

Emergence of odd elasticity in a microswimmer using deep reinforcement learning

Li-Shing Lin^{1,2}, Kento Yasuda², Kenta Ishimoto², and Shigeyuki Komura^{3,4,1,*}

¹Department of Chemistry, Graduate School of Science, Tokyo Metropolitan University, Tokyo 192-0397, Japan

²Research Institute for Mathematical Sciences, Kyoto University, Kyoto 606-8502, Japan

³Wenzhou Institute, University of Chinese Academy of Sciences, Wenzhou, Zhejiang 325001, China

⁴Oujiang Laboratory, Wenzhou, Zhejiang 325000, China



(Received 3 November 2023; accepted 5 June 2024; published 2 July 2024)

We use the Deep Q-Network with reinforcement learning to investigate the emergence of odd elasticity in an elastic microswimmer model. For an elastic microswimmer, it is challenging to obtain the optimized dynamics due to the intricate elasto-hydrodynamic interactions. However, our machine-trained model adopts a transition strategy (the waiting behavior) to optimize the locomotion. For the trained microswimmers, we evaluate the performance of the cycles by the product of the loop area (called *nonreciprocity*) and the loop frequency and show that the average swimming velocity is proportional to the performance. By calculating the force-displacement correlations, we obtain the effective odd elasticity of the microswimmer to characterize its nonreciprocal dynamics. This emergent odd elasticity is shown to be closely related to the loop frequency of the cyclic deformation. Our work demonstrates the utility of machine learning in achieving optimal dynamics for elastic microswimmers and introduces postanalysis methods to extract crucial physical quantities such as nonreciprocity and odd elasticity.

DOI: [10.1103/PhysRevResearch.6.033016](https://doi.org/10.1103/PhysRevResearch.6.033016)

I. INTRODUCTION

Active systems composed of self-driven units play a crucial role in biological processes as they are able to convert microscopic energy into macroscopic work [1–3]. To achieve sustainable work at the microscopic scale, active units must go through nonreciprocal cyclic motions [4,5]. For example, cyclic state transitions of enzymatic molecules are driven by catalytic chemical reactions [6,7], which can be described by simple coarse-grained models [8–10]. To evaluate the functionality of an enzyme, we defined a physical quantity called *nonreciprocity* that represents the area enclosed by a trajectory in the conformational space [11,12]. According to Purcell’s scallop theorem for microswimmers moving in a viscous fluid [13–15], the average swimming velocity is proportional to the nonreciprocity and the loop frequency of the cyclic body motion. It was also reported that the crawling speed of a cell on a substrate is determined by the nonreciprocity [16–18].

Recently, Scheibner *et al.* introduced the concept of odd elasticity which is useful for characterizing nonequilibrium active systems [19,20]. Odd elasticity, arising from anti-symmetric (odd) components of the elastic modulus tensor that violate the energy conservation law, can exist in active

materials [21–24], biological systems [25], and active robots [26–28]. We emphasize that the concept of odd elasticity is not limited to elastic materials but can be extended to various dynamical systems [29,30]. An illustrative example is that a microswimmer with odd elasticity can exhibit directional locomotion in the presence of thermal agitation [31]. In fact, the average velocity of an odd microswimmer is proportional to the odd elasticity. On the other hand, in the model of a stochastic enzyme, we have quantified the average work per cycle in terms of effective odd elasticity [12]. Notably, odd elasticity serves as a useful measure for characterizing nonequilibrium micromachines such as proteins, enzymes, microswimmers, and robots, regardless of their specific functions.

Despite the importance of odd elasticity in active systems, its physical origin still needs to be better understood [32]. One possibility is to use Onsager’s variational principle [33] to derive dynamical equations for an active system with odd elasticity [34]. The obtained nonreciprocal equations [35] manifest the physical origin of the odd elastic constant that is proportional to the nonequilibrium driving force [34]. On the other hand, odd elasticity may not be innate to micromachines but can be an ability acquired after many experiences and training processes. In this work, considering an elastic three-sphere microswimmer model [36–38], we utilize machine learning techniques to account for the emergence of odd elastic relation between its elastic components. With this approach, a microswimmer can automatically obtain the most efficient swimming strategy without prescribing any deformation dynamics.

In recent years, machine learning has been widely applied to active systems as a powerful tool to unravel the complexities of biological systems [39–42]. Notably, the application

*Contact author: komura@wiucas.ac.cn

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

of reinforcement learning techniques is capable and versatile in training various microswimmers. These methods have been used to navigate them through complex and dynamic environments with remarkable adaptability, such as path-planning in turbulent flows or noisy surroundings [43–46]. Furthermore, machine learning has been applied to optimize the local motion and intricate navigation of microswimmers with complex structures or higher degrees of freedom [47–49]. These approaches have been further extended to more difficult tasks, such as cooperative swimming and predation models [50–52].

The main aim of this article is to reveal the emergence of odd elasticity in an elastic microswimmer model by using the Deep Q-Network with reinforcement learning. Traditional kinematic models, such as rigidly connected swimmers, select paths in the deformation space (gait switching) assuming that the applied forces can instantaneously adapt to any prescribed motion [47,53,54]. However, the deformation of an elastic microswimmer cannot be prescribed [36]. Determining these dynamics requires a nuanced understanding of the elasto-hydrodynamic process, where traditional methods for achieving optimal control face significant challenges. These challenges are evident in the previous study [36], in which the swimming velocity of an elastic microswimmer decreases in its large-frequency regime when prescribed dynamics are assumed.

Unlike the prescribed dynamics model, our machine-learning approach successfully develops an optimal control strategy, adapting a transition (emergence of the waiting behavior) to avoid the velocity decrease. We note that such elasto-hydrodynamic systems, different from the study in Tsang *et al.* [47], usually require continuous state and action spaces to tackle. The discovered strategy transition using the waiting behavior emerges from the fluid-structure interactions in which distinct hydrodynamic modes with different time scales play important roles.

From the obtained numerical data, we quantify the performance and effective odd elasticity of the trained microswimmer. The estimated cycle performance, which is the product of the nonreciprocity (loop area) and the loop frequency, coincides with the swimming velocity by using a proper scale factor. We also demonstrate that the emergent odd elasticity of the microswimmer is closely related to the loop frequency of the cyclic deformation. The present work demonstrates the utility of machine learning in revealing various nonreciprocal phenomena in active systems.

In Sec. II, we review the model of an elastic microswimmer with prescribed dynamics [36]. In Sec. III, we explain the deep reinforcement learning technique to train the elastic microswimmer. In Sec. IV, we describe the physical properties of the fully trained microswimmer. In particular, we shall discuss the emergence of limit cycles, cycle performance, average velocity, and effective odd elasticity. In Sec. V, we briefly mention the training progression of an elastic microswimmer. A summary of our work and some discussion are given in Sec. VI.

II. ELASTIC MICROSWIMMER WITH PRESCRIBED DYNAMICS

We first review the elastic three-sphere microswimmer model introduced by the present authors [36,55–57] and

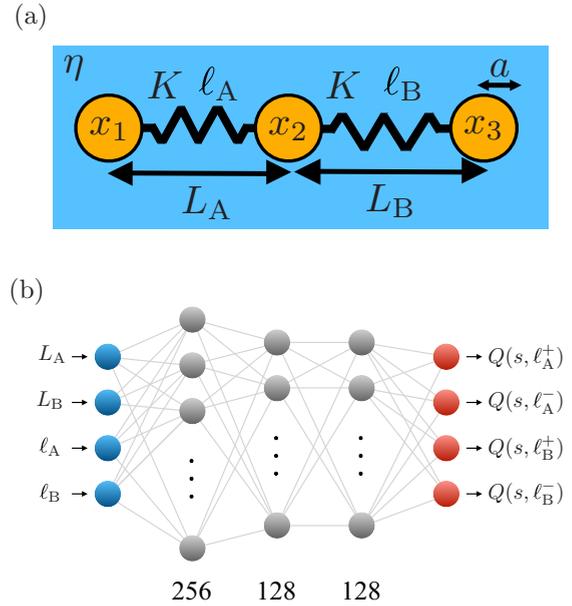


FIG. 1. (a) An elastic microswimmer consists of three spheres of radius a positioned along a one-dimensional coordinate system, denoted by x_i ($i = 1, 2, 3$). The three spheres are connected by two harmonic springs with the elastic constant K and the time-dependent natural lengths ℓ_α ($\alpha = A, B$). The spring extensions are denoted by $L_A = x_2 - x_1$ and $L_B = x_3 - x_2$. The microswimmer is immersed in a viscous fluid with the shear viscosity η and $\tau = 6\pi\eta a/K$ gives the hydrodynamic relaxation time. (b) Schematic of the neural network architecture that predicts Q values for each action from an input observation of L_α and ℓ_α . The blue and red columns represent the input and output layers, respectively. The gray columns are the three linear hidden layers with dimensions 256, 128, and 128.

others [37,38,58]. As illustrated in Fig. 1(a), the microswimmer consists of three spheres of radius a positioned along a one-dimensional coordinate system, denoted by x_i ($i = 1, 2, 3$). One can assume $x_1 < x_2 < x_3$ without loss of generality. Unlike the original three-sphere microswimmer model by Najafi and Golestanian [53,54], the three spheres are connected by two harmonic springs, each with a time-dependent natural length ℓ_α ($\alpha = A, B$) and having the same spring elastic constant K . Such an *elastic microswimmer* is immersed in a viscous fluid with the shear viscosity η . Although a and K can differ between the spheres and the springs, respectively [36,54,59], we consider here the symmetric case. This elastic microswimmer model [36] reduces to the original three-sphere model with rigid arms [53,54] when K is infinitely large.

When the spring lengths $L_A = x_2 - x_1$ and $L_B = x_3 - x_2$ deviate from their natural lengths ℓ_α , the elastic forces f_i acting on each sphere are given by

$$f_1 = K(L_A - \ell_A), \quad (1)$$

$$f_2 = -K(L_A - \ell_A) + K(L_B - \ell_B), \quad (2)$$

$$f_3 = -K(L_B - \ell_B). \quad (3)$$

Due to hydrodynamic interactions described by the Stokes mobility and the Oseen tensor, the forces f_i and the sphere

velocities $v_i = \dot{x}_i = dx_i/dt$ (dot indicates the time derivative) are related by [53,54]

$$v_1 = \frac{f_1}{6\pi\eta a} + \frac{f_2}{4\pi\eta L_A} + \frac{f_3}{4\pi\eta(L_A + L_B)}, \quad (4)$$

$$v_2 = \frac{f_1}{4\pi\eta L_A} + \frac{f_2}{6\pi\eta a} + \frac{f_3}{4\pi\eta L_B}, \quad (5)$$

$$v_3 = \frac{f_1}{4\pi\eta(L_A + L_B)} + \frac{f_2}{4\pi\eta L_B} + \frac{f_3}{6\pi\eta a}, \quad (6)$$

where the conditions $a/L_\alpha \ll 1$ are assumed. For the elastic microswimmer, the force-free condition, $f_1 + f_2 + f_3 = 0$, is automatically satisfied, ensuring a self-propelled motion without any external force.

When L_α and ℓ_α are given, the dynamics of the microswimmer are deterministic and the total swimming velocity $V = (v_1 + v_2 + v_3)/3$ is given by

$$V = \frac{1}{12\pi\eta} \left[\left(\frac{1}{L_A + L_B} - \frac{1}{L_B} \right) f_1 + \left(\frac{1}{L_A + L_B} - \frac{1}{L_A} \right) f_3 \right]. \quad (7)$$

For relatively small deformations of the springs, we can define the small displacements with respect to the average spring length ℓ as $u_\alpha = L_\alpha - \ell$ ($\alpha = A, B$). Within the small-amplitude approximation, $u_\alpha/\ell \ll 1$, Golestanian and Ajdari calculated the average swimming velocity of a three-sphere microswimmer up to the leading order as [54]

$$\bar{V} = \frac{7a}{24\ell^2} \overline{(u_A \dot{u}_B - \dot{u}_A u_B)}. \quad (8)$$

Here the averaging, indicated by the bar, is performed by time integration in a full cycle and further divided by the total time of a period. The above expression indicates that \bar{V} is determined by the product of the closed loop area and the loop frequency [60].

The explicit form of \bar{V} of an elastic microswimmer can be obtained by specifying a prescribed cyclic change in the natural spring lengths ℓ_α . Previously, we used the following sinusoidal forms [36,57]:

$$\ell_A = \ell + d_A \cos(\Omega t), \quad (9)$$

$$\ell_B = \ell + d_B \cos(\Omega t - \phi), \quad (10)$$

where ℓ is the constant natural length, d_α are the amplitudes of the oscillatory change, Ω is the common frequency, and ϕ is the phase difference between the two cyclic changes. When the natural lengths undergo this prescribed cycle, the spring lengths relax to their new natural lengths obeying Eqs. (1)–(6) with a hydrodynamic relaxation time $\tau = 6\pi\eta a/K$.

Then the average swimming velocity of an elastic microswimmer with the prescribed dynamics was calculated to be [36,57]

$$\bar{V} = \frac{7ad_A d_B}{24\ell^2 \tau} G(\hat{\Omega}) \sin \phi, \quad (11)$$

where $\hat{\Omega} = \Omega\tau$ is the dimensionless frequency and $G(\hat{\Omega})$ is the scaling function

$$G(\hat{\Omega}) = \frac{3\hat{\Omega}(3 + \hat{\Omega}^2)}{9 + 10\hat{\Omega}^2 + \hat{\Omega}^4}. \quad (12)$$

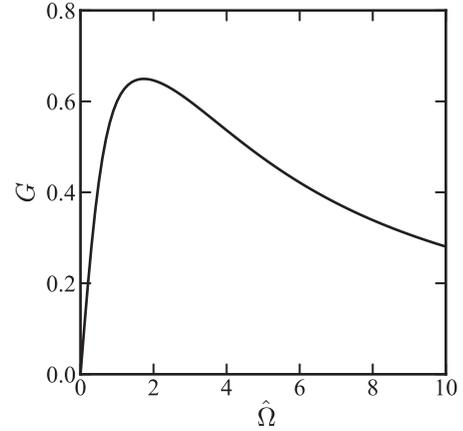


FIG. 2. The scaling function $G(\hat{\Omega})$ in Eq. (12) that describes the frequency dependence of the average swimming velocity \bar{V} of an elastic swimmer with prescribed dynamics of the natural lengths given by Eqs. (9) and (10).

We see that \bar{V} is nonzero when $\phi \neq 0$, corresponding to the nonreciprocal deformation, and $|\bar{V}|$ is maximized when $\phi = \pm\pi/2$. Notably, such nonreciprocal deformation can be effectively generated by assuming an antisymmetric stress-strain relation among the two springs [30,31]. This antisymmetric cross-correlation represents what we refer to as odd elasticity in this article.

In Fig. 2, we plot $G(\hat{\Omega})$ using Eq. (12). In the small-frequency limit, where $\hat{\Omega} \ll 1$, the average velocity increases as $\bar{V} \sim \hat{\Omega}$ [53,54]. However, in the larger-frequency limit, where $\hat{\Omega} \gg 1$, the average velocity decreases as $\bar{V} \sim \hat{\Omega}^{-1}$ [36,57]. The crossover frequency between these two regimes is approximately $\hat{\Omega} \approx 1$. In the large-frequency regime, the mechanical response is delayed because it takes time for the springs to relax to their natural lengths. Such a decrease in the swimming velocity is a drawback of elastic microswimmers with prescribed motion. A similar crossover behavior and a decrease in the average velocity were also predicted for the Najafi-Golestanian microswimmer model in a viscoelastic fluid [61,62].

III. ELASTIC MICROSWIMMER DIRECTED BY REINFORCEMENT LEARNING

Using an elastic microswimmer model, we apply a machine learning method to direct its movement rather than prescribing the motion of the natural lengths ℓ_α . We combine the Deep Q-Network (DQN) with reinforcement learning [63–65] to train the actuation of a microswimmer and obtain the optimized dynamics for the natural lengths. In particular, we shall investigate how a trained microswimmer adapts to a new strategy to avoid the decrease in the average swimming velocity when the actuation is faster than the hydrodynamic relaxation.

As shown in Fig. 1(b), our artificial intelligence (AI) uses the spring lengths L_α and the natural lengths ℓ_α as an observation (state) input and performs an action to change either ℓ_A or ℓ_B with an *actuation velocity* U , i.e., the rate of changes in the natural lengths. Specifically, the output action space is discrete

and consists of four actions ℓ_α^\pm corresponding to the increase and decrease of ℓ_α . When the natural lengths ℓ_α change, the spring lengths L_α tend to relax toward the new natural lengths according to Eqs. (1)–(6). Hereinafter, we choose the sphere radius a and the hydrodynamic relaxation time $\tau = 6\pi\eta a/K$ as the units for length and time, respectively. The dimensionless quantities are then denoted with a hat such as $\hat{L}_\alpha = L_\alpha/a$, $\hat{\ell}_\alpha = \ell_\alpha/a$, and $\hat{U} = U\tau/a$.

During the training process, we constrain the natural lengths in the range $8 \leq \hat{\ell}_\alpha \leq 12$ to ensure the conditions $a/L_\alpha \ll 1$ and $(L_\alpha - \ell)/\ell \ll 1$. This assignment simplifies the model and aligns with established methods, enabling valid comparisons with existing studies [31,53,54,57]. In our model, the actuation velocity \hat{U} is a control parameter and can be contrasted with the frequency $\hat{\Omega}$ in Sec. II. We train the microswimmers at different actuation velocities ranging from $\hat{U} = 0.1$ to 10, with an increment of 0.1. Each training session comprises 200 episodes, and each episode contains 1,200 decisions made by the AI.

To optimize the locomotion of the microswimmer, we initiate each episode with a random state and employ the epsilon greedy algorithm (EGA) to further balance between exploration and exploitation during training process [65]. Actions are determined successively by the AI in each decision step in which the natural lengths $\hat{\ell}_\alpha$ are changed by ± 1 with a given actuation velocity \hat{U} . Hence, consecutive decisions are made at intervals of $\Delta\hat{t}_{\text{dec}} = 1/\hat{U}$. Since the numerical time step $\Delta\hat{t}_{\text{num}} = 0.01$ is used to solve the hydrodynamic equations in Eqs. (1)–(6), the relation $\Delta\hat{t}_{\text{dec}}/\Delta\hat{t}_{\text{num}} = 100/\hat{U}$ ensures adequate numerical time steps between successive actions for $0.1 \leq \hat{U} \leq 10$.

In our model, successive actions are consistently taken at intervals of $\Delta\hat{t}_{\text{dec}}$. This means that the subsequent action is executed immediately after $\hat{\ell}_\alpha$ has evolved by ± 1 , irrespective of whether the spring lengths \hat{L}_α have fully relaxed to the new natural lengths ℓ_α or not. In specific situations, the machine can conduct a waiting strategy where it refrains from changing any natural lengths for $\Delta\hat{t}_{\text{dec}}$. Such a decision arises from actions that violate the natural length constraints. For example, taking $\hat{\ell}_A^+$ is not allowed when $\hat{\ell}_A = 12$. In this case, no change is applied to $\hat{\ell}_A$, and the consecutive action is conducted after $\Delta\hat{t}_{\text{dec}}$.

The training process is formulated as a Markov decision process (MDP) with the memoryless property [65], ensuring that the immediate reward depends only on the current state and action. Within the MDP method, the training of the DQN is guided by Bellman's equation that is used to iteratively update the prediction of the Q -value function (the expected cumulative future reward) in every decision step [63–65]. In our model, Bellman's equation is given by

$$Q(s_t, a_t) = r_t + \gamma \max_{a'} Q(s_{t+1}, a'), \quad (13)$$

where $Q(s_t, a_t)$ represents the Q value of taking an action $a_t \in \hat{\ell}_\alpha^\pm$ in state $s_t = (\hat{L}_\alpha, \hat{\ell}_\alpha)$ at time t and r_t denotes the reward obtained after taking action a_t in state s_t . The reward is defined as the positive displacement of the whole microswimmer during a decision step. The coefficient γ represents the discount factor that balances the importance between immediate rewards and future rewards. We choose a conventional value of

$\gamma = 0.99$ for farsightedness to prioritize long-term cumulative rewards [47–49,52]. By updating the network with Bellman's equation in every decision step, our DQN efficiently learns the optimal policy for controlling the dynamics of the natural lengths $\hat{\ell}_\alpha$.

In Fig. 3(a), we show the training curve of a microswimmer when the actuation velocity $\hat{U} = 2$. As a function of the trained episodes, we plot the total distance, namely the net displacement from the initial position that the microswimmer can achieve within each episode. During the training process, the total distance continuously increases over the episodes, indicating enhanced locomotion ability. After about 100 episodes of training, the microswimmer's performance approaches an optimal swimming distance within each single episode. The small fluctuations that remain after reaching the optimal total distance are due to random initial conditions and EGA used for the training. Details of the training progression will be discussed separately in Sec. V.

The characteristics of the training curve vary for swimmers with different values of \hat{U} . As \hat{U} increases, swimmers require more training episodes to achieve optimized swimming velocity. This trend becomes pronounced at higher values of \hat{U} , which are not included in this article due to the lack of additional physical interest (e.g., $\hat{U} \sim 20$ –30). On the other hand, the standard deviation of the training datasets shows increased sensitivity within the range $\hat{U} \in [0.1, 10]$. This increase in standard deviation is due to the complexity arising from the delayed response of the exact length L_α . Such a delay in response leads to a behavior transition, which will be further discussed in Sec. IV.

IV. FULLY TRAINED ELASTIC MICROSWIMMER

In this section, we discuss the dynamic properties of the microswimmers which have been trained for 200 episodes. Besides the emergence of limit cycles, we shall discuss the performance and effective odd elasticity of trained microswimmers.

A. Emergence of limit cycles

In Figs. 3(b) and 3(c), we present the cyclic motions of the fully trained swimmer when $\hat{U} = 2$. The black lines represent the deviation of the natural lengths from its average value, i.e., $\hat{d}_\alpha = \hat{\ell}_\alpha - \hat{\ell}$ ($\alpha = A, B$), where we chose $\hat{\ell} = 10$ because of the constraint $8 \leq \hat{\ell}_\alpha \leq 12$. The red lines represent the spring extensions $\hat{u}_\alpha = \hat{L}_\alpha - \hat{\ell}$ ($\alpha = A, B$). Both \hat{u}_A and \hat{u}_B exhibit periodic motions with a phase difference of approximately $\pi/2$, which corresponds to the maximum efficiency for swimming.

In Fig. 3(d), we present the configuration space trajectory of the same trained microswimmer over one cycle. The trajectory of the natural lengths \hat{d}_α (shown in black) forms a counterclockwise square in the range $-2 \leq \hat{d}_\alpha \leq 2$. The spring extensions \hat{u}_α (shown in red) also exhibit a limit cycle in the configuration space. This indicates that the fully trained elastic microswimmer has acquired steady nonreciprocal spring motion after training.

To investigate the dependence on the actuation velocity \hat{U} , we plot the configuration space trajectories for different \hat{U}

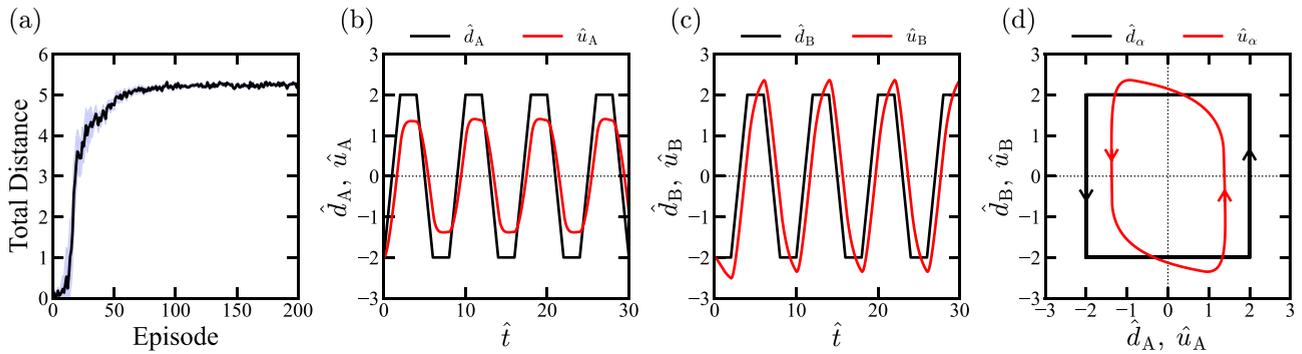


FIG. 3. (a) Training curve of an elastic microswimmer with an actuation velocity of $\hat{U} = 2$. The black line represents the total distance achieved by the microswimmer in each training episode for a single case. The blue shaded area indicates the standard deviation around the mean value obtained from 100 training instances of microswimmers with the same actuation velocity. (b), (c) Periodic dynamics of the natural length deviations $\hat{d}_\alpha = \hat{\ell}_\alpha - \hat{\ell}$ (black) and the spring extensions $\hat{u}_\alpha = \hat{L}_\alpha - \hat{\ell}$ (red) obtained from the fully trained microswimmer with 200 episodes of experience. The dimensionless time is defined by $\hat{t} = t/\tau$, and the parameters are $\hat{U} = 2$ and $\hat{\ell} = 10$. The oscillation phase difference between (b) and (c) is approximately $\pi/2$, with a fore-aft amplitude difference. (d) The configuration space trajectory of \hat{d}_α (black) and \hat{u}_α (red) when $\hat{U} = 2$. Both \hat{d}_α and \hat{u}_α form counterclockwise cyclic loops.

values ranging from 0.1 to 10 in Figs. 4(a)–4(h) [Figs. 4(d) and 3(d) are the same]. In all the cases, the natural lengths \hat{d}_α (shown in black) follow the same counterclockwise square shape. When the actuation velocity is small, such as $\hat{U} \leq 2$ in Figs. 4(a)–4(d), the hydrodynamic relaxation process can catch up with the change in the natural lengths, and the cycles of the spring extensions \hat{u}_α (shown in red) are close to the square-shaped trajectory. In this regime, the enclosed area

within the loop decreases when \hat{U} is increased, as we quantify in the next subsection.

For larger actuation velocity $\hat{U} > 2$, corresponding to Figs. 4(e)–4(h), the hydrodynamic relaxation becomes the slower mode. This situation is similar to the case of $\hat{\Omega} > 1$ in Sec. II. For the trained microswimmer, however, the AI adapts a waiting strategy once the natural lengths reach the maximum or minimum values ($\hat{d}_\alpha = \pm 2$) for the spring extensions \hat{u}_α

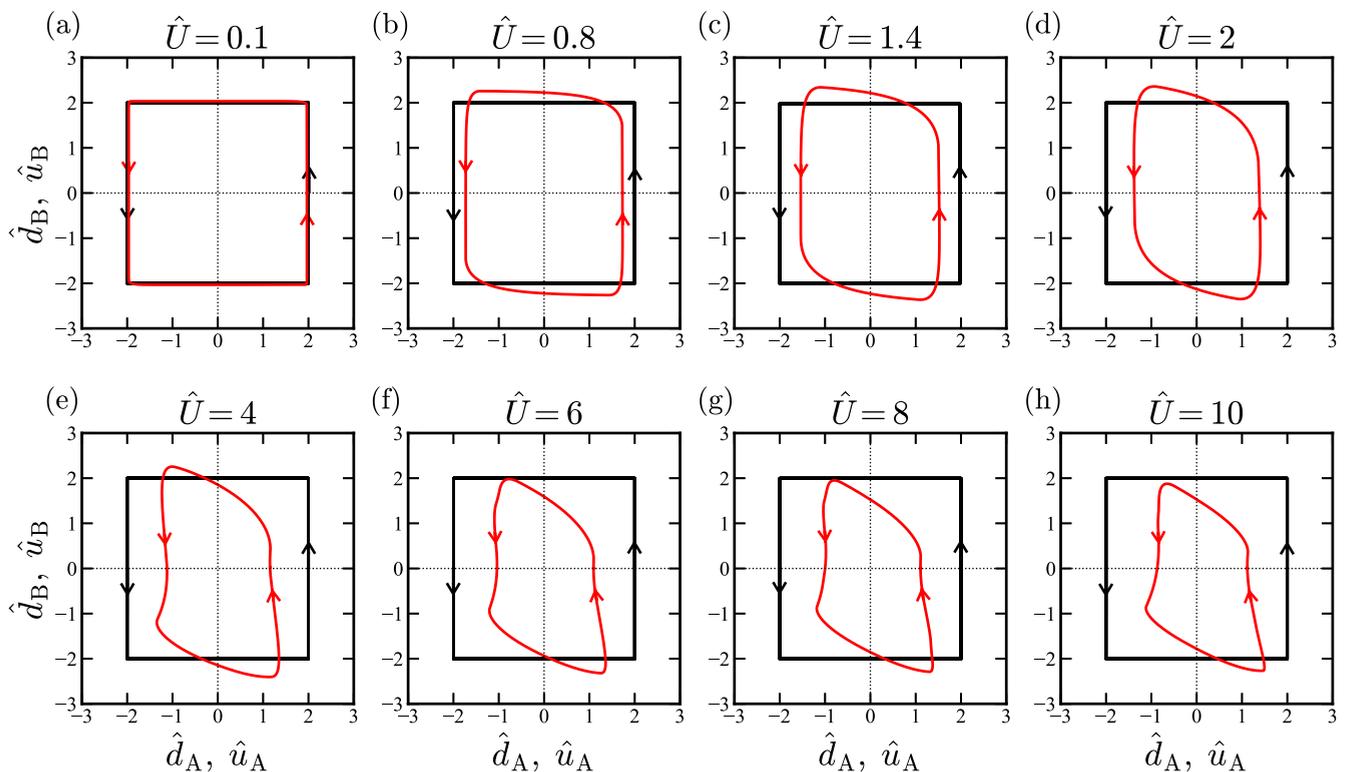


FIG. 4. Limit cycles of microswimmers trained with different \hat{U} values ranging from 0.1 to 10. Black and red lines represent the natural length deviations \hat{d}_α and the spring extensions \hat{u}_α , respectively. The plot in (d) for $\hat{U} = 2$ is identical to that in Fig. 3(d).

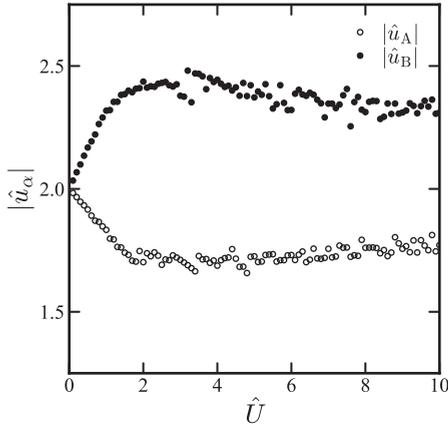


FIG. 5. The dimensionless spring amplitudes $|\hat{u}_\alpha|$ as a function of \hat{U} . The amplitudes exhibit fore-aft asymmetry.

to relax sufficiently. This waiting strategy allows the spring lengths to catch up with the large actuation velocity \hat{U} and to prevent the enclosed area from further decreasing. Since the dynamics are no longer dominated by the actuation velocity \hat{U} , the trajectories of \hat{u}_α become less squared and less symmetric for $\hat{U} > 2$. As a result, the distinctions between different loop shapes become less pronounced in Figs. 4(e)–4(h).

Another notable result in Fig. 4 is the fore-aft amplitude asymmetry between \hat{u}_A and \hat{u}_B . In Fig. 5, we plot the amplitudes $|\hat{u}_\alpha|$ as a function of \hat{U} . Starting from the value $|\hat{u}_\alpha| = 2$, $|\hat{u}_A|$ and $|\hat{u}_B|$ decreases and increases, respectively, as \hat{U} is increased. Notice that the maximum amplitude can be up to $|\hat{u}_B|/\hat{\ell} \approx 0.25$. When $\hat{U} > 2$, however, $|\hat{u}_\alpha|$ are almost independent of \hat{U} . This asymmetry in fore-aft amplitude is due to the nonreciprocal swimming cycle, which is associated with the asymmetric order in which ℓ_α changes.

B. Cycle performance and swimming velocity

Next we discuss the performance of the acquired cyclic motion and the average swimming velocity of the fully trained microswimmers. Following our work on catalytic enzymes

[11,12], we consider the following quantity called *nonreciprocity*:

$$R = \frac{1}{2} \int_0^T dt (u_A \dot{u}_B - \dot{u}_A u_B), \quad (14)$$

where T is the period of one cycle. We note again that R represents the area enclosed by the loop trajectory in the configuration space [60]. Then the dimensionless average velocity $\hat{V} = \bar{v}\tau/a$ obtained from Eq. (8) can be rewritten in terms of the dimensionless nonreciprocity $\hat{R} = R/a^2$ as

$$\hat{V} = \frac{7}{12\hat{\ell}^2} \frac{\hat{R}}{\hat{T}}, \quad (15)$$

where $\hat{T} = T/\tau$. We shall call $1/\hat{T}$ the dimensionless *loop frequency*. The quantity \hat{R}/\hat{T} , representing the area enclosed by the loop per unit time, thus determines the swimming velocity when the deformation is small.

In Fig. 6(a), we plot the nonreciprocity \hat{R} as a function of \hat{U} for fully trained microswimmers. As indicated in Fig. 4, the enclosed area \hat{R} decreases as \hat{U} increases. In Fig. 6(b), the loop frequency $1/\hat{T}$ is plotted as a function of \hat{U} . When $\hat{U} \leq 2$, the loop frequency increases linearly with \hat{U} (shown by the dotted line) to minimize the loop period. For $\hat{U} > 2$, however, the dependence of $1/\hat{T}$ on \hat{U} deviates from the linear relation. This is because the actuation velocity \hat{U} outpaces the hydrodynamic relaxation rate, and the AI adapts the waiting strategy to adjust to the slow hydrodynamic mode. The slope of the dashed line is $1/16$ ($1/\hat{T} = \hat{U}/16$), and hence the dimensionless waiting time at $\dot{d}_\alpha = \pm 2$ is estimated by $\hat{T} - 16/\hat{U}$. The finite waiting time appears as a transition at $\hat{U} \approx 2$. As plotted in Fig. 6(c), we find that the cycle performance, as measured by \hat{R}/\hat{T} , increases monotonically with \hat{U} and eventually approaches the value $\hat{R}/\hat{T} \approx 1.75$, bounded by the hydrodynamic relaxation process.

To check the validity of Eq. (15), we plot in Fig. 6(d) the average swimming velocity \hat{V} obtained from the actual displacement of the microswimmers. Both \hat{R}/\hat{T} and \hat{V} in Figs. 6(c) and 6(d), respectively, show almost the same dependence on \hat{U} except for a scaling factor. The obtained

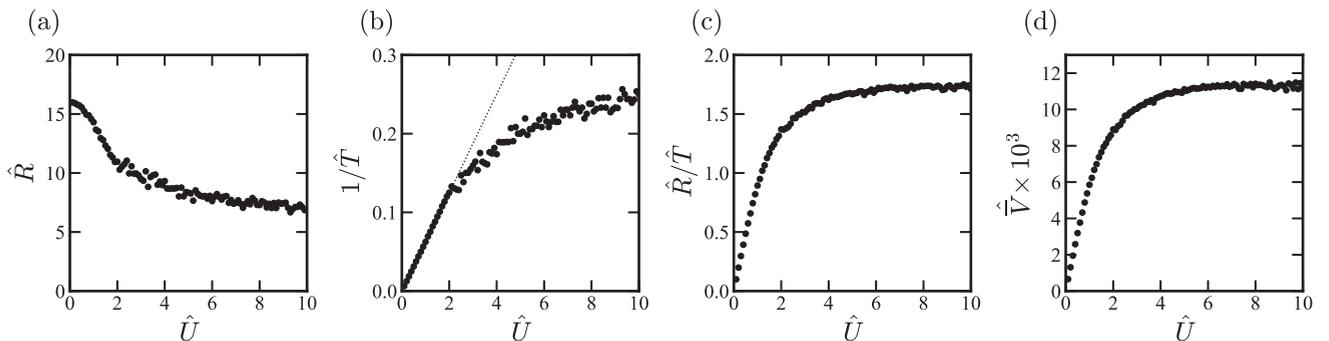


FIG. 6. Performance of the fully trained microswimmer at various actuation velocities \hat{U} . (a) The dimensionless nonreciprocity $\hat{R} = R/a^2$ defined in Eq. (14) as a function of \hat{U} . This quantity represents the enclosed area of the red loops made of u_α in Fig. 4. (b) The dimensionless loop frequency $1/\hat{T} = \tau/T$ as a function of \hat{U} . When $\hat{U} \leq 2$, the loop frequency increases linearly with \hat{U} as fitted by the dotted line ($1/\hat{T} = \hat{U}/16$). (c) The dimensionless cycle performance \hat{R}/\hat{T} [calculated from (a) and (b)] as a function of \hat{U} . (d) The dimensionless average swimming velocity $\hat{V} = \bar{v}\tau/a$ as a function of \hat{U} . Both \hat{R}/\hat{T} in (c) and \hat{V} in (d) show the same \hat{U} dependence except for a scaling factor. The ratio between these quantities is $\hat{V}\hat{T}/\hat{R} \approx 6.57 \times 10^{-3}$. The interpretation of this ratio is discussed in the text.

ratio from Figs. 6(c) and 6(d) is $\hat{V}\hat{T}/\hat{R} \approx 6.57 \times 10^{-3}$ within the studied \hat{U} range. This value can be compared with the geometrical prefactor $7/(12\hat{\ell}^2) \approx 5.83 \times 10^{-3}$ in Eq. (15) when $\hat{\ell} = 10$. A small difference between these prefactors comes from the assumption $u_\alpha/\ell \ll 1$ used in Eq. (8) or Eq. (15). As previously shown in Fig. 5, the spring deformations for our trained microswimmers can reach $\hat{u}_\alpha/\hat{\ell} \approx 0.2$ or larger. Higher-order contributions need to be incorporated into Eq. (15) to reproduce the average swimming velocity achieved by our swimmer model.

The behavior of \hat{V} in Fig. 6(d) for the trained microswimmer is in sharp contrast to that of an elastic microswimmer whose natural spring motions are prescribed. In Sec. II, we showed in Eq. (11) and Fig. 2 that the average velocity decreases when $\hat{\Omega} > 1$ [36,57]. For the fully trained microswimmer, however, an emergent waiting strategy appears when $\hat{U} > 2$, such that the swimming velocity does not decrease even at higher actuation velocities \hat{U} . This discovery of a strategy transition exemplifies how machine learning can reveal the intricate optimal dynamics of complex microswimmers.

C. Effective even and odd elasticities

Finally, we explain the method to extract the effective even and odd elasticities of the trained microswimmer from the numerical data. A direct way is to assume the following odd-elastic Hookean relations between the forces and the spring extensions [12,30,35]:

$$\begin{pmatrix} F_A \\ F_B \end{pmatrix} = \begin{pmatrix} k^e & k' + k^o \\ k' - k^o & k^e \end{pmatrix} \begin{pmatrix} u_A \\ u_B \end{pmatrix}. \quad (16)$$

In the above expression, the forces F_α are given by Stokes's law $F_\alpha = -6\pi\eta a\dot{u}_\alpha$ and $u_\alpha = L_\alpha - \ell$ as before. In the elastic matrix, k^e and k' represent the diagonal and off-diagonal even elasticities, respectively, while k^o represents the effective odd elasticity. The diagonal even elasticity k^e should be distinguished from the spring elastic constant K in the elastic microswimmer model.

Typically, a finite k^e results in damping oscillation. Given the fact that our data demonstrate sustained oscillations without amplitude decay, we assume in the following that $k^e = 0$. On the other hand, the off-diagonal even elasticity k' characterizes the fore-aft amplitude asymmetry as described in Fig. 5, and the odd elasticity k^o should quantify the nonreciprocal dynamics of the microswimmer [12,30].

To obtain k' and k^o from the numerical data, we calculate the cross-correlations $\overline{F_A u_B}$ and $\overline{F_B u_A}$ averaged over one period of the deformation cycle. This is possible since the trained dynamics for the two springs are periodic with the same frequency and a finite phase difference. Since these correlations are given by

$$\overline{F_A u_B} = (k' + k^o)\overline{u_B^2}, \quad (17)$$

$$\overline{F_B u_A} = (k' - k^o)\overline{u_A^2}, \quad (18)$$

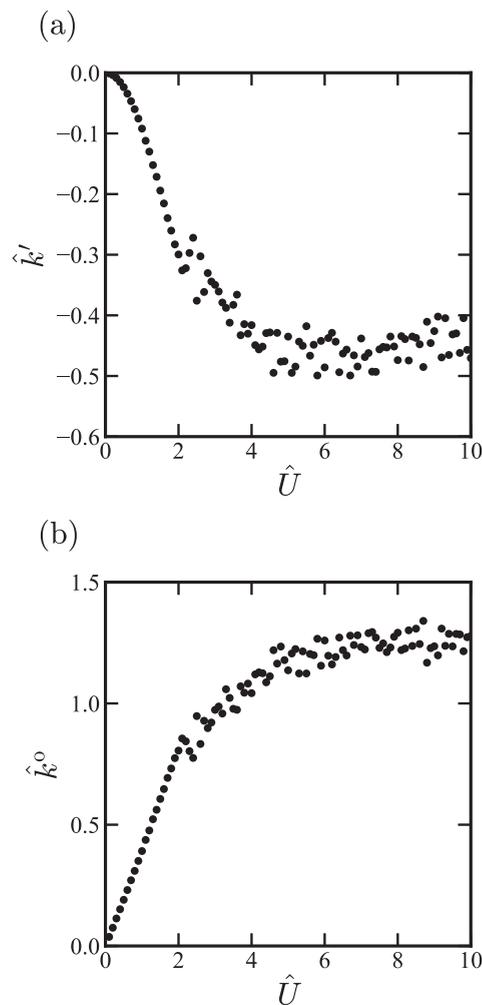


FIG. 7. (a) The effective off-diagonal even elasticity $\hat{k}' = k'/K$ for fully trained microswimmers as a function of \hat{U} [see Eq. (19)]. The \hat{U} dependence is similar to that of the nonreciprocal \hat{R} in Fig. 6(a). (b) The effective odd elasticity $\hat{k}^o = k^o/K$ as a function of \hat{U} [see Eq. (20)]. The \hat{U} dependence closely resembles that of the loop frequency $1/\hat{T}$ in Fig. 6(b).

the effective (off-diagonal) even and odd elastic coefficients can be obtained from

$$k' = \frac{1}{2} \left(\frac{\overline{F_A u_B}}{\overline{u_B^2}} + \frac{\overline{F_B u_A}}{\overline{u_A^2}} \right), \quad (19)$$

$$k^o = \frac{1}{2} \left(\frac{\overline{F_A u_B}}{\overline{u_B^2}} - \frac{\overline{F_B u_A}}{\overline{u_A^2}} \right), \quad (20)$$

where all the related correlations can be directly calculated from the numerical data.

In Figs. 7(a) and 7(b), we plot dimensionless even elasticity, $\hat{k}' = k'/K$ and odd elasticity, $\hat{k}^o = k^o/K$, respectively, as functions of \hat{U} . Both \hat{k}' and \hat{k}^o vanish when $\hat{U} = 0$. We recognize that the behavior of \hat{k}' is similar to that of \hat{R} in Fig. 6(a) (except a constant shift), while the data of \hat{k}^o closely resemble those of $1/\hat{T}$ in Fig. 6(b). These results indicate that \hat{k}' characterizes the extent of amplitude asymmetry that reduces the enclosed area, whereas \hat{k}^o directly corresponds to the loop frequency. Since $1/\hat{T} \propto \hat{k}^o$, we confirm that the

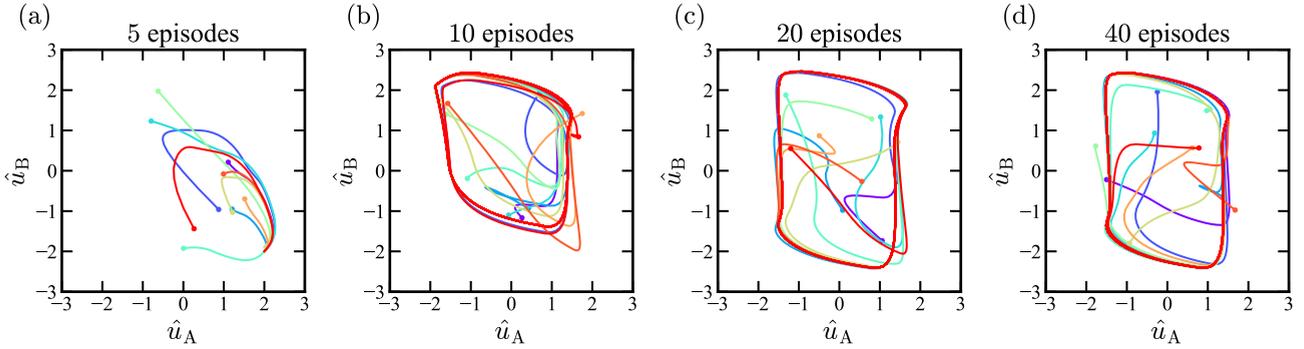


FIG. 8. The configuration space trajectories at different training stages; (a) 5, (b) 10, (c) 20, and (d) 40 episodes for a microswimmer with $\hat{U} = 2$. Each trajectory is initiated from a distinct initial state, indicated by different colored dots. Training beyond the stage (d) aims to approach the optimized limit cycle shown in Fig. 4(d) for 200 episodes.

average swimming velocity \bar{V} in Eq. (15) is proportional to the odd elasticity k^o , i.e., $\bar{V} \propto k^o$. Notably, the odd elasticity \hat{k}^o is proportional to the actuation velocity up to $\hat{U} \leq 2$, similarly to $1/\hat{T}$.

The fact that the effective odd elasticity \hat{k}^o is proportional to the loop frequency $1/\hat{T}$ is consistent with our model of a stochastic odd microswimmer in which the presence of odd elasticity was implemented [31,59]. In the odd microswimmer model, the probability flux forms a closed loop and the eigenvalues of the corresponding frequency matrix are proportional to the odd elasticity. For the time-correlation functions in general odd Langevin systems [30], it was shown that odd elasticity determines the frequency of the sinusoidal component both in their symmetric and antisymmetric parts. These results support the relationship wherein the work per cycle is generally determined by the product of the odd elasticity and the closed loop area in active systems [19,20].

V. TRAINING PROGRESSION OF ELASTIC MICROSWIMMER

In this section, we discuss briefly the dynamic evolution of swimming behavior during the training process. The plots in Figs. 8(a)–8(d) illustrate the configuration space trajectories \hat{u}_α ($\alpha = A, B$) at distinct stages of training in Fig. 3(a), namely 5, 10, 20, and 40 episodes, respectively, when $\hat{U} = 2$. Each trajectory starts from a different initial state denoted by colored dots.

In the initial stage of training, as shown in Fig. 8(a), the swimming behavior is relatively restricted. Starting from different initial states, the system only evolves towards the point $(2, -2)$, showing limited exploration of the motion possibilities. This early stage primarily focuses on optimizing short-term rewards where cyclic motion has not emerged yet. As the microswimmer gains more training experience, its behavior progressively improves. After approximately 10 episodes, as shown in Fig. 8(b), the microswimmer begins to explore long-term rewards accessible through cyclic motions. Enclosed loops emerge in the configuration space, and the microswimmer starts to recognize the dynamics required for sustainable cyclic locomotion.

In Fig. 8(c) at around 20 training episodes, the shape of loops becomes clearer and more refined. This stage shows the swimmer's ability to fine-tune its motion strategy toward

the optimal cyclic pattern to maximize long-term rewards. After training for 40 episodes, the microswimmer achieves the optimized limit cycle, as shown in Fig. 8(d). Importantly, this optimized cycle is robust across various initial conditions, and it will eventually approach the limit cycle shown in Fig. 4(d).

VI. SUMMARY AND DISCUSSION

Using deep reinforcement learning (Deep Q-Network), we have investigated how the effective odd elasticity emerges when optimizing the swimming ability of an elastic microswimmer [36–38]. One of the key findings is the optimized natural-length dynamics without the need for prescribed motion. Notably, we observed a strategy transition (the emergence of waiting behavior) when the actuation velocity $\hat{U} \approx 2$ (Fig. 4). For larger \hat{U} , the trained microswimmers adapt to the slow hydrodynamic relaxation and avoid the velocity decrease. This waiting strategy significantly improves the swimming ability compared to the elastic microswimmer with prescribed dynamics having large-frequency oscillations. Additionally, the trained microswimmer exhibits fore-aft asymmetry in the spring amplitudes (Fig. 5), which is generally difficult to predict and implement in the prescribed dynamics.

By calculating the force-displacement correlations for fully trained microswimmers, we have extracted the effective even and odd elasticities, k' and k^o , respectively (Fig. 7). We have shown that the U dependencies of k' and k^o closely resemble those of the nonreciprocity R in Fig. 6(a) and the loop frequency $1/T$ in Fig. 6(b), respectively. From the numerical data, we have further confirmed that the average swimming velocity \bar{V} is proportional to the cycle performance R/T [Figs. 6(c) and 6(d)], as predicted in Eq. (15). These results clarify the proportionality between the average velocity and odd elasticity, $\bar{V} \propto k^o$. Our study demonstrates the use of machine learning to reveal the emergence of odd elasticity in various active systems.

If we assume that the energy injected through the nonreciprocal process balances with the dissipation due to the net motion of a microswimmer, then its power (work per unit time) scales as $\dot{W} \sim (\eta a \bar{V}) \times \bar{V} \propto (k^o)^2$. Here $\eta a \bar{V}$ corresponds to the dissipative force, and we have used the relation $\bar{V} \propto k^o$ [12]. For an odd microswimmer, it was further shown that all the extracted work due to odd elasticity is converted into the entropy production rate [31]. Hence, odd elasticity

is useful to characterize nonreciprocal dynamics of active micromachines, including not only microswimmers but also other molecular motors [29,30]. With our analysis method, the work performance of these micromachines can be quantified by observing their deformation, even without any precise understanding of their specific functions.

In our machine learning model, we adopted the DQN algorithm with a discrete action space to train elastic microswimmers. Since the swimming ability is directly related to the effective strength of nonreciprocal forces, the optimized dynamics are to adjust one of the natural lengths to either its maximum or minimum value before changing the other. For such strategies, the microswimmer does not change the two natural lengths simultaneously, and thus continuous-action algorithms are unnecessary. To validate this argument, we have also employed other continuous-action algorithms, such as deep deterministic policy gradient [66] and soft actor-critic [67]. Both of these methods resulted in the same strategy as the discrete DQN algorithm. Although these continuous-action algorithms are helpful for more complicated microswimmers, the discrete DQN algorithm used in this work is sufficient for an elastic microswimmer moving in a one-dimensional space.

Microswimmers composed of biomaterials are often soft and exhibit viscoelastic responses. This softness is crucial not only because of the inevitable interaction with viscoelastic environments but also due to their versatile functionalities [68–74], such as mechanical signal sensing [75,76], cargo loading and unloading [77,78], and navigation through intricate channels [79]. On the other hand, when one considers manipulating the gait switching of a microswimmer from a simple potential field, our elastic model becomes more suitable. For example, if we consider manipulating the spheres through optical tweezers or a harmonic electromagnetic field [58,72], then the corresponding potential could be determined effectively through Eqs. (1)–(3).

We comment that the even and odd elasticities obtained from the numerical data are assumed to be linear. In a more general scenario, however, these elasticities can be nonlinear [20]. Since a diagonal positive elasticity generally leads to a decaying oscillation [34], nonlinearity is commonly required for odd-elastic systems to exhibit a stable limit cycle [26]. The inclusion of nonlinear elasticity can describe more general cases, such as the spontaneous onset of oscillations with a specific amplitude that is regulated by the ratio of linear and nonlinear even elasticities [27]. In the present work, the oscillation amplitudes are determined by the constraint that

clips the natural lengths, and hence the linear approach is more suitable.

In our study, a crossover between the elastic- and hydrodynamic-dominated limits generally exists between microswimmers' large- and small-frequency limits. For such systems, we have shown that a swimming transition strategy around the crossover actuation velocity is desired for the optimized performance because a simple prescribed periodic motion becomes less efficient in the large-frequency regime. Crucially, we emphasize a fundamental characteristic that exists in such systems: a significant delay between the applied control (ℓ_α) and the shape responses (L_α). This delay induces strategy transitions that can generally exist in various microswimmers.

Applications that use machine learning to optimize gait switching for microswimmer navigation have been widely studied in recent literatures [39–52]. These studies have provided valuable insights into the potential of machine learning to optimize the dynamics of microswimmers. In the current work, in addition to focusing on performance optimization, we have emphasized the possibility of extracting useful information such as nonreciprocity and odd elasticity. In future studies, we aim to establish universal relations between performance and odd elasticity for different types of microswimmers [15].

ACKNOWLEDGMENTS

We thank J. Shuai and X. Xu for the useful discussion. L.S.L. was supported by Tokyo Human Resources Fund for City Diplomacy. K.Y. acknowledges the support by a Grant-in-Aid for JSPS Fellows (No. 22KJ1640) from the Japan Society for the Promotion of Science (JSPS). K.I. acknowledges the JSPS, KAKENHI for Transformative Research Areas A (No. 21H05309) and the Japan Science and Technology Agency (JST), FOREST Grant (No. JPMJFR212N). S.K. acknowledges the support by the National Natural Science Foundation of China (Grants No. 12274098 and No. 12250710127) and the startup grant of Wenzhou Institute, University of Chinese Academy of Sciences (Grant No. WIUCASQD2021041). L.S.L., K.Y., and K.I. were supported by the Research Institute for Mathematical Sciences, an International Joint Usage/Research Center located in Kyoto University. K.Y., K.I., and S.K. acknowledge the support by the JSPS Core-to-Core Program “Advanced core-to-core network for the physics of self-organizing active matter (Grant No. JPJSCCA20230002).

[1] M. C. Marchetti, J. F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao, and R. A. Simha, Hydrodynamics of soft active matter, *Rev. Mod. Phys.* **85**, 1143 (2013).
 [2] G. Gompper, R. G. Winkler, T. Speck, A. Solon, C. Nardini, F. Peruani, H. Löwen, R. Golestanian, U. B. Kaupp, L. Alvarez *et al.*, The 2020 motile active matter roadmap, *J. Phys.: Condens. Matter* **32**, 193001 (2020).
 [3] S. Shankar, A. Souslov, M. J. Bowick, M. C. Marchetti, and V. Vitelli, Topological active matter, *Nat. Rev. Phys.* **4**, 380 (2022).

[4] K. K. Dey, Dynamic coupling at low Reynolds number, *Angew. Chem. Int. Ed.* **58**, 2208 (2019).
 [5] Y. Hosaka and S. Komura, Nonequilibrium transport induced by biological nanomachines, *Biophys. Rev. Lett.* **17**, 51 (2022).
 [6] K. K. Dey, F. Wong, A. Altemose, and A. Sen, Catalytic motors—Quo vadimus? *Curr. Opin. Colloid Interface Sci.* **21**, 4 (2016).
 [7] H. Y. Aviram, M. Pirchi, H. Mazal, Y. Barak, I. Riven, and G. Haran, Direct observation of ultrafast large-scale dynamics of

- an enzyme under turnover conditions, *Proc. Natl. Acad. Sci. USA* **115**, 3243 (2018).
- [8] C. Echeverria, Y. Togashi, A. S. Mikhailov, and R. Kapral, A mesoscopic model for protein enzymatic dynamics in solution, *Phys. Chem. Chem. Phys.* **13**, 10527 (2011).
- [9] A. S. Mikhailov and R. Kapral, Hydrodynamic collective effects of active protein machines in solution and lipid bilayers, *Proc. Natl. Acad. Sci. USA* **112**, E3639 (2015).
- [10] Y. Hosaka, S. Komura, and A. S. Mikhailov, Mechanochemical enzymes and protein machines as hydrodynamic force dipoles: The active dimer model, *Soft Matter* **16**, 10734 (2020).
- [11] K. Yasuda and S. Komura, Nonreciprocity of a micromachine driven by a catalytic chemical reaction, *Phys. Rev. E* **103**, 062113 (2021).
- [12] A. Kobayashi, K. Yasuda, K. Ishimoto, L.-S. Lin, I. Sou, Y. Hosaka, and S. Komura, Odd elasticity of a catalytic micromachine, *J. Phys. Soc. Jpn.* **92**, 074801 (2023).
- [13] E. M. Purcell, Life at low Reynolds number, *Am. J. Phys.* **45**, 3 (1977).
- [14] E. Lauga and T. R. Powers, The hydrodynamics of swimming microorganisms, *Rep. Prog. Phys.* **72**, 096601 (2009).
- [15] E. Lauga, *The Fluid Dynamics of Cell Motility* (Cambridge University Press, Cambridge, UK, 2020).
- [16] M. Leoni and P. Sens, Polarization of cells and soft objects driven by mechanical interactions: Consequences for migration and chemotaxis, *Phys. Rev. E* **91**, 022720 (2015).
- [17] M. Leoni and P. Sens, Model of cell crawling controlled by mechanosensitive adhesion, *Phys. Rev. Lett.* **118**, 228101 (2017).
- [18] M. Tarama and R. Yamamoto, Mechanics of cell crawling by means of force-free cyclic motion, *J. Phys. Soc. Jpn.* **87**, 044803 (2018).
- [19] C. Scheibner, A. Souslov, D. Banerjee, P. Surówka, W. T. Irvine, and V. Vitelli, Odd elasticity, *Nat. Phys.* **16**, 475 (2020).
- [20] M. Fruchart, C. Scheibner, and V. Vitelli, Odd viscosity and odd elasticity, *Annu. Rev. Condens. Matter Phys.* **14**, 471 (2023).
- [21] L. Braverman, C. Scheibner, B. VanSaders, and V. Vitelli, Topological defects in solids with odd elasticity, *Phys. Rev. Lett.* **127**, 268001 (2021).
- [22] E. S. Bililign, F. B. Usabiaga, Y. A. Ganan, A. Poncet, V. Soni, S. Magkiriadou, M. J. Shelley, D. Bartolo, and W. T. Irvine, Motile dislocations knead odd crystals into whorls, *Nat. Phys.* **18**, 212 (2022).
- [23] P. Surówka, A. Souslov, F. Jülicher, and D. Banerjee, Odd Cosserat elasticity in active materials, *Phys. Rev. E* **108**, 064609 (2023).
- [24] M. Fossati, C. Scheibner, M. Fruchart, and V. Vitelli, Odd elasticity and topological waves in active surfaces, *Phys. Rev. E* **109**, 024608 (2024).
- [25] T. H. Tan, A. Mietke, J. Li, Y. Chen, H. Higinbotham, P. J. Foster, S. Gokhale, J. Dunkel, and N. Fakhri, Odd dynamics of living chiral crystals, *Nature (London)* **607**, 287 (2022).
- [26] K. Ishimoto, C. Moreau, and K. Yasuda, Self-organized swimming with odd elasticity, *Phys. Rev. E* **105**, 064603 (2022).
- [27] K. Ishimoto, C. Moreau, and K. Yasuda, Odd elastohydrodynamics: Non-reciprocal living material in a viscous fluid, *PRX Life* **1**, 023002 (2023).
- [28] M. Brandenbourger, C. Scheibner, J. Veenstra, V. Vitelli, and C. Coulais, Limit cycles turn active matter into robots, [arXiv:2108.08837](https://arxiv.org/abs/2108.08837).
- [29] K. Yasuda, A. Kobayashi, L.-S. Lin, Y. Hosaka, I. Sou, and S. Komura, The Onsager-Machlup integral for non-reciprocal systems with odd elasticity, *J. Phys. Soc. Jpn.* **91**, 015001 (2022).
- [30] K. Yasuda, K. Ishimoto, A. Kobayashi, L.-S. Lin, I. Sou, Y. Hosaka, and S. Komura, Time-correlation functions for odd Langevin systems, *J. Chem. Phys.* **157**, 095101 (2022).
- [31] K. Yasuda, Y. Hosaka, I. Sou, and S. Komura, Odd microswimmer, *J. Phys. Soc. Jpn.* **90**, 075001 (2021).
- [32] N. Mitarai, H. Hayakawa, and H. Nakanishi, Collisional granular flow as a micropolar fluid, *Phys. Rev. Lett.* **88**, 174301 (2002).
- [33] M. Doi, *Soft Matter Physics* (Oxford University Press, Oxford, 2013).
- [34] L.-S. Lin, K. Yasuda, K. Ishimoto, Y. Hosaka, and S. Komura, Onsager's variational principle for nonreciprocal systems with odd elasticity, *J. Phys. Soc. Jpn.* **92**, 033001 (2023).
- [35] M. Fruchart, R. Hanai, P. B. Littlewood, and V. Vitelli, Non-reciprocal phase transitions, *Nature (London)* **592**, 363 (2021).
- [36] K. Yasuda, Y. Hosaka, M. Kuroda, R. Okamoto, and S. Komura, Elastic three-sphere microswimmer in a viscous fluid, *J. Phys. Soc. Jpn.* **86**, 093801 (2017).
- [37] J. Pande and A.-S. Smith, Forces and shapes as determinants of micro-swimming: Effect on synchronisation and the utilisation of drag, *Soft Matter* **11**, 2364 (2015).
- [38] J. Pande, L. Merchant, T. Krüger, J. Harting, and A.-S. Smith, Setting the pace of microswimmers: When increasing viscosity speeds up self-propulsion, *New J. Phys.* **19**, 053024 (2017).
- [39] F. Cichos, K. Gustavsson, B. Mehlig, and G. Volpe, Machine learning for active matter, *Nat. Mach. Intell.* **2**, 94 (2020).
- [40] H. Stark, Artificial microswimmers get smart, *Sci. Robot.* **6**, eabh1977 (2021).
- [41] B. Hartl, M. Hübl, G. Kahl, and A. Zöttl, Microswimmers learning chemotaxis with genetic algorithms, *Proc. Natl. Acad. Sci. USA* **118**, e2019683118 (2021).
- [42] M. Putzke and H. Stark, Optimal navigation of a smart active particle: Directional and distance sensing, *Eur. Phys. J. E* **46**, 48 (2023).
- [43] E. Schneider and H. Stark, Optimal steering of a smart active particle, *Europhys. Lett.* **127**, 64003 (2019).
- [44] 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).
- [45] S. Muiños-Landin, A. Fischer, V. Holubec, and F. Cichos, Reinforcement learning with artificial microswimmers, *Sci. Robot.* **6**, eabd9285 (2021).
- [46] M. Nasiri, H. Löwen, and B. Liebchen, Optimal active particle navigation meets machine learning, *Europhys. Lett.* **142**, 17001 (2023).
- [47] A. C. H. Tsang, P. W. Tong, S. Nallan, and O. S. Pak, Self-learning how to swim at low Reynolds number, *Phys. Rev. Fluids* **5**, 074101 (2020).
- [48] Z. Zou, Y. Liu, Y.-N. Young, O. S. Pak, and A. C. H. Tsang, Gait switching and targeted navigation of microswimmers via deep reinforcement learning, *Commun. Phys.* **5**, 158 (2022).
- [49] K. Qin, Z. Zou, L. Zhu, and O. S. Pak, Reinforcement learning of a multi-link swimmer at low Reynolds numbers, *Phys. Fluids* **35**, 032003 (2023).
- [50] 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).

- [51] G. Zhu, W. Z. Fang, and L. Zhu, Optimizing low-Reynolds-number predation via optimal control and reinforcement learning, *J. Fluid Mech.* **944**, A3 (2022).
- [52] Y. Liu, Z. Zou, O. S. Pak, and A. C. H. Tsang, Learning to cooperate for low-Reynolds-number swimming: A model problem for gait coordination, *Sci. Rep.* **13**, 9397 (2023).
- [53] A. Najafi and R. Golestanian, Simple swimmer at low Reynolds number: Three linked spheres, *Phys. Rev. E* **69**, 062901 (2004).
- [54] R. Golestanian and A. Ajdari, Analytic results for the three-sphere swimmer at low Reynolds number, *Phys. Rev. E* **77**, 036308 (2008).
- [55] M. Kuroda, K. Yasuda, and S. Komura, Hydrodynamic interaction between two elastic microswimmers, *J. Phys. Soc. Jpn.* **88**, 054804 (2019).
- [56] K. Era, Y. Koyano, Y. Hosaka, K. Yasuda, H. Kitahata, and S. Komura, Autonomous elastic microswimmer, *Europhys. Lett.* **133**, 34001 (2021).
- [57] K. Yasuda, Y. Hosaka, and S. Komura, Generalized three-sphere microswimmers, *J. Phys. Soc. Jpn.* **92**, 121008 (2023).
- [58] G. Grosjean, M. Hubert, G. Lagubeau, and N. Vandewalle, Realization of the Najafi-Golestanian microswimmer, *Phys. Rev. E* **94**, 021101(R) (2016).
- [59] A. Kobayashi, K. Yasuda, L.-S. Lin, I. Sou, Y. Hosaka, and S. Komura, Simulations of odd microswimmers, *J. Phys. Soc. Jpn.* **92**, 034803 (2023).
- [60] A. Shapere and F. Wilczek, Geometry of self-propulsion at low Reynolds number, *J. Fluid Mech.* **198**, 557 (1989).
- [61] K. Yasuda, R. Okamoto, and S. Komura, Swimmer-microrheology, *J. Phys. Soc. Jpn.* **86**, 043801 (2017).
- [62] K. Yasuda, M. Kuroda, and S. Komura, Reciprocal microswimmers in a viscoelastic fluid, *Phys. Fluids* **32**, 093102 (2020).
- [63] C. J. C. H. Watkins and P. Dayan, Q-learning, *Mach. Learn.* **8**, 279 (1992).
- [64] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, Playing Atari with deep reinforcement learning, [arXiv:1312.5602](https://arxiv.org/abs/1312.5602).
- [65] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction* (MIT Press, Cambridge, MA, 2018).
- [66] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, Continuous control with deep reinforcement learning, [arXiv:1509.02971](https://arxiv.org/abs/1509.02971).
- [67] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor, in *Proceedings of the 35th International Conference on Machine Learning* (2018), Vol. 80, pp. 1861–1870.
- [68] D. Schamel, A. G. Mark, J. G. Gibbs, C. Miksch, K. I. Morozov, A. M. Leshansky, and P. Fischer, Nanopropellers and their actuation in complex viscoelastic media, *ACS Nano* **8**, 8794 (2014).
- [69] D. Walker, B. T. Käs Dorf, H. H. Jeong, O. Lieleg, and P. Fischer, Enzymatically active biomimetic micropropellers for the penetration of mucin gels, *Sci. Adv.* **1**, e1500501 (2015).
- [70] Z. Wu, J. Troll, H.-H. Jeong, Q. Wei, M. Stang, F. Ziemssen, Z. Wang, M. Dong, S. Schnichels, T. Qiu, and P. Fischer, A swarm of slippery micropropellers penetrates the vitreous body of the eye, *Sci. Adv.* **4**, eaat4388 (2018).
- [71] M. Medina-Sánchez, V. Magdanz, M. Guix, V. M. Fomin, and O. G. Schmidt, Swimming microrobots: Soft, reconfigurable, and smart, *Adv. Funct. Mater.* **28**, 1707228 (2018).
- [72] H.-W. Huang, F. E. Uslu, P. Katsamba, E. Lauga, M. S. Sakar, and B. J. Nelson, Adaptive locomotion of artificial microswimmers, *Sci. Adv.* **5**, eaau1532 (2019).
- [73] F. Striggow, M. Medina-Sánchez, G. K. Auernhammer, V. Magdanz, B. M. Friedrich, and O. G. Schmidt, Sperm-driven micromotors moving in oviduct fluid and viscoelastic media, *Small* **16**, 2000213 (2020).
- [74] F. Soto, E. Karshalev, F. Zhang, B. E. F. de Avila, A. Nourhani, and J. Wang, Smart materials for microrobots, *Chem. Rev.* **122**, 5365 (2022).
- [75] T. Xu, J. Zhang, M. Salehizadeh, O. Onaizah, and E. Diller, Millimeter-scale flexible robots with programmable three-dimensional magnetization and motions, *Sci. Robot.* **4**, eaav4494 (2019).
- [76] X. Huo, Y. Wu, A. Boymelgreen, and G. Yossifon, Analysis of cargo loading modes and capacity of an electrically-powered active carrier, *Langmuir* **36**, 6963 (2020).
- [77] Y. Alapan, O. Yasa, O. Schauer, J. Giltinan, A. F. Tabak, V. Sourjik, and M. Sitti, Soft erythrocyte-based bacterial microswimmers for cargo delivery, *Sci. Robot.* **3**, eaar4423 (2018).
- [78] H. Xu, M. Medina-Sánchez, M. F. Maitz, C. Werner, and O. G. Schmidt, Sperm micromotors for cargo delivery through flowing blood, *ACS Nano* **14**, 2982 (2020).
- [79] H. R. Vutukuri, M. Hoore, C. Abaurrea-Velasco, L. van Buren, A. Dutto, T. Auth, D. A. Fedosov, G. Gompper, and J. Vermant, Active particles induce large shape deformations in giant lipid vesicles, *Nature (London)* **586**, 52 (2020).