An information-theoretic approach to study hydrodynamic interactions in schooling fish

Peng Zhang[a], Elizabeth Krasner[a], Sean D. Peterson[a,b], and Maurizio Porfiri[a,c]

[a]Department of Mechanical and Aerospace Engineering, New York University Tandon School of Engineering, Six MetroTech Center, Brooklyn, NY 11201, USA
[b]Mechanical and Mechatronics Engineering, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L 3G1, Canada
[c]Department of Biomedical Engineering, New York University Tandon School of Engineering, Six MetroTech Center, Brooklyn, NY 11201, USA

ABSTRACT

Understanding the role of hydrodynamic interactions in fish swimming may help explain why and how fish swim in schools. In this work, we designed controlled experiments to study fish swimming in a disturbed flow. Specifically, we recorded the tail beat frequency of a fish swimming in the presence of an actively-controlled airfoil pitching at varying frequencies. We propose an information-theoretic approach to quantify the influence of the motion of the pitching airfoil on the animal swimming. The theoretical framework developed in this work may inform future investigations on the mechanisms underlying schooling in groups.

Keywords: Hydrodynamic interaction, information theory, Shannon entropy, swimming, transfer entropy

1. INTRODUCTION

Schooling behavior is prevalent among fish species.[1,2] Swimming in schools can help individual fish escape from predators, search for food, and mate. The emergence of fish schools has often been associated with the energetic advantages of coordinated swimming.[3,4] The advantage of swimming in schools was first explained theoretically by Weihs.[3] Within Weihs' explanation, fish swimming in a school would preferentially form a diamond pattern, in which fish in a following position could reduce the cost of swimming through interaction with vortices. The ability of fish interacting with vortices has been supported by experimental evidence.[4,5]

However, conflicting findings have challenged our understanding of the role of hydrodynamic interactions in coordinated swimming.[6,7] In particular, a recent study has demonstrated that, at high swimming speed, fish adopt a phalanx pattern, in which they align their body laterally and there is no leader or follower,[7] thereby challenging the possibility of interaction through vortices.

The lack of data-driven techniques has hindered our ability to disentangle causes from effects in coordinated swimming. Information theory can offer a model-free framework for the quantification of interactions between fish. Since the seminal work of Shannon,[8] information theory has been applied to a wide range of research areas, facilitating the study of the effect of drug exposure in social behavior,[9] analysis of financial markets,[10] and energy flow in the climate system,[11] to name a few.

Here, we propose a series of controlled experiments to study fish swimming in the presence of an actively-controlled airfoil. Specifically, we examine live fish swimming against a water flow in the presence of a pitching airfoil positioned in the upstream, simulating the tail beating of a leading fish. Both airfoil pitching and fish swimming are recorded using a high speed camera. The information-theoretic construct of transfer entropy is employed as a quantitative measure for the influence of the airfoil on fish swimming. The theoretical framework examined in this study may help elucidate the role of hydrodynamics in fish swimming, which could inform future investigations on the hydrodynamic advantage of swimming in schools.

Further author information: (Send correspondence to M.P.)
M.P.: E-mail: mporfiri@nyu.edu, Telephone: 1 646 997 3681

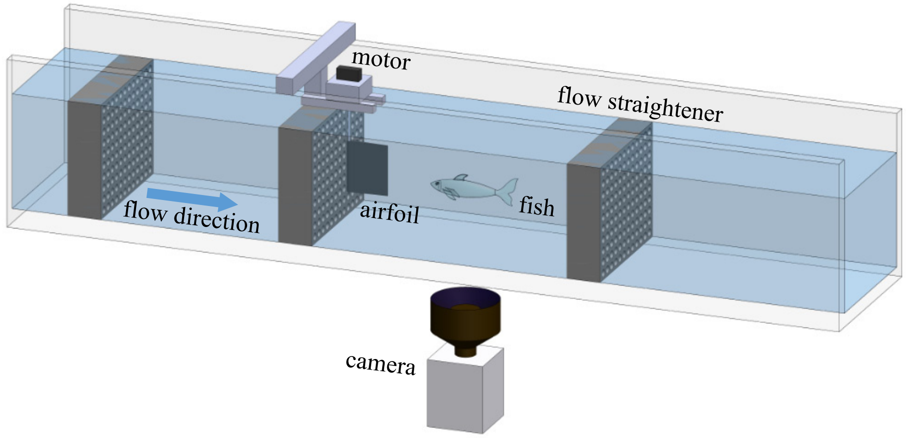

Figure 1. Schematic of the experimental setup. For visualization purposes, the top panel and the metal meshes on the side walls are not shown in the schematic.

2. EXPERIMENTAL METHOD

2.1 Experimental setup

The experimental setup was built on a water tunnel with a $240 \times 15 \times 15$ cm (length$\times$height$\times$width) test section. A swimming area with a length of $L = 30$ cm was book-ended by two honeycomb flow straighteners, as shown in Fig. 1. To mitigate potential confounds due to undesired low fluid velocity regions near the wall, metal meshes were placed on both side walls of the swimming area. The top of the test section was sealed with a white acrylic board, maintaining the height of the test section at approximately 10 cm. The panel created a uniform white background to facilitate image analysis.

A 3D-printed NACA 0012 airfoil with a chord length of 5 cm and a span of 8 cm was placed in the tunnel, upstream of the swimming area. The chord length of the airfoil was chosen to be approximately the body length of the fish used in our experiments. A servomotor was connected to the airfoil through a metal rod to actuate the pitching of the airfoil. The servomotor was controlled by an Arduino Uno microcontroller (Arduino Uno, Arduino, Italy), which was programmed in MATLAB (version R2018a).

A fluorescent light was used to illuminate the test section from the top of the swimming area. All experimental trials were recorded by a Nikon D7000 camera at 30 frames per second with a resolution of 1280×720 pixels and a field of view of $37 \times 21\,\mathrm{cm}^2$.

Prior to the experimental trials, particle image velocimetry (PIV) was conducted to visualize the flow field in the test section in the absence of the airfoil and the fish. PIV results showed uniform velocity profiles for flow speed tested at 5.0 cm/s.

2.2 Experimental procedure

Experiments were performed in accordance with relevant guidelines and regulations approved by the University Animal Welfare Committee (UAWC) of New York University under protocol number 13-1424. The experimental results described herein are part of a wider study, which will be published as a journal paper.

A total of 20 adult giant danios (*Devario aequipinnatus*) were used in our study, each with a body length between 5 and 7 cm. To investigate the influence of a background mean flow, in our experiments, we considered both the no flow condition, $U_0 = 0$ cm/s, and a moderate background flow condition, $U_1 = 5.0$ cm/s, corresponding to approximately one body length per second. At the beginning of each trial, the flow speed (U) in the tunnel was set to zero, and a fish was transported into the swimming area of the water tunnel, followed immediately by a first session of habituation of 5 min. The flow speed was then changed to the desired level of either U_0 or U_1, and was maintained for a second habituation period of 2.5 min. The airfoil was then actuated to pitch between $\pm 5°$ for a third habituation session of 2.5 min. The same condition was maintained for a testing period of 5 min. A total of 10 fish were randomly picked and tested at each flow speed.

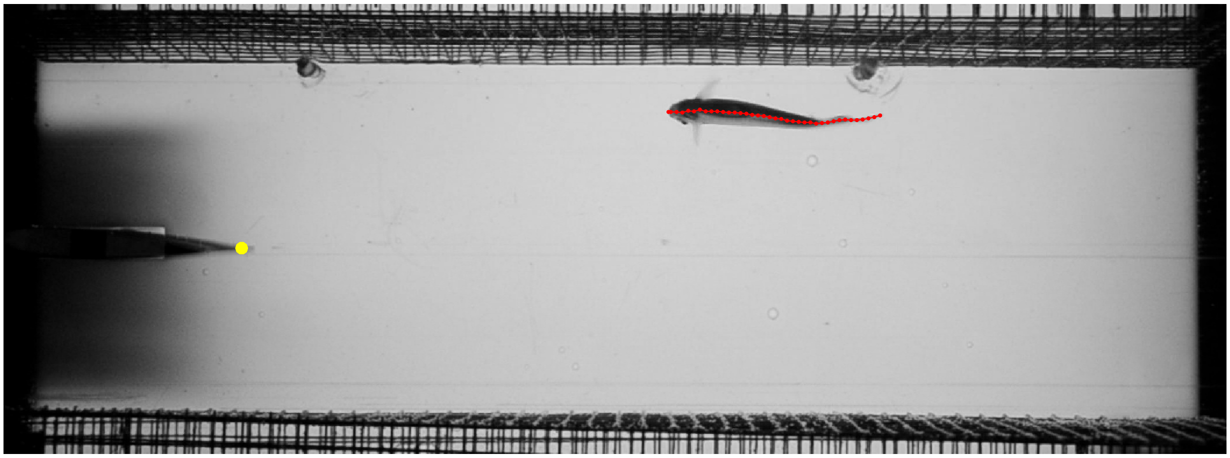

Figure 2. Camera view of the swim tunnel during an experimental test, showing a fish swimming downstream of the airfoil. The center line of the fish body is indicated by red connected dots. The trailing edge of the airfoil is marked by a yellow dot.

To create an information-rich dynamic interaction with hydrodynamically-salient changes, the airfoil was actuated to pitch in the following pattern. During each trial, the pitching frequency of the airfoil was randomly varied among three possible values, $\{\overline{\mathrm{TBF}} - 3\Delta\mathrm{TBF}, \overline{\mathrm{TBF}}, \overline{\mathrm{TBF}} + 3\Delta\mathrm{TBF}\}$. The airfoil was actuated to pitch 10 cycles at each frequency, before switching to the next random value. $\overline{\mathrm{TBF}}$ and $\Delta\mathrm{TBF}$ are the mean and standard deviation of the tail beat frequency of the fish swimming at the corresponding flow speed in the absence of the airfoil, respectively. To determine the values of $\overline{\mathrm{TBF}}$ and $\Delta\mathrm{TBF}$, we conducted pilot tests prior to the experimental trials, which suggested $\overline{\mathrm{TBF}} \approx 3.00\,\mathrm{Hz}$ and $\Delta\mathrm{TBF} \approx 0.25\,\mathrm{Hz}$.

2.3 Data analysis

The recorded videos of the 5 min testing period of the experiment were analyzed using a custom-built MATLAB program, from which time series of the airfoil pitching frequency and the fish tail beat frequency were extracted. Briefly, each frame of the video was binarized based on predetermined thresholds, such that the dark fish body and the trailing edge of the airfoil were converted to black, and the light background to white. The heading and the tail orientation were identified from the tracked centerline of the fish body, and the tail beat angle of the fish was defined as the angle between the heading and tail orientation. The pitching angle of the airfoil was determined from the trailing edge position. Figure 2 displays an exemplary frame of the tracked fish body centerline and the trailing edge position of the airfoil.

The two time series of the airfoil pitching and fish tail beat angles were then converted into series of pitching and tail beat frequencies. The time series of the airfoil pitching angle was discretized into segments of 10 cycles, and each segment was represented by an averaged frequency value. The tail beat angle of the fish was discretized with the same time step as the airfoil pitching angle. Within each segment, we computed an average tail beat frequency by dividing the total number of tail beats by the time span of the segment.

3. THEORETICAL BACKGROUND

Here, we illustrate the dynamics of the airfoil and the fish by their pitching and tail beat frequency, respectively. We consider stationary discrete stochastic processes, $X = \{X_t\}_{t\in\mathbb{Z}^+}$ and $Y = \{Y_t\}_{t\in\mathbb{Z}^+}$, to describe the time series of the airfoil pitching frequency and fish tail beat frequency, respectively.

Within information theory, the uncertainty of the fish tail beat frequency at time step t, Y_t, can be measured through the notion of entropy, originally defined by Shannon[8] as

$$H(Y_t) = -\sum_{y_t\in\Omega^Y} \Pr\{Y_t = y_t\} \log_2 \Pr\{Y_t = y_t\}, \tag{1}$$

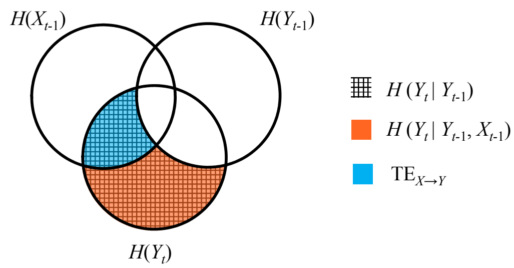

Figure 3. Venn diagram for key information-theoretic quantities.

where $\Pr\{\cdot\}$ denotes probability, y_t is a realization of Y_t, and the set Ω^Y is the sample space that contains all possible realizations of Y_t. The entropy of the airfoil pitching frequency, $H(X_t)$, can be defined similar to Eq. (1). Large values of $H(Y_t)$ would indicate that fish randomly change its tail beat without a preferred beat frequency, while small values of entropy could be associated with fish swimming at a preferred tail beat frequency. Similar to Eq. (1), the joint and conditional entropy of the two random variables, X_t and Y_t, can be defined as

$$H(X_t, Y_t) = -\sum_{x_t\in\Omega^{X_t}, y_t\in\Omega^{Y_t}} \Pr\{X_t = x_t, Y_t = y_t\} \log_2 \Pr\{X_t = x_t, Y_t = y_t\}, \tag{2a}$$

$$H(X_t|Y_t) = -\sum_{x_t\in\Omega^{X_t}, y_t\in\Omega^{Y_t}} \Pr\{X_t = x_t, Y_t = y_t\} \log_2 \Pr\{X_t = x_t|Y_t = y_t\}. \tag{2b}$$

Figure 3 illustrates a few exemplary values of entropy and conditional entropy based on the definitions in Eqs. (1) and (2).

Information flow from the airfoil to the fish can be quantified using the notion of transfer entropy,[12] which is defined as

$$\begin{aligned}\mathrm{TE}_{X\to Y} &= H(Y_t|Y_{t-1}) - H(Y_t|Y_{t-1}, X_{t-1})\\ &= H(Y_t, Y_{t-1}) - H(Y_{t-1}) - H(Y_t, Y_{t-1}, X_{t-1}) + H(Y_{t-1}, X_{t-1})\\ &= \sum_{y_t, y_{t-1}, x_{t-1}} \Pr\{Y_t = y_t, Y_{t-1} = y_{t-1}, X_{t-1} = x_{t-1}\} \log_2 \frac{\Pr\{Y_t = y_t|Y_{t-1} = y_{t-1}, X_{t-1} = x_{t-1}\}}{\Pr\{Y_t = y_t|Y_{t-1} = y_{t-1}\}}\end{aligned} \tag{3}$$

A graphical illustration of transfer entropy is shown in Fig. 3 (blue region). Transfer entropy in Eq. (3) measures the reduction in the uncertainty in predicting the future state of fish tail beat frequency (Y_t) from its present, due to additional knowledge about the present of the airfoil pitching frequency (X_t). In alignment with Wiener's principle of causality,[13] this reduction can be attributed to an information flow from X to Y. For example, if the airfoil does not influence the fish, then additional information of X_{t-1} does not help the prediction of Y_t, that is, $H(Y_t|Y_{t-1}) = H(Y_t|Y_{t-1}, X_{t-1})$, and $\mathrm{TE}_{X\to Y}$ would be zero. On the other hand, if the airfoil influences the fish, then $H(Y_t|Y_{t-1}) > H(Y_t|Y_{t-1}, X_{t-1})$, and $\mathrm{TE}_{X\to Y}$ would be positive.

To estimate the probability mass functions in Eq. (3) , we adopt a symbolic representation of the data.[14,15] Specifically, we introduce a set of binary symbols of 1 and -1, such that for a time series $\{y_t\}_{t=1}^N$ of length N, consecutive data points are associated with the symbol 1 if $y_t \le y_{t+1}$, and -1 if $y_t > y_{t+1}$. Using binary symbols, the computation of transfer entropy in Eq. (3) requires the estimation for the probability of at most $2^3 = 8$ possible triplets.

4. RESULTS

To examine the role of the pitching airfoil in predicting fish swimming, we computed transfer entropy from the airfoil pitching to the fish swimming, $\mathrm{TE}_{X\to Y}$. To test if the values of $\mathrm{TE}_{X\to Y}$ are significantly greater than chance, we compare the values of $\mathrm{TE}_{X\to Y}$ with surrogate data sets. Specifically, under each flow speed, we randomly pair the 10 time series of the airfoil data with the fish data by shuffling their order. We compute 10 values of $\mathrm{TE}_{X\to Y}$ from all random pairs, which are then averaged to obtain a surrogate data point. The same process is repeated 20000 times to generate 20000 surrogate data points, which produced an empirical probability distributions for $\mathrm{TE}_{X\to Y}$. A significance level of 0.05 is used to test if the mean values of $\mathrm{TE}_{X\to Y}$ from the real pairs are significantly greater than chance.

Comparison of $\mathrm{TE}_{X\to Y}$ with surrogate data is shown in Fig. 4. We find that the value of $\mathrm{TE}_{X\to Y}$ at $U = U_0$ is significantly greater than chance ($p = 0.011$) (Fig. 4(a)). In contrast, at $U = U_1$, the magnitude of $\mathrm{TE}_{X\to Y}$ is slightly smaller, and is not significantly greater than change ($p = 0.064$) (Fig. 4(b)).

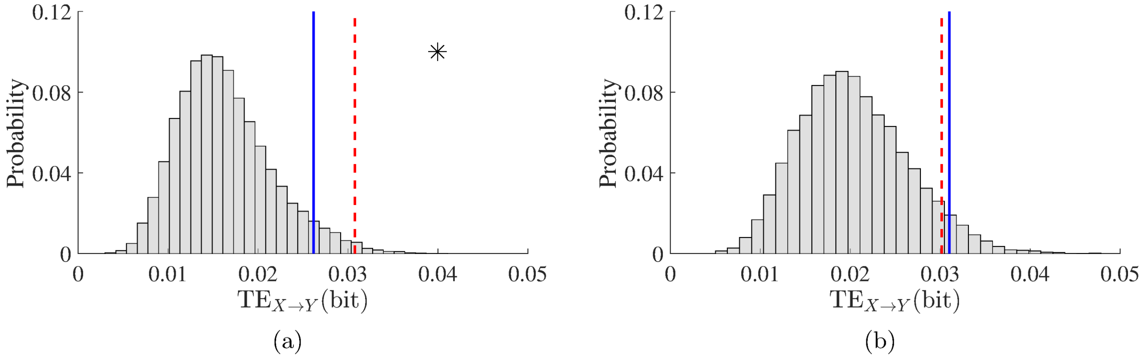

Figure 4. Mean values of $\mathrm{TE}_{X\to Y}$ (red dashed lines) compared with surrogate data set for (a) $U = U_0$ and (b) $U = U_1$. Blue solid lines indicate the 95% quantile of the probability distributions. An asterisk indicates significance with $p \leq 0.050$.

5. DISCUSSION

In this work, we examined the possibility of using the information-theoretic construct of transfer entropy for the study of fish swimming. We designed a series of controlled experiments in which a fish swam in the downstream of an actively pitching airfoil. An information-rich interaction between the airfoil and the fish was generated through randomly switching the pitching frequency of the airfoil. Through the computation of transfer entropy using time series of the airfoil pitching frequency and fish tail beating frequency as inputs, we quantified the extent to which fish swimming could be explained by the pitching of the airfoil.

We found that fish swimming was significantly influenced by the pitching of the airfoil only in the absence of a mean flow. When the flow was increased to one body length per second, no significant influence of the airfoil pitching on fish swimming was found. The interaction between the fish and the airfoil in the absence of a mean background flow could possibly be attributed to fish actively modulating their tail beat in response to coherent flow structures generated by the pitching airfoil.[16,17]

The weak interaction between the airfoil and the fish at the non-zero flow speed may be explained by a competing need of fish to maintain balance and to achieve efficient swimming. Compared to swimming in placid water, maintaining body balance in a mean background flow requires a higher level of energy expenditure to respond to increased drag and more frequent turbulent flow events.[18] Consequently, fish may prioritize the need of maintaining their balance over achieving a favorable hydrodynamic interaction in the flow.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under Grant # CMMI-1505832 and Grant # CMMI-1433670.

REFERENCES

[1] Pitcher, T. J., [*The Behaviour of Teleost Fishes*], Croom Helm Ltd, London & Sydney (1986).
[2] Krause, J., Ruxton, G. D., and Ruxton, G. D., [*Living in Groups*], Oxford University Press (2002).
[3] Weihs, D., "Hydromechanics of fish schooling," *Nature* **241**(5387), 290–291 (1973).

[4] Herskin, J. and Steffensen, J. F., "Energy savings in sea bass swimming in a school: measurements of tail beat frequency and oxygen consumption at different swimming speeds," *Journal of Fish Biology* **53**(2), 366–376 (1998).

[5] Pitcher, T., Magurran, A., and Edwards, J., "Schooling mackerel and herring choose neighbours of similar size," *Marine Biology* **86**(3), 319–322 (1985).

[6] Liao, J. C., "A review of fish swimming mechanics and behaviour in altered flows," *Philosophical Transactions of the Royal Society B: Biological Sciences* **362**(1487), 1973–1993 (2007).

[7] Ashraf, I., Bradshaw, H., Ha, T.-T., Halloy, J., Godoy-Diana, R., and Thiria, B., "Simple phalanx pattern leads to energy saving in cohesive fish schooling," *Proceedings of the National Academy of Sciences* **114**(36), 9599–9604 (2017).

[8] Shannon, C. E., "A mathematical theory of communication," *Bell System Technical Journal* **27** (1948).

[9] Paulus, M. P., Geyer, M. A., Gold, L. H., and Mandell, A. J., "Application of entropy measures derived from the ergodic theory of dynamical systems to rat locomotor behavior," *Proceedings of the National Academy of Sciences* **87**(2), 723–727 (1990).

[10] Marschinski, R. and Kantz, H., "Analysing the information flow between financial time series," *The European Physical Journal B - Condensed Matter and Complex Systems* **30**(2), 275–281 (2002).

[11] Donges, J. F., Zou, Y., Marwan, N., and Kurths, J., "The backbone of the climate network," *Europhysics Letters* **87**(4), 48007 (2009).

[12] Schreiber, T., "Measuring information transfer," *Physical Review Letters* **85**(2), 461–464 (2000).

[13] Wiener, N., "The theory of prediction," *Modern Mathematics for Engineers* **1**, 125–139 (1956).

[14] López, F., Matilla-García, M., Mur, J., and Ruiz Marín, M., "A non-parametric spatial independence test using symbolic entropy," *Regional Science and Urban Economics* **40**(2), 106–115 (2010).

[15] Conaway, M. R., "Analysis of repeated categorical measurements with conditional likelihood methods," *Journal of the American Statistical Association* **84**(405), 53–62 (1989).

[16] Liao, J. C., Beal, D. N., Lauder, G. V., and Triantafyllou, M. S., "The Kármán gait: novel body kinematics of rainbow trout swimming in a vortex street," *Journal of Experimental Biology* **206**(6), 1059–1073 (2003).

[17] Liao, J. C., Beal, D. N., Lauder, G. V., and Triantafyllou, M. S., "Fish exploiting vortices decrease muscle activity," *Science* **302**(5650), 1566–1569 (2003).

[18] Lupandin, A. I., "Effect of flow turbulence on swimming speed of fish," *Biology Bulletin* **32**, 461–466 (Sep 2005).