

LETTER | JULY 01 2025

## Optimal navigation of magnetic artificial microswimmers in blood capillaries with deep reinforcement learning

Lucas Amoudruz  ; Sergey Litvinov  ; Petros Koumoutsakos  



*Physics of Fluids* 37, 071703 (2025)  
<https://doi.org/10.1063/5.0274623>



### Articles You May Be Interested In

Stable upstream swimming of a swarm of puller microswimmers

*Physics of Fluids* (June 2025)

Can flocking aid the path planning of microswimmers in turbulent flows?

*Physics of Fluids* (April 2025)

Synthesis of magneto-responsive microswimmers for biomedical applications

*AIP Advances* (February 2021)



**Physics of Fluids**  
Special Topics  
Open for Submissions

[Learn More](#)

 AIP Publishing

# Optimal navigation of magnetic artificial microswimmers in blood capillaries with deep reinforcement learning

Cite as: Phys. Fluids **37**, 071703 (2025); doi: [10.1063/5.0274623](https://doi.org/10.1063/5.0274623)

Submitted: 7 April 2025 · Accepted: 13 May 2025 ·

Published Online: 1 July 2025



View Online



Export Citation



CrossMark

Lucas Amoudruz,  Sergey Litvinov,  and Petros Koumoutsakos<sup>a)</sup> 

## AFFILIATIONS

School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA

<sup>a)</sup>Author to whom correspondence should be addressed: [petros@seas.harvard.edu](mailto:petros@seas.harvard.edu)

## ABSTRACT

Biomedical applications, such as targeted drug delivery, microsurgery, and sensing, rely on reaching precise areas within the body in a minimally invasive way. Artificial bacterial flagella (ABFs) have emerged as potential tools for this task by navigating through the circulatory system with the help of external magnetic fields. While their swimming characteristics are well understood in simple settings, their controlled navigation through realistic capillary networks remains a significant challenge due to the complexity of blood flow and the high computational cost of detailed simulations. We address this challenge by conducting numerical simulations of ABFs in retinal capillaries, propelled by an external magnetic field. The simulations are based on a validated blood model that predicts the dynamics of individual red blood cells and their hydrodynamic interactions with ABFs. The magnetic field follows a control policy that brings the ABF to a prescribed target. The control policy is learned with an actor-critic, off-policy reinforcement learning algorithm coupled with a reduced-order model of the system. We show that the same policy robustly guides the ABF to a prescribed target in both the reduced-order model and the fine-grained blood simulations. This approach is suitable for designing robust control policies for personalized medicine at moderate computational cost.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0274623>

Targeted drug delivery, microsurgery, and microsensing represent important areas of research aimed at revolutionizing precision medicine.<sup>1,2</sup> These challenging tasks require precise access to target areas in the body, which are often difficult to reach in a noninvasive way. During the past decade, artificial microswimmers have emerged as potent candidates to navigate specific regions of the human body through the circulatory system or tissues.<sup>3</sup> A particular form of artificial microswimmers, known as artificial bacterial flagella (ABFs), features a cork-screw-shaped body propelled by rotating magnetic fields.<sup>4</sup> ABFs have been used in various applications, including navigation through ex vivo bovine eye tissues,<sup>5</sup> assisting spermatozoa in reaching ovocytes,<sup>6</sup> enhancing nanoparticle transport,<sup>7</sup> and carrying drugs for cancer therapy.<sup>8</sup>

The swimming properties of ABFs are well understood in the idealized scenario of unbounded, viscous Newtonian fluids,<sup>9,10</sup> near walls,<sup>11</sup> and near other swimmers.<sup>12</sup> Venugopalan *et al.* demonstrated that ABFs can propel in low-concentrated blood suspensions.<sup>13</sup> Alapan *et al.* guided microrollers in physiologically relevant blood concentrations with a rotating magnetic field.<sup>14</sup> Qi *et al.* conducted

numerical simulations of these rollers through blood in straight pipes.<sup>15</sup> However, the swimming characteristics of ABFs within the bloodstream have not been thoroughly investigated. The control of microswimmers for path planning has been the subject of recent advances with the emergence of reinforcement learning (RL). Muñoz-Landin *et al.* used RL in an experimental setup to control a microswimmer toward a target, without background flow.<sup>16</sup> Similarly, several studies applied RL to guide microswimmers<sup>17</sup> through background flows<sup>18–22</sup> and perform independent control of multiple ABFs<sup>23,24</sup> to reach a target or follow a prescribed path.<sup>25</sup> Yang *et al.* achieved path-planning of point particles between obstacles represented by red blood cells (RBCs).<sup>26</sup> However, the numerical models used in these studies do not reflect the complex flow patterns present in capillaries, the finite size of the swimmer, and the presence of deformable blood cells, which are known to strongly affect blood flow.<sup>27</sup>

In this work, we model and simulate an ABF that evolves in a retinal capillary network. In particular, we learn a control policy that guides the ABF to a prescribed target by imposing an external

magnetic field. The geometric representation of the environment is reconstructed from a fundus image of a retinal capillary network.<sup>28</sup> The simulations include an accurate RBC model that was extensively validated on experimental data,<sup>29</sup> coupled with the dissipative particle dynamics (DPD) method to resolve the hydrodynamics at the micro-scale.<sup>30</sup> The control policy is learned with the actor-critic, off policy RL algorithm V-RACER.<sup>31</sup> To substantially reduce the computational cost associated with expensive simulations of the environment, the RL agent is trained on a reduced order model of fine-grained blood simulations. The approach of training the RL agent on a reduced order model has been used in the context of robotics to reduce the number of experiments or expensive simulations.<sup>32,33</sup> We demonstrate that the policy is robust to noise and transferable to the fine-grained and expensive simulation environment. These results indicate that RL is a promising approach to design control policies with the level of robustness required in biomedical applications. Furthermore, this work demonstrates that ABFs can evolve within the intricate geometry and flow fields of the retinal capillaries to reach precise locations with minimal invasiveness.

The simulations consist of an ABF evolving through a network of capillaries filled with blood at a hematocrit  $Ht = 25\%$ . The ABF has a radius of  $2 \mu\text{m}$  and a length of  $18.37 \mu\text{m}$ , within the range of the ABFs presented in the literature.<sup>9,10</sup> Blood is modeled as RBC membranes suspended in plasma and enclosing the cytosol, a fluid five times more viscous than plasma.<sup>34</sup> Both plasma and cytosol are viscous Newtonian fluids and are modeled with the DPD method.<sup>30</sup> RBC membranes evolve according to bending forces, and shear forces that are zero when the membrane takes the shape of the stress-free state of the RBC cytoskeleton.<sup>29,35,36</sup> An example of simulation of an ABF swimming through a straight tube filled with blood is shown in Fig. 1.

The bending energy of the membrane is given by

$$U_{bending} = 2\kappa_b \oint H^2 dA, \quad (1)$$

where the integral is taken over the membrane surface,  $\kappa_b$  is the bending modulus, and  $H$  is the local mean curvature. The energy related to the membrane deformation with respect to the stress-free shape (SFS) is given by

$$\begin{aligned} U_{in-plane} = & \frac{K_x}{2} \oint (\alpha^2 + a_3 \alpha^3 + a_4 \alpha^4) dA_0 \\ & + \mu \oint (\beta + b_1 \alpha \beta + b_2 \beta^2) dA_0, \end{aligned} \quad (2)$$

where the integral is taken over the surface of SFS,  $\alpha$  and  $\beta$  are the local dilation and shear strain invariants of the membrane, respectively,  $K_x$

is the dilation elastic modulus,  $\mu$  is the shear elastic modulus, and the coefficients  $a_3$ ,  $a_4$ ,  $b_1$ , and  $b_2$  are parameters that control the non-linearity of the membrane elasticity for large deformations.<sup>36</sup>

In addition, the area and volume of the RBC membranes are restricted by penalization energies,<sup>37</sup>

$$U_{area} = k_A \frac{(A - A_0)^2}{A_0}, \quad U_{volume} = k_V \frac{(V - V_0)^2}{V_0}, \quad (3)$$

where  $A_0$  and  $V_0$  are the target area and volume of the cell and  $A$  and  $V$  are the area and volume of the cell, respectively. Furthermore,  $k_A$  and  $k_V$  are coefficients that were chosen empirically to enforce small variations of the area and volume over time, within 1%.

Each membrane surface is discretized into a triangle mesh composed of  $N_v$  vertices with positions  $\mathbf{r}_i$ , velocities  $\mathbf{v}_i$ , and mass  $m_i$ ,  $i = 1, 2, \dots, N_v$ , evolving through time according to Newton's law of motion. The forces acting on each particle are computed by taking the negative gradient of the discretized energies listed above with respect to particle positions. Finally, the membrane viscosity is modeled through pairwise forces between particles that share an edge,

$$\mathbf{f}_{ij}^{visc} = -\frac{4\eta_m}{\sqrt{3}} (\mathbf{v}_{ij} \cdot \mathbf{e}_{ij}) \mathbf{e}_{ij}, \quad (4)$$

where  $\eta_m$  is the membrane viscosity,  $\mathbf{v}_{ij} = \mathbf{v}_i - \mathbf{v}_j$ , and  $\mathbf{e}_{ij}$  is the unit vector between  $\mathbf{r}_i$  and  $\mathbf{r}_j$ . The details about the fluid-structure interactions and model parameters are described in Ref. 38 with extensive validation against the experimental data.<sup>29</sup>

The ABF is represented as a set of frozen particles and a surface moving as a rigid body. The surrounding solvent particles interact with the ABF particles through DPD forces and are bounced-back from the surface. The magnetic moment  $\mathbf{m}$  of the ABF remains constant in the reference frame of the swimmer, perpendicular to its principal axis. The ABF is immersed in an external, uniform magnetic field  $\mathbf{B}$ , creating a magnetic torque  $\mathbf{T} = \mathbf{m} \times \mathbf{B}$ . This torque, combined with the cork-screw shape of the ABF's tail, causes the ABF to propel along its main axis.<sup>9</sup> The external magnetic field varies over time and will be the controlled quantity used to stir the ABF toward its target.

The boundaries of the capillaries were generated from a human retinal vasculature fundus image in Ghassemi *et al.*<sup>28</sup> (see the supplementary material). The walls are formed by a layer of frozen DPD particles, of width larger than the interaction cutoff. Furthermore, the DPD particles are bounced-back from the wall surface. The parameters of the DPD forces and interactions between every objects of the simulations are detailed in Amoudruz.<sup>38</sup> To drive the flow, an external body force is applied to the DPD particles, following the approach described in Yazdani *et al.*<sup>39</sup> The body force is derived from the pressure gradient

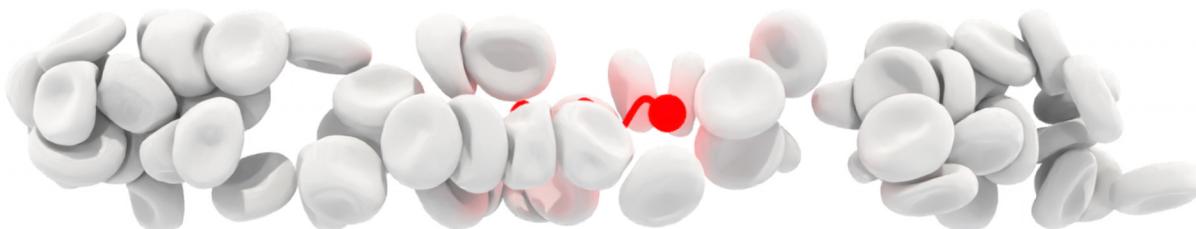


FIG. 1. Simulation snapshot of an ABF inside a periodic tube filled with blood. Tube boundaries not shown for visualization purpose.

obtained by solving the Stokes equation in the same geometry. The flow is solved with *Aphros*,<sup>40</sup> assuming a Newtonian fluid, no-slip boundary conditions at the geometry boundaries, and a constant velocity at the inlet. In the fine-grained blood simulations, the body forces were scaled to match the Reynolds number at the inlet of the domain. Blood simulations are performed with *Mirheo*,<sup>41</sup> a high-performance package for microfluidic simulations on multi-graphics processing unit (GPU) architectures.

Figure 2 shows the swimming speed of a single ABF in a viscous fluid against the rotation frequency of the external magnetic field. The ABF used for this validation case is a helix of diameter  $5.294 \mu\text{m}$  and length  $15.75 \mu\text{m}$  with three turns and an ellipsoidal cross section with main diameters  $1.192$  and  $2.231 \mu\text{m}$ . This shape was reproduced to match that of the ABF in experiments from Mhanna *et al.*<sup>42</sup> The magnetic moment of the swimmer was set to  $1 \times 10^{-14} \text{ NmT}^{-1}$ , tuned to match the experimental data, and the magnetic field magnitude to  $3 \text{ mT}$ . Simulations are in good agreement with the experiments by Mhanna *et al.*,<sup>42</sup> validating the DPD method to simulate swimming ABFs in fluids at low Reynolds numbers.

In this work, the ABF has a helical tail and a spherical head, designed to prevent the ABF from getting stuck in RBCs. To increase the swimming speed of the ABF, both the magnetic field frequency and the magnetic moment of the ABF were increased to  $f = 1 \text{ kHz}$  and  $1 \times 10^{-11} \text{ NmT}^{-1}$ , respectively. These adjustments enable the ABF to reach swimming speeds comparable to the inlet flow velocity  $U_{in} = 1 \text{ mm s}^{-1}$ . Such a high magnetic moment can be achieved by using a magnetic head composed, for example, of nickel, which has a magnetization of  $M_{Ni} = 4.9 \times 10^5 \text{ Am}^{-1}$ . For a spherical head of radius  $R = 2 \mu\text{m}$ , this yields a magnetic moment of  $m = M_{Ni} 4\pi R^3 / 3 \approx 1.6 \times 10^{-11} \text{ NmT}^{-1}$ , demonstrating that the values used in simulations are within experimentally achievable values. Importantly, even at this increased rotation frequency  $f = 1 \text{ kHz}$ , the Reynolds number remains low,  $\text{Re} = R^2 f / \nu \approx 0.004$ , justifying the use of the Stokes equation. Therefore, the validation shown in Fig. 2 indicates that the DPD method remains appropriate to predict the evolution of ABFs rotating with an angular frequency of  $f = 1 \text{ Hz}$ .

In this section, we describe the procedure used to learn, with RL, a control policy for guiding the ABF through the retinal capillary network. The goal is to bring the ABF from the network entry point to a

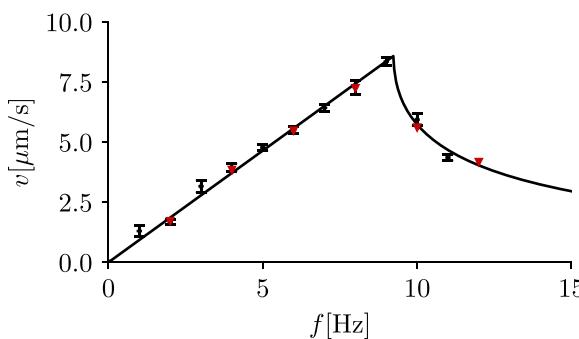


FIG. 2. Swimming speed of the ABF against the rotation frequency of the external magnetic field. Triangles are obtained with DPD simulations, crosses are from the experiments of Mhanna *et al.*,<sup>42</sup> and solid line is a fit to the experimental data with an ODE model<sup>43,44</sup> (see the supplementary material).

specified target location. This is achieved by varying the external magnetic field over time. RL algorithms typically require tens of thousands to millions of episodes to converge. However, each fine-grained simulation of the ABF in the capillary network has a high computational cost of 64 P100 GPUs over 24 h. Therefore, it is intractable to train the RL agent directly on the high-fidelity model. Instead, the RL agent is coupled with a reduced-order model during training.

The reduced order model represents the ABF as a self-propelling point particle advected by a background velocity field  $\mathbf{u}$ . The background velocity field is obtained by solving the Stokes equation within the geometry of the retinal network. The Stokes equation was computed with the grid-based solver *Aphros*. In addition, collisions between the ABF and surrounding RBCs are modeled by a stochastic term, resulting in the stochastic differential equation (SDE),

$$\dot{\mathbf{x}} = \mathbf{u}(\mathbf{x}) + U\mathbf{p} + \sqrt{D}\xi, \quad (5)$$

where  $\mathbf{p}$  and  $U$  are the direction and magnitude of the ABF self-propelling velocity, respectively.  $D$  is a diffusion coefficient, and  $\xi$  is a Gaussian white noise vector. In addition, the particle bounces back from the wall boundaries. The parameters  $D$  and  $U$  were calibrated from high-fidelity blood flow simulations of a single ABF in a straight pipe of radius typically found in capillaries (see the supplementary material). We neglect the rotation dynamics of the ABF, as the reorientation timescale  $1/f = 10^{-3} \text{ s}$  is 20 times faster than the typical timescale associated with flow velocity gradients,  $2R_{in}/U_{in} = 2 \times 10^{-2} \text{ s}$ .

The system is advanced in time with a piecewise constant action  $\mathbf{p}$ , updated every  $\Delta t$  units of time. The direction  $\mathbf{p}$  is computed from the ABF's position  $\mathbf{x}$  through the control policy,  $\mathbf{p} = \hat{\mathbf{p}}/|\hat{\mathbf{p}}|$ ,  $\hat{\mathbf{p}} = \pi(\mathbf{x})$ . Each episode ends if the simulation time exceeds a maximum time  $T_{max}$  or if the ABF reaches the target within a distance  $\delta$ . The initial position of the ABF at the beginning of each episode is sampled from a set of positions that are along the downstream vessels of the target (see the supplementary material). The reward at step  $t$  is expressed as

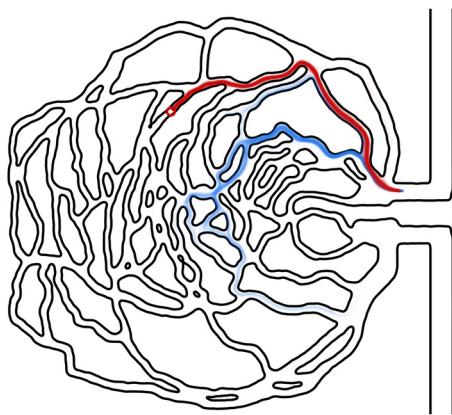
$$r_t = -C\Delta t + ||\mathbf{x}_{t-1} - \mathbf{x}_{\text{target}}|| - ||\mathbf{x}_t - \mathbf{x}_{\text{target}}||, \quad (6)$$

where  $C > 0$  is a constant and  $\mathbf{x}_{\text{target}}$  is the target position. The first term in Eq. (6) penalizes long trajectories, while the second term is a reward shaping that is positive when the ABF progresses toward its target.<sup>45</sup>

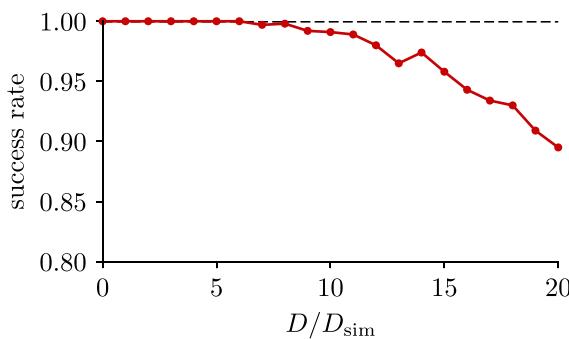
We train the policy using V-RACER,<sup>31</sup> an actor-critic off-policy RL algorithm that was applied to falling objects,<sup>46</sup> magnetic microswimmers,<sup>23</sup> and self-propelling fish.<sup>47,48</sup> The algorithm is implemented in *Korali*.<sup>49</sup> Each episode consists of about 500 experiences, and we train the agent over 10 000 episodes.

To test the policy, we generate 100 trajectories with Eq. (5), both with the trained policy and for passive tracers ( $U=0$ ). These trajectories are shown in Fig. 3. Passive tracers are much more sensitive to noise than controlled swimmers, and their trajectories end at various locations in the geometry. Passive tracers do not reach the target, as opposed to swimmers that follow the RL policy.

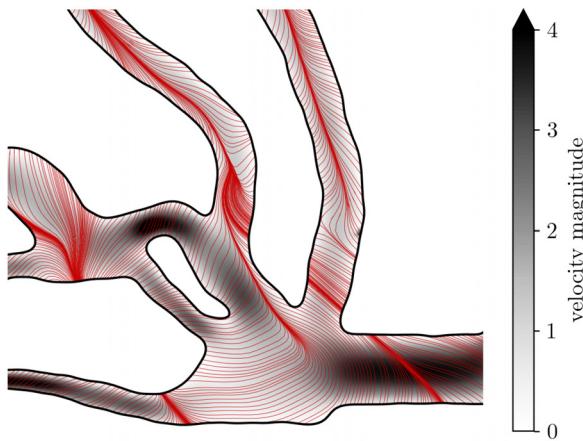
The robustness of the control policy is assessed by varying the magnitude of noise  $D$  and measuring the success rate, defined as the number of times the particle reaches the target out of 1000 trials divided by the number of trials. Figure 4 demonstrates that the control policy brings the swimmer to its target more than 98% of the time when  $D < 10D_{\text{sim}}$ , and 100% of the time when the noise is comparable to that measured from the fine-grained simulations,  $D_{\text{sim}}$ . Increasing the noise level further results in lower success rates. However, the



**FIG. 3.** Trajectories of passive tracers (blue) and controlled swimmers (red) obtained from 100 random seeds, with the reduced order model and  $D = D_{\text{sim}}$ . The circle represents the target. The inlet and outlet have a diameter of  $20 \mu\text{m}$ .



**FIG. 4.** Success rate of the control policy against the noise level  $D$  relative to that measured from the fine-grained simulations  $D_{\text{sim}}$ .



**FIG. 5.** Streamlines of the policy at two bifurcations along the optimal path. The direction chosen by the agent is parallel to the streamlines. The background colors indicate the background velocity magnitude.

success rate remains above 85% when  $D \leq 20D_{\text{sim}}$ . In comparison, passive tracers never reached the target for any values  $D \in [0, 20D_{\text{sim}}]$  in our experiments.

[Figure 5](#) shows the policy at bifurcations. The streamlines of the directions obtained from the policy,  $\mathbf{p} = \hat{\mathbf{p}}/|\hat{\mathbf{p}}|$ ,  $\hat{\mathbf{p}} = \pi(\mathbf{x})$ , converge to a line that connects the starting point to the target. This line seems to favor regions where the background velocity is the largest and aligned with the swimming direction, thus decreasing the overall travel time of the swimmer. When the swimmer is away from this line, the agent chooses a direction that brings the swimmer back to the line. This mechanism explains the robustness of the control policy to external perturbations.

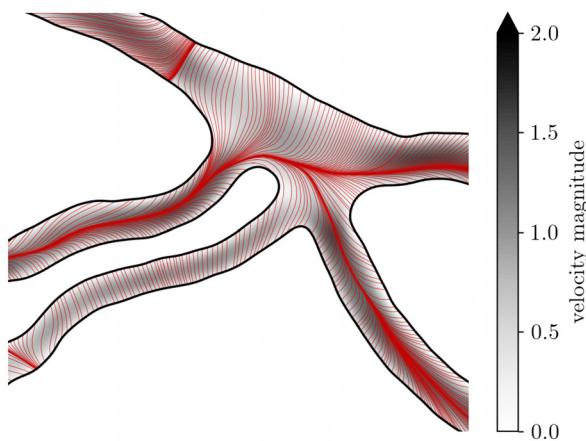
[Figure 6](#) shows the streamlines of the policy, together with the state-value function estimated during learning. The state-value function takes low values in regions where the swimmer cannot recover due to strong flow fields. Hence, in these regions, the policy is not meaningful. In regions away from the optimal path and with a higher state-value function, the policy has a similar behavior to the policy along the optimal path: the streamlines converge near the centerline of the vessel, where the velocity is higher.

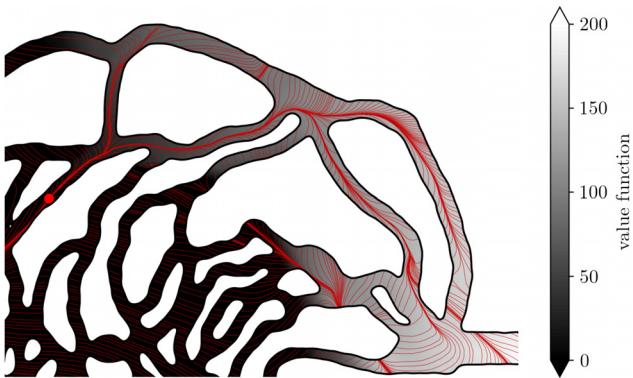
The control policy trained on the reduced-order model is now tested in the fine-grained model. The policy maps the swimmer's position to the swimming direction  $\mathbf{p}$ . However, in the fine-grained simulations, the swimming direction is not set directly and we instead control the magnetic field. Given that ABFs align perpendicularly to the plane of the magnetic field's rotation, we adjust the magnetic field as

$$\mathbf{B}(t) = BR_x(\mathbf{p}) \begin{pmatrix} 0 \\ \cos \omega t \\ \sin \omega t \end{pmatrix}, \quad (7)$$

where  $B$  and  $\omega$  are the magnitude and frequency of rotation of the magnetic field, respectively, and  $R_x(\mathbf{p})$  is the rotation that transforms the vector  $\mathbf{e}_x$  into the swimming direction  $\mathbf{p}$  with axis of rotation  $\mathbf{e}_x \times \mathbf{p}$ . The direction  $\mathbf{p}$  is computed from the policy evaluated at the ABF's center of mass, and the magnetic field is adapted at every time step of the simulation.

[Figure 7](#) (multimedia view) shows that the ABF successfully reaches the target using the policy that was previously learned from





**FIG. 6.** Streamlines of the policy within the capillaries (solid lines) and state value function (shades of grey). The direction chosen by the agent is parallel to the streamlines. The target is represented with the red disk.

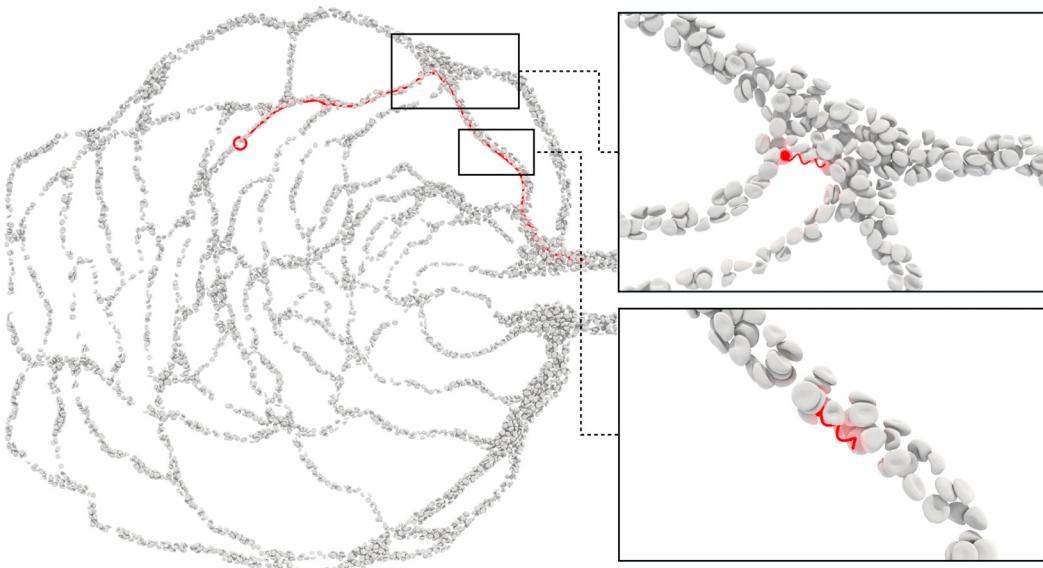
the reduced-order model. The ABF follows a trajectory similar to that observed in the reduced-order model. Additionally, the ABF tends to align with the flow direction and the centerline of the capillaries, suggesting that the trajectory minimizes the travel time. Away from bifurcations, the policy seems to keep the ABF near the center of the capillaries, where the flow is faster. We observe that the ABF remains close to the wall boundaries in the vicinity of bifurcations. This strategy offers two benefits: the ABF is closer to the correct downstream branch, reducing the risk of taking the wrong bifurcation, and the ABF has a larger swimming speed relative to the flow, allowing finer control in these critical areas.

We emphasize that the policy was trained on an environment that is different from the fine-grained simulations but has a similar qualitative response. In the reduced-order model, we assume that the

swimmer immediately reorients to the control output, which is not the case in the fine-grained blood model. Furthermore, the velocity field is different from that obtained from the Stokes solver with Newtonian assumption. Finally, the swimming speed of the ABF with respect to the fluid depends on the local hematocrit, the configuration of the surrounding RBCs, and the geometry of the capillaries, which we ignore in the reduced order model. A comparison of the ABF's coordinates against time between the DPD simulations and the reduced order model is shown in the [supplementary material](#). In both cases, the ABF follows the same trajectory and has a similar travel time. We observe small differences in the swimming velocity that may be due to the surrounding geometry, wall effects, and local hematocrit. Despite these differences, the policy learned by the RL agent in the reduced-order model is successful in the fine-grained simulations. This transfer indicates that RL is an effective method to learn policies that are robust to changes in the environment.

In this study, we assume that the position of the ABF can be estimated in real time. However, in practice, the localization of micrometer-sized objects *in vivo* remains a significant challenge. Promising techniques, such as magnetic particle imaging (MPI),<sup>50</sup> ultrasound imaging,<sup>51</sup> and fluorescence imaging,<sup>52</sup> have shown potential for tracking magnetic microswimmers in biological environments, but further advances in spatial resolution are necessary to enable their practical application.<sup>53</sup>

Furthermore, in this work, we have considered a single ABF, while it would be beneficial to control multiple ABFs to achieve targeted drug delivery or microsurgery. Controlling multiple ABFs of different shapes with a uniform magnetic field has been explored in the absence of wall boundaries,<sup>23</sup> but this approach is currently limited to a few ABFs. Another potential approach to control a swarm of ABFs consists in exploiting the collective behavior of micro-objects through their magnetic interactions,<sup>54</sup> although the current research considers



**FIG. 7.** Trajectory of the ABF in the blood simulations with feedback control via the RL policy. The ABF reaches the target (red circle) and follows a trajectory (red line) similar to that obtained with the reduced order model. The zoomed-in views show the ABF swimming in the capillaries. Multimedia available online.

ideal setups of microrollers close to flat surfaces and no background flows.

We have performed simulations of the evolution of ABFs in the bloodstream through complex capillaries. The simulations employ a fine-grained blood model coupled with the DPD method to resolve the flow mechanics. The ABF is propelled via an external magnetic field, controlled by a RL agent with the aim of guiding the ABF toward a prescribed target. The agent was trained on a reduced order model that was calibrated from fine-grained blood simulations. This approach is much less computationally demanding than the use of blood simulations in the training phase. The control policy is robust to external noise in the reduced-order environment and is successful nearly 100% of the time for relatively large perturbations. Interestingly, the same control policy is also successful in fine-grained blood simulations.

These results demonstrate that control policies trained on reduced-order models are potent candidates for control in fine-grained simulations and possibly in real conditions, while requiring moderate computational resources. Furthermore, this approach enables the design of personalized control strategies based on patient-specific data, as we have shown in this case by using the reconstruction of capillary networks generated from fundus images. This method can be generalized to the navigation of microswimmers for targeted drug delivery and microsurgery at precise locations, potentially becoming a critical tool for personalized medicine.

## SUPPLEMENTARY MATERIAL

See the [supplementary material](#) for details about the generation of the geometry file, a description about the calibration of the reduced-order model from DPD simulations, the parameters of the reinforcement learning algorithm, a comparison between the models, and analysis of the motion of the ABF in the DPD simulations.

We acknowledge the computational resources granted by the Swiss National Supercomputing Center (CSCS) under the project ID “s1160.”

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

### Author Contributions

**Lucas Amoudruz:** Conceptualization (equal); Investigation (equal); Methodology (equal); Software (equal); Validation (equal); Visualization (equal); Writing – original draft (equal). **Sergey Litvinov:** Conceptualization (equal); Investigation (equal); Methodology (equal); Software (equal); Validation (equal); Writing – original draft (equal). **Petros Koumoutsakos:** Conceptualization (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Project administration (equal); Supervision (equal); Writing – original draft (equal).

## DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

## REFERENCES

- <sup>1</sup>F. Soto, J. Wang, R. Ahmed, and U. Demirci, “Medical micro/nanorobots in precision medicine,” *Adv. Sci.* **7**, 2002203 (2020).
- <sup>2</sup>M. T. Manzari, Y. Shamay, H. Kiguchi, N. Rosen, M. Scaltriti, and D. A. Heller, “Targeted drug delivery strategies for precision medicines,” *Nat. Rev. Mater.* **6**, 351 (2021).
- <sup>3</sup>A.-I. Bunea and R. Taboryski, “Recent advances in microswimmers for biomedical applications,” *Micromachines* **11**, 1048 (2020).
- <sup>4</sup>R. Dreyfus, J. Baudry, M. L. Roper, M. Fermigier, H. A. Stone, and J. Bibette, “Microscopic artificial swimmers,” *Nature* **437**, 862 (2005).
- <sup>5</sup>J. Yu, D. Jin, K.-F. Chan, Q. Wang, K. Yuan, and L. Zhang, “Active generation and magnetic actuation of microrobotic swarms in bio-fluids,” *Nat. Commun.* **10**, 5631 (2019).
- <sup>6</sup>M. Medina-Sánchez, L. Schwarz, A. K. Meyer, F. Hebenstreit, and O. G. Schmidt, “Cellular cargo delivery: Toward assisted fertilization by sperm-carrying micromotors,” *Nano Lett.* **16**, 555 (2016).
- <sup>7</sup>S. Schuerle, A. P. Soleimany, T. Yeh, G. Anand, M. Häberli, H. Fleming, N. Mirkhani, F. Qiu, S. Hauert, X. Wang *et al.*, “Synthetic and living micropropellers for convection-enhanced nanoparticle transport,” *Sci. Adv.* **5**, eaav4803 (2019).
- <sup>8</sup>X. Wang, J. Cai, L. Sun, S. Zhang, D. Gong, X. Li, S. Yue, L. Feng, and D. Zhang, “Facile fabrication of magnetic microrobots based on Spirulina templates for targeted delivery and synergistic chemo-photothermal therapy,” *ACS Appl. Mater. Interfaces* **11**, 4745 (2019).
- <sup>9</sup>L. Zhang, J. J. Abbott, L. Dong, B. E. Kratochvil, D. Bell, and B. J. Nelson, “Artificial bacterial flagella: Fabrication and magnetic control,” *Appl. Phys. Lett.* **94**, 064107 (2009).
- <sup>10</sup>L. Zhang, J. J. Abbott, L. Dong, K. E. Peyer, B. E. Kratochvil, H. Zhang, C. Bergeles, and B. J. Nelson, “Characterizing the swimming properties of artificial bacterial flagella,” *Nano Lett.* **9**, 3663 (2009).
- <sup>11</sup>M. Pal, I. Fouxon, A. M. Leshansky, and A. Ghosh, “Fluid flow induced by helical microswimmers in bulk and near walls,” *Phys. Rev. Res.* **4**, 033069 (2022).
- <sup>12</sup>R. Dey, C. M. Buness, B. V. Hokmabad, C. Jin, and C. C. Maass, “Oscillatory rheotaxis of artificial swimmers in microchannels,” *Nat. Commun.* **13**, 2952 (2022).
- <sup>13</sup>P. L. Venugopalan, R. Sai, Y. Chandorkar, B. Basu, S. Shivashankar, and A. Ghosh, “Conformal cytocompatible ferrite coatings facilitate the realization of a nanovoyager in human blood,” *Nano Lett.* **14**, 1968 (2014).
- <sup>14</sup>Y. Alapan, U. Bozuyuk, P. Erkoc, A. C. Karacakol, and M. Sitti, “Multifunctional surface microrollers for targeted cargo delivery in physiological blood flow,” *Sci. Rob.* **5**, eaba5726 (2020).
- <sup>15</sup>X. Qi, S. Wang, S. Ma, K. Han, X. Bian, and X. Li, “Quantitative prediction of rolling dynamics of leukocyte-inspired microroller in blood flow,” *Phys. Fluids* **33**, 121908 (2021).
- <sup>16</sup>S. Muiños-Landin, A. Fischer, V. Holubec, and F. Cichos, “Reinforcement learning with artificial microswimmers,” *Sci. Rob.* **6**, eabd9285 (2021).
- <sup>17</sup>C. Mo, G. Li, and X. Bian, “Challenges and attempts to make intelligent microswimmers,” *Front. Phys.* **11**, 1279883 (2023).
- <sup>18</sup>S. Colabrese, K. Gustavsson, A. Celani, and L. Biferale, “Flow navigation by smart microswimmers via reinforcement learning,” *Phys. Rev. Lett.* **118**, 158004 (2017).
- <sup>19</sup>S. Colabrese, K. Gustavsson, A. Celani, and L. Biferale, “Smart inertial particles,” *Phys. Rev. Fluids* **3**, 084301 (2018).
- <sup>20</sup>L. Biferale, F. Bonaccorso, M. Buzzicotti, P. Clark Di Leoni, and K. Gustavsson, “Zermelo’s problem: Optimal point-to-point navigation in 2D turbulent flows using reinforcement learning,” *Chaos* **29**, 103138 (2019).
- <sup>21</sup>J. K. Alageshan, A. K. Verma, J. Bec, and R. Pandit, “Machine learning strategies for path-planning microswimmers in turbulent flows,” *Phys. Rev. E* **101**, 043110 (2020).
- <sup>22</sup>F. Borra, L. Biferale, M. Cencini, and A. Celani, “Reinforcement learning for pursuit and evasion of microswimmers at low Reynolds number,” *Phys. Rev. Fluids* **7**, 023103 (2022).
- <sup>23</sup>L. Amoudruz and P. Koumoutsakos, “Independent control and path planning of microswimmers with a uniform magnetic field,” *Adv. Intell. Syst.* **4**, 2100183 (2022).
- <sup>24</sup>P. Karnakov, L. Amoudruz, and P. Koumoutsakos, “Optimal navigation in microfluidics via the optimization of a discrete loss,” *Phys. Rev. Lett.* **134**, 044001 (2025).

- <sup>25</sup>Y. Liu, Z. Wang, and A. C. Tsang, "Reinforcement learning selects multimodal locomotion strategies for bioinspired microswimmers," *Soft Matter* **21**, 2363 (2025).
- <sup>26</sup>Y. Yang, M. A. Bevan, and B. Li, "Hierarchical planning with deep reinforcement learning for 3D navigation of microrobots in blood vessels," *Adv. Intell. Syst.* **4**, 2200168 (2022).
- <sup>27</sup>L. Amoudruz, A. Economides, and P. Koumoutsakos, "The volume of healthy red blood cells is optimal for advective oxygen transport in arterioles," *Biophys. J.* **123**, 1289 (2024).
- <sup>28</sup>P. Ghassemi, J. Wang, A. J. Melchiorri, J. C. Ramella-Roman, S. A. Mathews, J. C. Coburn, B. S. Sorg, Y. Chen, and T. J. Pfefer, "Rapid prototyping of biomimetic vascular phantoms for hyperspectral reflectance imaging," *J. Biomed. Opt.* **20**, 1 (2015).
- <sup>29</sup>L. Amoudruz, A. Economides, G. Arampatzis, and P. Koumoutsakos, "The stress-free state of human erythrocytes: Data-driven inference of a transferable RBC model," *Biophys. J.* **122**(8), 1517–1525 (2023).
- <sup>30</sup>R. D. Groot and P. B. Warren, "Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation," *J. Chem. Phys.* **107**, 4423 (1997).
- <sup>31</sup>G. Novati and P. Koumoutsakos, in *Proceedings of the 36th International Conference on Machine Learning* (PMLR, 2019).
- <sup>32</sup>A. Marco, F. Berkenkamp, P. Hennig, A. P. Schoellig, A. Krause, S. Schaal, and S. Trimpe, in *2017 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, 2017), pp. 1557–1563.
- <sup>33</sup>X. B. Peng, M. Andrychowicz, W. Zaremba, and P. Abbeel, in *2018 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, 2018), pp. 3803–3810.
- <sup>34</sup>A. K. Dasanna, J. Mauer, G. Gompper, and D. A. Fedosov, "Importance of viscosity contrast for the motion of erythrocytes in microcapillaries," *Front. Phys.* **9**, 666913 (2021).
- <sup>35</sup>F. Jülicher, "The morphology of vesicles of higher topological genus: Conformal degeneracy and conformal modes," *J. Phys. II* **6**, 1797 (1996).
- <sup>36</sup>G. Lim, H. W. M. Wortis, and R. Mukhopadhyay, "Red blood cell shapes and shape transformations: Newtonian mechanics of a composite membrane," *Soft Matter* **4**, 83–139 (2008).
- <sup>37</sup>D. A. Fedosov, "Multiscale modeling of blood flow and soft matter," Ph.D. thesis (Citeseer, 2010).
- <sup>38</sup>L. Amoudruz, "Simulations and control of artificial microswimmers in blood," Ph.D. thesis (ETH Zurich, 2022).
- <sup>39</sup>A. Yazdani, M. Deng, B. Caswell, and G. E. Karniadakis, "Flow in complex domains simulated by dissipative particle dynamics driven by geometry-specific body-forces," *J. Comput. Phys.* **305**, 906 (2016).
- <sup>40</sup>P. Karnakov, F. Wermelinger, S. Litvinov, and P. Koumoutsakos, in *Proceedings of the Platform for Advanced Scientific Computing Conference* (PASC, 2020), pp. 1–10.
- <sup>41</sup>D. Alexeev, L. Amoudruz, S. Litvinov, and P. Koumoutsakos, "Mirheo: High-performance mesoscale simulations for microfluidics," *Comput. Phys. Commun.* **254**, 107298 (2020).
- <sup>42</sup>R. Mhanna, F. Qiu, L. Zhang, Y. Ding, K. Sugihara, M. Zenobi-Wong, and B. J. Nelson, "Artificial bacterial flagella for remote-controlled targeted single-cell drug delivery," *Small* **10**, 1953 (2014).
- <sup>43</sup>D. Schamel, M. Pfeifer, J. G. Gibbs, B. Miksch, A. G. Mark, and P. Fischer, "Chiral colloidal molecules and observation of the propeller effect," *J. Am. Chem. Soc.* **135**, 12353 (2013).
- <sup>44</sup>P. J. Vach, N. Brun, M. Bennet, L. Bertinetti, M. Widdrat, J. Baumgartner, S. Klumpp, P. Fratzl, and D. Faire, "Selecting for function: Solution synthesis of magnetic nanopropellers," *Nano Lett.* **13**, 5373 (2013).
- <sup>45</sup>A. Y. Ng, D. Harada, and S. Russell, in *ICML* (Citeseer, 1999), Vol. 99, pp. 278–287.
- <sup>46</sup>G. Novati, L. Mahadevan, and P. Koumoutsakos, "Controlled gliding and perching through deep-reinforcement-learning," *Phys. Rev. Fluids* **4**, 093902 (2019).
- <sup>47</sup>S. Verma, G. Novati, and P. Koumoutsakos, "Efficient collective swimming by harnessing vortices through deep reinforcement learning," *Proc. Natl. Acad. Sci.* **115**, 201800923 (2018).
- <sup>48</sup>I. Mandralis, P. Weber, G. Novati, and P. Koumoutsakos, "Learning swimming escape patterns for larval fish under energy constraints," *Phys. Rev. Fluids* **6**, 093101 (2021).
- <sup>49</sup>S. M. Martin, D. Wälchli, G. Arampatzis, A. E. Economides, P. Karnakov, and P. Koumoutsakos, "Korali: Efficient and scalable software framework for Bayesian uncertainty quantification and stochastic optimization," *Comput. Methods Appl. Mech. Eng.* **389**, 114264 (2022).
- <sup>50</sup>A. C. Bakenecker, A. von Gladiss, T. Friedrich, U. Heinen, H. Lehr, K. Lüdtke-Buzug, and T. M. Buzug, "Actuation and visualization of a magnetically coated swimmer with magnetic particle imaging," *J. Magn. Magn. Mater.* **473**, 495 (2019).
- <sup>51</sup>I. S. Khalil, P. Ferreira, R. Eleutério, C. L. de Korte, and S. Misra, in *2014 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, 2014), pp. 3807–3812.
- <sup>52</sup>Q. Wang, Q. Wang, Z. Ning, K. F. Chan, J. Jiang, Y. Wang, L. Su, S. Jiang, B. Wang, B. Y. M. Ip *et al.*, "Tracking and navigation of a microswarm under laser speckle contrast imaging for targeted delivery," *Sci. Robot.* **9**, eadh1978 (2024).
- <sup>53</sup>S. Pané, J. Puigmartí-Luis, C. Bergeles, X.-Z. Chen, E. Pellicer, J. Sort, V. Počepcová, A. Ferreira, and B. J. Nelson, "Imaging Technologies for Biomedical Micro- and Nanoswimmers," *Adv. Mater. Technol.* **4**, 1800575 (2019).
- <sup>54</sup>G. Kokot and A. Snezhko, "Manipulation of emergent vortices in swarms of magnetic rollers," *Nat. Commun.* **9**, 2344 (2018).