Efficient collective swimming by harnessing vortices through deep reinforcement learning

Siddhartha Verma¹,², Guido Novati¹, and Petros Koumoutsakos¹,²

¹Computational Science and Engineering Laboratory, ETH Zürich, CH-8092 Zürich, Switzerland

Edited by James A. Sethian, University of California, Berkeley, CA, and approved April 25, 2018 (received for review January 22, 2018)

Fish in schooling formations navigate complex flow fields replete with mechanical energy in the vortex wakes of their companions. Their schooling behavior has been associated with evolutionary advantages including energy savings, yet the underlying physical mechanisms remain unknown. We show that fish can improve their sustained propulsive efficiency by placing themselves in appropriate locations in the wake of other swimmers and intercepting judiciously their shed vortices. This swimming strategy leads to collective energy savings and is revealed through a combination of high-fidelity flow simulations with a deep reinforcement learning (RL) algorithm. The RL algorithm relies on a policy defined by deep, recurrent neural nets, with long–short-term memory cells, that are essential for capturing the unsteadiness of the two-way interactions between the fish and the vortical flow field. Surprisingly, we find that swimming in-line with a leader is not associated with energetic benefits for the follower. Instead, “smart swimmers” place themselves at off-center positions, with respect to the axis of the leader(s) and deform their body to synchronize with the momentum of the oncoming vortices, thus enhancing their swimming efficiency at no cost to the leader(s). The results confirm that fish may harvest energy deposited in vortices and support the conjecture that swimming in formation is energetically advantageous. Moreover, this study demonstrates that deep RL can produce navigation algorithms for complex unsteady and vortical flow fields, with promising implications for energy savings in autonomous robotic swarms.

Significance

Can fish reduce their energy expenditure by schooling? We answer affirmatively this longstanding question by combining state-of-the-art direct numerical simulations of the 3D Navier–Stokes equations with reinforcement learning, using recurrent neural networks with long short-term memory cells to account for the unsteadiness of the flow field. Surprisingly, we find that swimming behind a leader is not always associated with energetic benefits for the follower. In turn, we demonstrate that fish can improve their sustained propulsive efficiency by placing themselves at appropriate locations in the wake of other swimmers and intercepting their wake vortices judiciously. The results show that autonomous, “smart” swimmers may exploit unsteady flow fields to reap substantial energetic benefits and have promising implications for robotic swarms.

Author contributions: S.V., G.N., and P.K. designed research; G.N. performed research; S.V. analyzed data; and S.V., G.N., and P.K. wrote the paper.

The authors declare no conflict of interest.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1800923115/-/DCSupplemental.

¹S.V. and G.N. contributed equally to this work.
²To whom correspondence should be addressed. Email: petros@ethz.ch.

This article is a PNAS Direct Submission.

Published under the PNAS license.

¹S.V. and G.N. contributed equally to this work.
²To whom correspondence should be addressed. Email: petros@ethz.ch.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1800923115/-/DCSupplemental.
Efficient Autonomous Swimmers

We first analyze the kinematics of swimmers IS\(\eta\) and IS\(\delta\) (Fig. 2), which were described previously, and were trained to attain specific high-level objectives via deep RL (see Methods). In both cases, the swimmer trails a leader representing an adult zebrafish of length \(L\), swimming steadily at a velocity \(U\), with tailbeat period \(T\) [Reynolds number \(Re = L^2/(TV) \approx 5000\)]. After training, we observe that IS\(\delta\) is able to maintain its position behind the leader quite effectively (\(\Delta y \approx 0\); Fig. 2D), in accordance to its reward (\(R_\delta = 1 - \vert \Delta y \vert / L\)). Surprisingly, IS\(\eta\) with a reward function proportional to swimming efficiency (\(R_\eta = \eta\)), also settles close to the center of the leader’s wake (Fig. 2D and Movie S4), although it receives no reward related to its relative position. This decision to interact actively with the unsteady wake has significant energetic implications, as described later in the text. Both IS\(\delta\) and IS\(\eta\) maintain a distance of \(\Delta x \approx 2.2L\) from their respective leaders (Fig. 2C). IS\(\eta\) shows a greater proclivity to maintain this separation and intercepts the periodically shed wake vortices just after they have been fully formed and detach from the leader’s tail. In addition to \(\Delta x \approx 2.2L\), there is an additional point of stability at \(\Delta x = 1.5\) (Fig. 2E). The difference 0.7L matches the distance between vortices in the wake of the leader. In both positions, the lateral motion of the follower’s head is synchronized with the flow velocity in the leader’s wake, thus inducing minimal disturbance on the oncoming flow field. We note that a similar synchronization with the flow velocity has been observed when trout minimize muscle use by interacting with vortex columns in a cylinder’s wake (13). IS\(\eta\) undergoes relatively minor body deformation while maneuvering (Fig. 2F), whereas IS\(\delta\) executes aggressive turns involving large body curvature. Trout interacting with cylinder wakes exhibit increased body curvature (27), which is contrary to the behavior displayed by IS\(\eta\). The difference may be ascribed to the widely spaced vortex columns generated by large-diameter cylinders used in the experimental study; weaving in and out of comparatively smaller vortices generated by like-sized fish encountered in a school (Fig. 2B) would entail excessive energy consumption.

We note that maintaining \(\Delta y = 0\) requires significant effort by IS\(\delta\) (SI Appendix, Fig. S2D), which is expected, as this swimmer’s reward (\(R_\delta\)) is insensitive to energy expenditure. One of our previous studies (33) demonstrated that minimizing lateral displacement led to enhanced swimming efficiency (compared with the leader), albeit with noticeable deviation from \(\Delta y = 0\). This conclusion is markedly different from our current observation and can be attributed to the use of improved learning techniques which are better able to achieve the specified goal. In the present study, recurrent neural networks augmented with “long short-term memory” cells (SI Appendix, Fig. S3) help encode time dependencies in the value function and by allowing flapping or swimming motions that closely mimic the interaction of real animals with their environment. This makes them invaluable for investigating concepts that may be carried over readily to bioinspired robotic applications, with minimal modification. This consideration has motivated our present study, where we expand on our earlier work (33), combining RL with direct numerical simulations (DNSs) of the Navier-Stokes (NS) equations for self-propelled autonomous swimmers. We first investigate 2D swimmers in a tandem configuration to scrutinize the strategy adopted by the RL algorithm for attaining the specified goals. Based on the observed behavior and the physical intuition we gain from examining these smart swimmers, we formulate simplified rules for implementing active control in significantly more complex 3D systems. This reverse-engineering approach allows us to determine simple and effective control rules from a data-driven perspective, without having to rely on simplistic models which may introduce errors owing to underlying assumptions.

Deep RL for Swimmers

RL (29) has been introduced to identify navigation policies in several model systems of vortex dipoles, soaring birds and microswimmers (30–32). These studies often rely on simplified representations of organisms interacting with their environment, which allows them to model animal locomotion with reduced physical complexity and manageable computational cost. However, the simplifying assumptions inherent in such models often do not account for feedback of the animals’ motion on the environment. High-fidelity numerical simulations, although significantly more computationally demanding, can account for such important considerations to a greater extent, for instance, swimmers take decisions by virtue of deep RL, using visual cues from their environment (Fig. 2A). The solitary swimmers SS\(\eta\) and SS\(\delta\) execute actions identical to IS\(\eta\) and IS\(\delta\), respectively, and serve as “control” configurations to assess how the absence of a leader’s wake impacts swimming-energetics.
produce far more robust smart swimmers than simpler feedforward networks (35). The performance of our deep recurrent network is compared with that of a feedforward network in SI Appendix, Fig. S4 and indicates that the deep network is better able to achieve the goal of in-line following, but at the penalty of increased energy expenditure. As a result, IS$_S$ succeeds in correcting for oscillations about $\Delta y = 0$ much more effectively by undergoing severe body undulations (Fig. 2F), leading to increased costs (SI Appendix, Fig. S2). These observations confirm that following a leader indiscriminately can be disadvantageous if energetic considerations are not taken into account. Thus, it is unlikely that strict in-line swimming is used as a collective-swimming strategy in nature, and fish presumably adopt a strategy closer to that of IS$_N$, by coordinating their motion with the wake flow. We note that patterns similar to the ones reported in this study have been observed in a recent experimental study (17). The behavior of swimmer IS$_S$ is also compared qualitatively to that of a real fish following a companion in Movie S6, and we observe that the motion of IS$_S$ resembles the swimming behavior of the live follower quite well.

**Intercepting Vortices for Efficient Swimming**

To determine the impact of wake-induced interactions on swimming performance, we compare energetics data for IS$_S$ and SS$_S$ in Fig. 3. The swimming efficiency of IS$_S$, is significantly higher than that of SS$_S$ (Fig. 3A), and the cost of transport (CoT), which represents energy spent for traversing a unit distance, is lower (Fig. 3B). Over a duration of 10 tail-beat periods (from $t = 20$ to $t = 30$; SI Appendix, Fig. S2) IS$_S$ experiences a 11% increase in average speed compared with SS$_S$, a 32% increase in average swimming efficiency and a 36% decrease in CoT. The benefit for IS$_S$ results from both a 29% reduction in effort required for deforming its body against flow-induced forces ($P_{D(isf)}$) and a 53% increase in average thrust power ($P_{Thrust}$). Performance differences between IS$_S$ and SS$_S$ exist solely due to the presence/absence of a preceding wake, since both swimmers undergo identical body undulations throughout the simulations. Comparing the swimming efficiency and power values of four distinct swimmers (SI Appendix, Fig. S2 and Table S1), we confirm that IS$_N$ and SS$_N$ are considerably more energetically efficient than either IS$_S$ or SS$_S$.

The efficient swimming of IS$_N$ [e.g., point $\eta_{max}(A)$ in Fig. 3A] is attributed to the synchronized motion of its head with the lateral flow velocity generated by the wake vortices of the leader (Movie S4s). This mechanism is evidenced by the correlation curve shown in Fig. 2E and by the co-alignment of velocity vectors close to the head in Fig. 4A and B. As shown in Movie S7, IS$_S$ intercepts the oncoming vortices in a slightly skewed manner, splitting each vortex into a stronger ($W_{2U}$, Fig. 4A) and a weaker fragment ($W_{2L}$). The vortices interact with the swimmer’s own boundary layer to generate “lifted vortices” ($L_i$), which in turn generate secondary vorticity ($S_i$) close to the body. Meanwhile, the wake and lifted vortices created during the previous half-period, $W_{2U}$, $W_{2L}$, and $L_2$, have traveled downstream along the body. This sequence of events alternates periodically between the upper (right lateral) and lower (left lateral) surfaces, as seen in Movie S7. Interactions of IS$_N$ with the flow field at
wake or the lifted vortices weaken this distribution by generating vorticity of opposite sign (e.g., secondary vorticity visible in narrow regions between the fish surface and vortices $L_1$, $W_{1U}$, $L_2$, and $L_3$) and create high-speed areas visible as bright spots in Fig. 4A, Lower. The resulting low-pressure region exerts a suction force on the surface of the swimmer (Fig. 4B, Upper), which assists body undulations when the force vectors coincide with the deformation velocity (Fig. 4B, Lower) or increases the effort required when they are counteraligned. The detailed impact of these interactions is demonstrated in Fig. 4C–F. On the lower surface, $W_{1L}$ generates a suction force oriented in the same direction as the deformation velocity ($0 < s < 0.2L$ in Fig. 4B), resulting in negative $P_{Def}$ (Fig. 4E) and favorable $P_{Thrust}$ (Fig. 4F). On the upper surface, the lifted vortex $L_1$ increases the effort required for deforming the body (positive peak in Fig. 4C at $s = 0.2L$), but is beneficial in terms of producing large positive thrust power (Fig. 4D). Moreover, as $L_1$ progresses along the body, it results in a prominent reduction in $P_{Def}$ over the next half-period, similar to the negative peak produced by the lifted vortex $L_2$ ($s = 0.55L$ in Fig. 4E). The average $P_{Def}$ on both the upper and lower surfaces is predominantly negative (i.e., beneficial), in contrast to the minimum swimming efficiency instance $\eta_{max}(D)$, where a mostly positive $P_{Def}$ distribution signifies substantial effort required for deforming the body (SI Appendix, Fig. S5). We observe noticeable drag on the upper surface close to $s = 0$ (Fig. 4B, Upper and Fig. 4D), attributed to the high-pressure region forming in front of the swimmer’s head. Forces induced by $W_{1U}$ are both beneficial and detrimental in terms of generating thrust power ($0 < s < 0.2L$ in Fig. 4F), whereas forces induced by $L_2$ primarily increase drag but assist in body deformation (Fig. 4E). The tail section ($s = 0.8L$ to $1L$) does not contribute noticeably to either thrust or deformation power at the instant of maximum swimming efficiency.

Energy-Saving Mechanisms in Coordinated Swimming

The most discernible behavior of $\dot{S}_S$ is the synchronization of its head movement with the wake flow. However, the most prominent reduction in deformation power occurs near the midsection of the body ($0.4 \leq s \leq 0.7$ in Fig. 4C and E). This indicates that the technique devised by $\dot{S}_S$ is markedly different from energy-conserving mechanisms implied in theoretical (6, 34) and computational (20) work, namely, drag reduction attributed to reduced relative velocity in the flow and thrust increase owing to the “channelling effect.” In fact, the predominant energetics gain (i.e., negative $P_{Def}$) occurs in areas of high

![Fig. 3](image-url) - Energetics data for a smart follower maximizing its swimming efficiency. Swimming efficiency ($A$) and CoT ($B$) for $\dot{S}_S$ (solid blue line) and $SS_0$ (dash-double-dot black line), normalized with respect to the CoT of a steady solitary swimmer. Four instances of maximum and minimum efficiency, which occur periodically throughout the simulation at times ($nT_\eta + 0.12$), ($nT_\eta + 0.37$), ($nT_\eta + 0.62$), ($nT_\eta + 0.87$), have been highlighted. $T_\eta = 1$ denotes an integral multiple. $\Delta$ represents the constant tail-beat period of the swimmers, whereas $n$ represents an integral multiple. $\eta$ at point $E$ ($t \approx 27.5, \eta = 0.86$) results from a erroneous maneuver at $t \approx 26.5$ (Movie S7), which reveals the existence of a time delay between actions and their consequences.

![Fig. 4](image-url) - Flow field and flow-induced forces for $\dot{S}_S$, corresponding to maximum efficiency. (A) Vorticity field (red, positive; blue, negative) with velocity vectors shown as black arrows (Upper) and velocity magnitude shown in Lower (bright, high speed; dark, low speed). The snapshots correspond to $t = 26.12$, i.e., point $\eta_{max}$ (A) in Fig. 3A. Snapshots correspond to every 0.2L along the body center line for reference. The wake vortices intercepted by the follower ($W_{1U}$, $W_{2L}$), the lifted vortices created by interaction of the body with the flow ($L_1$, $L_2$, and $L_3$), and secondary vorticity $S_2$ generated by $L_1$ have been annotated. (B) Flow-induced force vectors (Upper) and body deformation velocity (Lower) at $t = 26.12$. (C and D) Deformation power (C) and thrust power (D) (with negative values indicating drag power) acting on the upper surface of follower. The red line indicates the average over 10 different snapshots ranging from $t = 30.12$ to $t = 39.12$. The envelope signifies the SD among the 10 snapshots. (E and F) Deformation power (E) and thrust power (F) on the lower (left lateral) surface of the swimmer.

We observe that the swimmer’s upper surface is covered in a layer of negative vorticity (and vice versa for the lower surface) (Fig. 4A, Upper) owing to the no-slip boundary condition. The points $\eta_{max}(D)$ and $\eta_{min}(E)$ in Fig. 3A are analyzed separately in SI Appendix, Figs. S5 and S6.

[20] Verma et al.
leader where the flow velocity would match a follower’s head motion (SI Appendix, Fig. S7). A feedback controller is then used to regulate the undulations of two followers to maintain these target coordinates on either branch of the diverging wake, as shown in Fig. 1B and Movie S1. We note that a fish following in-line behind the leader would not benefit in the present 3D simulations, since the region behind the leader remains quiescent owing to the diverging wake. The controlled motion yields an 11% increase in average swimming efficiency for each of the followers (Fig. 5A) and a 5% reduction in each of their CoT. Overall, the group experiences a 7.4% increase in efficiency when compared with three isolated noninteracting swimmers. The mechanism of energy savings closely resembles that observed for the 2D swimmer; an oncoming WR (Fig. 5B) interacts with the deforming body to generate a “lifter-vortex” ring (LR; Fig. 5C). As this new ring proceeds along the length of the body, it modulates the follower’s swimming efficiency as observed in Fig. 5. Remarkably, the positioning of the lifted ring at the instants of minimum and maximum swimming efficiency resembles the corresponding positioning of lifted vortices in the 2D case; a slight dip in efficiency corresponds to lifted vortices interacting with the anterior section of the body (Fig. 5C and SI Appendix, Fig. S5), whereas an increase occurs upon their interaction with the midsection (Figs. 4 and S5D).

These results showcase the capability of machine learning, and deep RL in particular, for discovering effective solutions to complex physical problems with inherent spatial and temporal nonlinearities, in a completely data-driven and model-free manner. Deep RL is especially useful in scenarios where decisions must be taken adaptively in response to a dynamically evolving environment, and the best control strategy may not be evident a priori due to unpredictable time delay between actions and their effect. This necessitates the use of recurrent networks capable of encoding time dependencies, which can have a demonstrable impact on the physical outcome, as shown in SI Appendix, Fig. S4. In conclusion, we demonstrate that deep RL can produce efficient navigation algorithms for use in complex flow fields, which in turn can be used to formulate control rules that are effective in decidedly more complex settings and thus have promising implications for energy savings in autonomous robotic swarms.

Methods

We perform 2D and 3D simulations of multiple self-propelled swimmers using wavelet adapted vortex methods to discretize the velocity-vorticity form of the NS equations (in 2D) and their velocity pressure form along with the pressure-projection method (in 3D) using finite differences on a uniform computational grid. The swimmers adapt their motion using deep RL. The learning process is greatly accelerated by using recurrent neural networks with long short-term memory as a surrogate of the value function for the smart swimmer. Details regarding the simulation methods and the RL algorithm are provided in SI Appendix.

ACKNOWLEDGMENTS. This work was supported by European Research Council Advanced Investigator Award 341117 and Swiss National Science Foundation Sinergia Award CRSII3 147675. Computational resources were provided by Swiss National Supercomputing Centre (CSCS) Project s658.

7. Shaw E (1978) Schooiling fishes: The school, a truly egalitarian form of organization in which all members of the group are alike in influence, offers substantial benefits to its participants. Am Sci 66:166–175.


