macri, S., and Porfiri, M. (2014). Sociality modulates the effects of ethanol in zebra fish. *Alcoholism: Clinical and Experimental Research*, 38(7):2096–2104.
- Larsson, M. (2012). Why do fish school? *Current Zoology*, 58(1):116–128.
- Meinsma, G. (1995). Elementary proof of the Routh-Hurwitz test. *Systems & Control Letters*, 25(4):237–242.
- Miller, N. Y. and Gerlai, R. (2011). Shoaling in zebrafish: what we don't know. *Reviews in the Neurosciences*, 22(1):17–25.
- 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.
- Mwaffo, V., Zhang, P., Cruz, S. R., and Porfiri, M. (2017). Zebrafish swimming in the flow: a particle image velocimetry study. *PeerJ*, 5:e4041.
- Orger, M. B. and de Polavieja, G. G. (2017). Zebrafish behavior: opportunities and challenges. *Annual Review of Neuroscience*, 40:125–147.
- Partridge, B. and Pitcher, T. (1979). Evidence against a hydrodynamic function for fish schools. *Nature*, 279(5712):418–419.
- Perna, A., Grégoire, G., and Mann, R. P. (2014). On the duality between interaction responses and mutual positions in flocking and schooling. *Movement Ecology*, 2(1):22.
- Pikovsky, A., Kurths, J., Rosenblum, M., and Kurths, J. (2003). *Synchronization: a universal concept in nonlinear sciences*, volume 12. Cambridge university press.
- 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.
- Reebs, S. (2001). Influence of body size on leadership in shoals of golden shiners, *notemigonus crysoleucas*. *Behaviour*, 138(7):797–809.
- Tankov, P. (2003). *Financial modelling with jump processes*. CRC press.
- Tchieu, A. A., Kanso, E., and Newton, P. K. (2012). The finite-dipole dynamical system. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 468(2146):3006–3026.
- Vicsek, T. and Zafeiris, A. (2012). Collective motion. *Physics Reports*, 517(3–4):71–140.
- Weih, D. (1973). Hydromechanics of fish schooling. *Nature*, 241(5387):290–291.
- Zienkiewicz, A., Barton, D., Porfiri, M., and Di Bernardo, M. (2015a). Leadership emergence in a data-driven model of zebrafish shoals with speed modulation. *The European Physical Journal Special Topics*, 224(17):3343–3360.
- Zienkiewicz, A., Barton, D. A., Porfiri, M., and Di Bernardo, M. (2015b). Data-driven stochastic modelling of zebrafish locomotion. *Journal of Mathematical Biology*, 71(5):1081–1105.