






## Collective motion in large deviations of active particles

Yann-Edwin Keta <sup>1,2,3</sup> Étienne Fodor <sup>1,4</sup> Frédéric van Wijland <sup>2</sup> Michael E. Cates <sup>1</sup> and Robert L. Jack <sup>1,5</sup>

<sup>1</sup>*Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, United Kingdom*

<sup>2</sup>*Université de Paris, Laboratoire Matière et Systèmes Complexes (MSC), UMR 7057 CNRS, F-75205 Paris, France*

<sup>3</sup>*Département de Physique, École normale supérieure de Lyon, 69364 Lyon Cedex 07, France*

<sup>4</sup>*Department of Physics and Materials Science, University of Luxembourg, L-1511 Luxembourg*

<sup>5</sup>*Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom*



(Received 17 September 2020; accepted 12 January 2021; published 4 February 2021)

We analyze collective motion that occurs during rare (large deviation) events in systems of active particles, both numerically and analytically. We discuss the associated dynamical phase transition to collective motion, which occurs when the active work is biased towards larger values, and is associated with alignment of particles' orientations. A finite biasing field is needed to induce spontaneous symmetry breaking, even in large systems. Particle alignment is computed exactly for a system of two particles. For many-particle systems, we analyze the symmetry breaking by an optimal-control representation of the biased dynamics, and we propose a fluctuating hydrodynamic theory that captures the emergence of polar order in the biased state.

DOI: [10.1103/PhysRevE.103.022603](https://doi.org/10.1103/PhysRevE.103.022603)

### I. INTRODUCTION

#### A. Motivation

Active matter emerged in the last decades as a novel class of nonequilibrium soft systems where every constituent consumes and dissipates energy to produce a self-propelled motion [1–3]. It includes both living and social systems, such as swarms of bacteria [4,5], bird flocks [6,7], and human crowds [8,9], as well as synthetic systems, such as vibrated particles [10,11] and self-catalytic colloids in a fuel bath [12,13]. In these experimental systems, the combination of self-propulsion and interaction can lead to collective behavior without any equilibrium equivalent. Collective motion with orientational order [10,11] and the spontaneous formation of particle clusters despite the absence of attractive interactions [12,13] are celebrated examples.

Minimal models have been proposed to capture these collective effects, with a view to identifying the essential ingredients of the dynamics which delineate generic classes of active matter. The emergence of collective motion is generally described by the Vicsek model in terms of aligning active particles [14], and its equivalent Toner-Tu model at hydrodynamic level [15,16], whereas purely repulsive active particles yielding a motility-induced phase separation (MIPS) are usually considered to reproduce the behavior of isotropic self-propelled particles [17–19]. To characterize the structure and dynamics, thermodynamic tools inspired by equilibrium have been proposed, such as pressure [20–22], others focus specifically on the deviation from equilibrium, such as the irreversibility of the dynamics [23–28] and the dissipation of energy [29–33].

Several recent studies focused on large deviations of active matter [32–42]. They consider transient rare events where

the system does not behave ergodically, in the sense that time-averaged quantities differ significantly from the corresponding ensemble averages, even for long times. In these nonequilibrium systems, natural time-averaged quantities include measures of entropy production, dissipation, and work [32–34,37,40–42], which are associated with breaking of time-reversal symmetry, and whose large deviation behavior obeys fluctuation theorems [43–45]. The associated rare events are often accompanied by collective effects, and may also lead to dynamical phase transitions, where symmetry is spontaneously broken and atypical trajectories differ significantly from the typical ones [46–50]. Numerical techniques can be used to analyze these transient events by introducing a bias parameter which controls the distance from the typical dynamics [51,52]. They open the door to studying the microscopic mechanisms that stabilize atypical collective behaviors. These techniques have already proved successful to unveil dynamical transitions in glassy dynamics [53–55] and high-dimensional chaotic chains [56,57].

In active matter, large deviation studies have already revealed several kinds of dynamical phase transition. For example, one may observe macroscopic clusters of particles [32,33,37] and dynamically arrested states [37]. These transitions are similar to those found in passive fluids [49,54,55,58]. Moreover, it was shown recently [37] that large deviations of isotropic active particles can be associated with collective motion; this effect is not found in passive systems. The physical origin of this collective motion is that orientational order reduces the frequency with which particles collide, which facilitates their motion. Given the established links between large deviations and optimal-control theory [49,50,59], this result indicates that fluctuations in orientational order might be harnessed, to enable efficient particle transport in

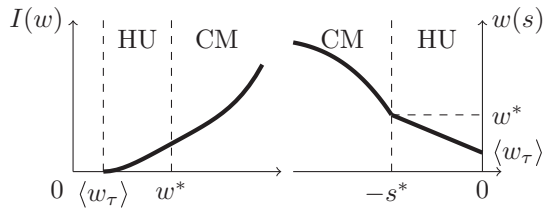


FIG. 1. Schematic behavior of (left) the rate function, following Fig. 1(a) of [37] and Fig. 2, and (right) the active work  $w_\tau$  as a function of its conjugate field  $s$ , following Fig. 1(b) of [37] and Figs. 3 and 4. Vertical dashed lines delimit the two regimes identified: CM  $\equiv$  collectively moving state, HU  $\equiv$  isotropic hyperuniform state. For positive  $s$  (or equivalently  $w < \langle w_\tau \rangle$ ), the system is phase separated and arrested [37]; that case is not discussed here.

nonequilibrium systems. It also means that long-ranged orientational order appears in large deviation events, despite the absence of any microscopic interactions that favor alignment. This stands in contrast to the usual (and intuitive) expectation that collective motion emerges as a result of particle alignment.

### B. Summary of main results

In this paper, we extend the analysis of collective motion in Ref. [37], improving our understanding of this collective motion phase, including the mechanism of spontaneous symmetry breaking, the location of the phase transition, and the relationship of the collective motion with the hydrodynamic dynamics of the system. We consider active Brownian particles (ABPs) as a popular model of overdamped self-propelled particles [17,18]. For a long time interval of duration  $\tau$ , we focus on the time-averaged rate of the active work per particle  $w_\tau$ , which quantifies how much the self-propulsion forces of particles translate into actual displacement. The ensemble-averaged rate is  $\langle w_\tau \rangle$ . The active work is a measure of the efficiency of self-propulsion, and is linked with dissipation and entropy production; it also obeys a fluctuation theorem [37]. Full definitions are given in Sec. II, below.

Consistent with previous observations of collective motion [37], we focus on large deviations where the active work is increased. The resulting picture is summarized in Fig. 1, as a function of the active work  $w$ , and also its conjugate field  $s$ . We restrict to situations where the steady state of the system is spatially homogeneous, so the activity of the particles is not enough to cause MIPS. We find that spontaneous symmetry breaking occurs for values of the active work  $w_\tau$  beyond a threshold  $w^*$  that is strictly greater than its average value  $\langle w_\tau \rangle$ . There is a corresponding threshold for the biasing field, in that collective motion takes place for  $s < -s^*$  (this sign convention is chosen so that  $s^* > 0$ , it means that the transition takes place at  $s = -s^*$  and not  $s = s^*$ ). This result is supported by a finite-size scaling analysis.

For ABPs, an important parameter is the rotational diffusion constant  $D_r$  which determines the correlation time of the self-propulsion force. We find that  $s^* \sim D_r$  for small  $D_r$ , which is the regime where the system differs strongly from a passive fluid. A consequence of this analysis is that the system is an isotropic fluid for  $-s^* < s < 0$ . Generic arguments

[49,50,60] based on coupling between large deviations and hydrodynamic modes mean that this phase is hyperuniform [61]. (The collective motion phase may also be expected to have a similar property but that question is not addressed here.)

The fact that  $s^*$  is strictly positive (not zero) resolves an open question from [37]. If  $s^* = 0$ , then symmetry breaking would be present for all  $w_\tau > \langle w_\tau \rangle$ , in sufficiently large systems. Such transitions can occur (always in the thermodynamic limit) [37,53]: using the optimal-control formulation of large deviations [49,50,59], they imply that a system's behavior may be changed qualitatively by very weak control forces, or by applying significant forces to a very small fraction of particles [37,50,60]. This is not the case for the collective motion transition found here: sustaining long-ranged orientational order requires control forces of order unity.

To analyze the mechanism of symmetry breaking, we first solve exactly a system of two active run-and-tumble particles (RTPs) to demonstrate that particles naturally align during large-deviation events. Turning to many-particle systems, we exploit connections between large-deviation theory and optimal control theory [49,50,59], and we also develop a Landau-Ginzburg theory for the symmetry-breaking transition, which includes both the orientational order of the ABPs, and their hydrodynamic density fluctuations. This gives a detailed description of the collective motion phase.

The structure of the of the paper is as follows: Sec. II describes the models and outlines the theoretical background; Sec. III presents numerical results for collective motion; Sec. IV analyzes the two-particle case; Sec. V explores the mechanism using large-deviation bounds based on controlled systems with orientational interactions; Sec. VI describes a Landau-Ginzburg theory for the collective motion transition. Conclusions are summarized in Sec. VII and several appendices contain additional technical information.

## II. MODEL AND METHODS

### A. Active Brownian particles

We consider  $N$  active Brownian particles (ABPs) in two spatial dimensions [37]. Their positions and orientations are  $\mathbf{r}_i$  and  $\theta_i$ . We define an orientation vector  $\mathbf{u}(\theta_i) = (\cos \theta_i, \sin \theta_i)$  which we sometimes abbreviate simply as  $\mathbf{u}_i$ . The particles are self-propelled with (bare) speed  $v_0$ , they interact through a Weeks-Chandler-Andersen (WCA) interaction potential  $V(r)$  with range  $\sigma$  and strength  $\epsilon_0$ . Define

$$U = \frac{1}{T} \sum_{1 \leq i < j \leq N} V(|\mathbf{r}_i - \mathbf{r}_j|) \quad (1)$$

as the dimensionless potential energy, which has been rescaled by the temperature  $T$ . We take Boltzmann's constant  $k_B = 1$ . The equations of motion are

$$\begin{aligned} \dot{\mathbf{r}}_i &= v_0 \mathbf{u}(\theta_i) - D \nabla_i U + \sqrt{2D} \boldsymbol{\eta}_i, \\ \dot{\theta}_i &= \sqrt{2D_r} \xi_i, \end{aligned} \quad (2)$$

where  $\boldsymbol{\eta}_i, \xi_i$  are zero-mean unit-variance Gaussian white noises, and  $D, D_r$  are translational and rotational diffusivities. The combination  $D \nabla_i U$  in Eq. (2) should be interpreted as the

product of a mobility  $\mu$  and the gradient of potential energy; here we have used  $\mu = D/T$  by the fluctuation-dissipation theorem.

For consistency with Refs. [18,37], we set  $D_r = 3D/\sigma^2$  in accordance with the Stokes-Einstein-Debye relation. The particles are contained in a periodic box of size  $L \times L$ , the dimensionless measure of density is  $\phi = N\pi\sigma^2/(4L^2)$ . For numerical work we consider a single density  $\phi = 0.65$ ; results for other densities are similar [37]. We also take  $T = \varepsilon_0$ .

For a single isolated particle, the effect of the self-propulsion force is that the particle follows a persistent random walk with persistence length  $l_p = v_0/D_r$ . Dividing this length by the particle diameter defines an important dimensionless control parameter

$$\tilde{l}_p = \frac{v_0}{\sigma D_r}, \quad (3)$$

which determines the effect of the active self-propulsion. For  $\tilde{l}_p \rightarrow 0$  we have a passive (equilibrium) system. For large  $\tilde{l}_p \gtrsim 15$  the self-propulsion leads to motility-induced phase separation (MIPS) [19]. We note that since  $D \propto D_r$ , increasing  $\tilde{l}_p$  changes the balance between the self-propulsion term and the repulsive (WCA) forces in Eq. (2). This means that large  $\tilde{l}_p$  tends to make particles overlap more: they appear to be softer.

When presenting numerical results, we take  $\sigma = 1$  as the unit of length and we fix the time unit by setting also  $v_0 = 1$ . For theoretical calculations, we retain  $v_0$  and  $\sigma$  as explicit quantities.

### B. Dissipation and active work

We define the instantaneous dissipated power from a purely mechanical argument as the rate of work that the particles exert on the solvent [45,62]

$$\dot{\mathcal{W}} = \sum_i \dot{\mathbf{r}}_i \circ \frac{1}{D} (\dot{\mathbf{r}}_i - \sqrt{2D}\boldsymbol{\eta}_i), \quad (4)$$

where  $\circ$  is a Stratonovich product. We have absorbed a factor of  $T$  into  $\mathcal{W}$ , to obtain a reduced (dimensionless) work. Here and in the following, sums run over all particles, unless otherwise stated. Using (2) and taking a time average, we find

$$\frac{1}{\tau} \int_0^\tau \dot{\mathcal{W}}(t) dt = \frac{Nv_0^2}{D} w_\tau + \frac{1}{\tau} [U(\tau) - U(0)], \quad (5)$$

where

$$w_\tau = \frac{1}{v_0 N \tau} \sum_i \int_0^\tau \mathbf{u}(\theta_i) \circ d\mathbf{r}_i \quad (6)$$

is the (reduced) active work per particle [37]. This is a natural measure of how efficiently active forces create motion. It is normalized such that  $\langle w_\tau \rangle = 1$  in the dilute limit  $\phi \rightarrow 0$ , while  $\langle w_\tau \rangle = 0$  for a completely jammed system. For a steady state, the term involving  $U$  in Eq. (5) is zero on average, so the average dissipation is fully determined by the average of the active work.

This active work  $w_\tau$  is also related to the entropy production rate in the full  $\{\mathbf{r}_i, \theta_i\}$  configuration space (which considers self-propulsion as a quantity that is even under time

reversal [27,37]). This differs in general from the entropy production measured in position space  $\{\mathbf{r}_i\}$  [23,63,64].

Since (2) has three separate contributions, there is a natural decomposition of the active work

$$w_\tau = 1 + w_{f,\tau} + w_{\eta,\tau}, \quad (7)$$

where the constant term stems from the product of the self-propulsion direction with itself and

$$w_{f,\tau} = \frac{-D}{v_0 N \tau} \sum_i \int_0^\tau \mathbf{u}(\theta_i) \cdot \nabla_i U dt, \quad (8)$$

$$w_{\eta,\tau} = \frac{1}{v_0 N \tau} \sum_i \int_0^\tau \mathbf{u}(\theta_i) \circ \sqrt{2D}\boldsymbol{\eta}_i dt. \quad (9)$$

On average  $\langle w_{\eta,\tau} \rangle = 0$ , so  $\langle w_\tau \rangle = 1 + \langle w_{f,\tau} \rangle$ . The quantity  $w_{f,\tau}$  is negative on average because collisions between particles tend to involve particle orientation vectors being antiparallel to the interparticle force. In the following, we consider situations where the system self-organizes to reduce collisions, in which case  $w_{f,\tau}$  becomes less negative (it increases towards zero).

### C. Large deviations

For any given  $N$ , the active work  $w_\tau$  satisfies a large-deviation principle in the limit of large  $\tau$  [37]:

$$p(w_\tau) \asymp \exp[-\tau N I(w_\tau)], \quad (10)$$

where  $I(w_\tau)$  is a scaled rate function. We define the scaled cumulant generating function (SCGF)

$$\psi(s) = \lim_{\tau \rightarrow \infty} \frac{1}{N\tau} \ln \langle \exp(-sN\tau w_\tau) \rangle, \quad (11)$$

related to  $I(w_\tau)$  by Legendre transformation [see Eq. (14) below], and where we have introduced a biasing parameter  $s$ . The SCGF can be obtained by solving an eigenvalue problem (see Appendix A).

There is a useful analogy between this dynamical large deviation formalism and equilibrium statistical mechanics. We recall the central features of this analogy (see also [46,47,50]). Trajectories of our two-dimensional system are analogous to configurations of a  $(2+1)$ -dimensional system. Also, the biasing field  $s$  corresponds to a thermodynamic field conjugate to the active work. The SCGF  $\psi(s)$  corresponds to the free energy density, and is thus sometimes referred to as the dynamical free energy. Any singularity in this function is a signature of a dynamical phase transition. In particular, we focus below on phase transitions where rotational symmetry of the system is spontaneously broken.

Continuing with this analogy, the average in Eq. (11) corresponds to a partition function for Boltzmann-type distribution of trajectories. Averages with respect to this distribution take the form

$$\langle \mathcal{A} \rangle_s = \frac{\langle \mathcal{A} e^{-sN\tau w_\tau} \rangle}{\langle e^{-sN\tau w_\tau} \rangle}, \quad (12)$$

where  $\mathcal{A}$  is a dynamical observable. Numerical computation of such averages is challenging in general: it is comparable to computing a thermodynamic average at temperature  $T$  by reweighting from an equilibrium system with temperature  $T' \neq T$ . To achieve this, we evolve simultaneously a

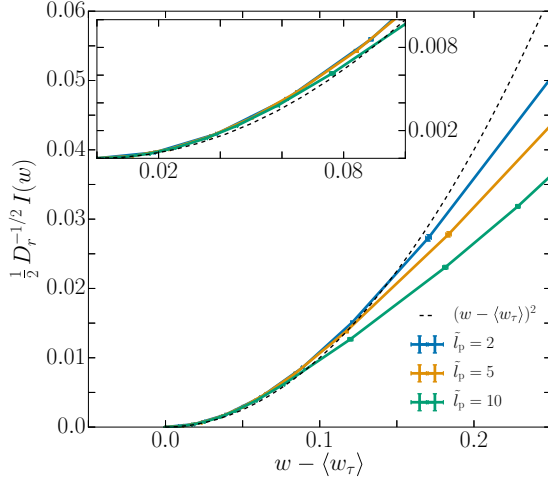


FIG. 2. Rate function  $I(w)$  computed with (14) rescaled by  $D_r^{1/2}$  for persistence lengths  $\tilde{l}_p = 2, 5, 10$ . The inset is a magnified version of the behavior for small  $w - \langle w_\tau \rangle$ . Parameter values:  $N = 50$ ,  $\phi = 0.65$ ,  $n_c = 10^3$ ,  $t_{\max} = 10^3$ .

large population of copies of the system to generate “biased ensembles” by cloning and deleting some of these copies at regular steps in order to enforce the dynamical effective Boltzmann distribution. This method, known as a cloning algorithm [51,65], allows estimation of averages like (11) and (12) with a cost that scales linearly in  $\tau$ , allowing direct access to the large- $\tau$  limit. We implement it following [52,66], using a modified equation of motion to evolve the clones, over a maximal time period  $t_{\max}$ . Details are given in Appendix B.

Of particular interest is the quantity

$$w(s) = \lim_{\tau \rightarrow \infty} \langle w_\tau \rangle_s \quad (13)$$

which obeys  $w(s) = -\psi'(s)$ . Since  $\psi$  is convex [47], this is a decreasing function of  $s$ , we denote its inverse by  $s(w)$ . The rate function  $I$  is related to the SCGF by Legendre transform, in particular,

$$I(w) = -ws(w) - \psi(s(w)) \quad (14)$$

which allows computation of the rate function from the output of a cloning simulation.

Figure 2 shows the rate function of the active work  $w \geq \langle w_\tau \rangle$ , for different persistence lengths  $\tilde{l}_p$ . The vertical axis has been scaled by  $D_r^{-1/2}$  which leads to data collapse near the minimum. The rate function is minimal (and equal to zero) at  $w = \langle w_\tau \rangle$  and its curvature there is related to the variance of the active work as

$$\frac{1}{I''(\langle w_\tau \rangle)} = \lim_{\tau \rightarrow \infty} \tau N [\langle w_\tau^2 \rangle - \langle w_\tau \rangle^2]. \quad (15)$$

Our data suggest that this variance is proportional to  $D_r^{-1/2}$ . As  $w$  increases from  $\langle w_\tau \rangle$ , the rate function deviates from a quadratic form, in particular, its curvature decreases, showing that large fluctuations of  $w_\tau$  are less unlikely than a simple Gaussian approximation would predict. This is related to a dynamical phase transition, as we now explain.

### III. EVIDENCE FOR SYMMETRY BREAKING

After introducing collective motion (Sec. III A), we present numerical evidence for a spontaneous breaking of rotational symmetry, on biasing towards larger values of the active work (Sec. III B). We then show that for  $\tilde{l}_p \gtrsim 2$ , the dependence on persistence length can be rationalized by comparison with a simple controlled system (Sec. III C), but the situation for  $\tilde{l}_p = 1$  is more complex, with several competing effects in play (Sec. III D).

#### A. Collective motion and symmetry breaking

This work builds on Ref. [37], which investigated the same biased ensembles. For  $s > 0$ , i.e., biasing towards trajectories of low active work, one finds a coexistence of a dense jammed, arrested domain with a dilute vapor, given the name phase-separated arrest (PSA). For  $s < 0$ , i.e., biasing towards trajectories of high active work, collective motion is found with aligned propulsion directions, despite the absence of aligning interactions microscopically. In this work, we consider exclusively trajectories with positive fluctuations of the active work ( $s < 0$ ). Reference [37] focused on a system whose parameters lie (as  $N \rightarrow \infty$ ) within the MIPS region,  $\tilde{l}_p = 40$  and  $\phi = 0.65$  (see Ref. [18] for the full phase diagram of the system). Compared with [37], we focus here on lower persistence lengths  $\tilde{l}_p$ , such that the unbiased behavior of the system is that of a homogeneous active fluid.

The physical reason for collective motion when  $s < 0$  is that if particles all travel in the same direction with speed  $v_0$ , they collide much less frequently, so  $w_{f,\tau}$  is increased. In particular, if the  $\mathbf{u}_i$  are random unit vectors, then there are large relative velocities between particles (because  $|\mathbf{u}_i - \mathbf{u}_j|$  is typically of order unity). On the other hand, perfectly aligned orientations lead to  $|\mathbf{u}_i - \mathbf{u}_j| = 0$ , so the only sources of relative motion are the passive noises  $\boldsymbol{\eta}_i, \boldsymbol{\eta}_j$ . The larger the relative velocities, the more often the particles collide, leading to smaller (more negative) values of  $w_{f,\tau}$ . Hence, collective motion is a natural mechanism for increasing  $w_\tau$ .

It is also notable that  $w_\tau$  is closely related to the ratio  $v(\rho)/v_0$  that appears in theories of MIPS [22], and measures the reduction in particle speed due to collisions. This further emphasizes that larger active work corresponds to reduced collisions.

Since the collective motion phase is associated with spontaneous breaking of rotational symmetry, it is natural to identify an order parameter

$$\mathbf{v} = \frac{1}{N} \sum_i \mathbf{u}_i. \quad (16)$$

The particle orientations  $\mathbf{u}_i$  evolve independently of their positions, so the steady state distribution of  $\mathbf{v}$  is simply the distribution of the average of  $N$  random unit vectors. It is convenient to define also the time average of the modulus of the order parameter:

$$\bar{v}_\tau = \frac{1}{\tau} \int_0^\tau |\mathbf{v}(t)| dt. \quad (17)$$

For large times one has  $\langle \bar{v}_\tau \rangle_s = \langle |\mathbf{v}| \rangle_s$ .



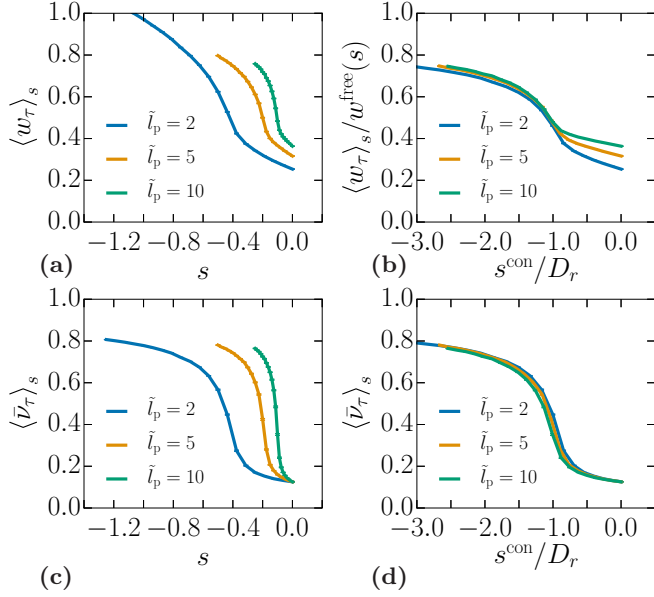


FIG. 3. (a) Biased average of the active work  $\langle w_\tau \rangle_s$ . (b) The data from (a) plotted using rescaled variables from (23) and (25), leading to data collapse. (c) Biased average of the polarization norm  $\langle |\mathbf{v}_\tau \rangle_s$  [see Eq. (17)] as a function of the rescaled biasing parameter  $s^{\text{con}}/D_r$  [see Eq. (25)]. (d) The data from (c) plotted using rescaled variables, showing data collapse. Parameter values (all panels):  $N = 50$ ,  $\phi = 0.65$ ,  $n_c = 10^3$ ,  $t_{\text{max}} = 10^3$ .

For large  $N$ , the central limit theorem means that  $p(\mathbf{v}) \rightarrow (N/\pi)e^{-N|\mathbf{v}|^2}$  (in distribution) and hence

$$\langle \bar{v}_\tau \rangle_s \simeq \frac{1}{2} \sqrt{\frac{\pi}{N}} \quad (18)$$

which tends to zero as  $N \rightarrow \infty$ . On the other hand, symmetry-broken states have

$$\langle \bar{v}_\tau \rangle_s = O(1) \quad (19)$$

as  $N \rightarrow \infty$ .

As  $N \rightarrow \infty$ , the limiting value of  $\langle \bar{v}_\tau \rangle_s$  is zero throughout the isotropic phase, but nonzero in the collective motion phase. This leads to a singularity at the transition point  $-s^*$ , as expected for an order parameter. However, in finite systems, the quantity  $\langle |\mathbf{v}_\tau \rangle_s$  is always positive and has a smooth (analytic) dependence on the field  $s$ . To identify the phase transition in numerical studies, we use that the finite-size scaling behavior (18) and (19) is different in the two phases.

### B. Results for $\tilde{l}_p \geq 2$

Recall that the persistence length  $\tilde{l}_p$  measures the strength of the active self-propulsion, compared to passive diffusion. Figure 3 shows results obtained by cloning for  $\tilde{l}_p \geq 2$ , where the active propulsion is significant. Note that since we focus throughout on  $s < 0$ , the point corresponding to the unbiased (natural) dynamics is on the right of the graphs and the strength of the bias increases from right to left. As the bias becomes more negative, both the active work and the orientational order parameter increase slowly at first, before

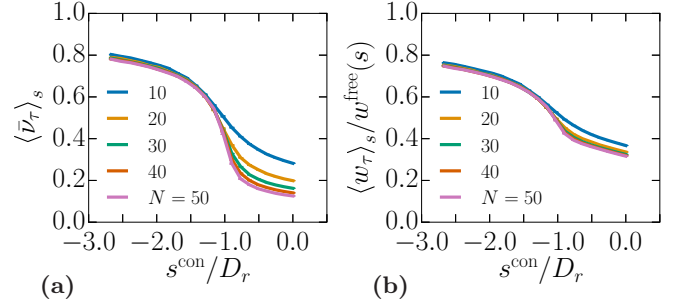


FIG. 4. Biased averages of the active work  $\langle w \rangle_s$  (a) and of the polarization norm  $\langle |\mathbf{v}| \rangle_s$  (b) as functions of the rescaled biasing parameter  $s^{\text{con}}/D_r$  [see Eq. (25)] for number of particles  $N = 10, 20, 30, 40, 50$ . Parameter values:  $\tilde{l}_p = 5$ ,  $\phi = 0.65$ ,  $n_c = 10^3$ ,  $t_{\text{max}} = 10^3$ .

showing a more rapid increase. Similar to (15),

$$w'(s) = - \lim_{\tau \rightarrow \infty} \tau N [\langle w_\tau^2 \rangle_s - \langle w_\tau \rangle_s^2] \quad (20)$$

so a rapid change in  $w(s)$  corresponds to a large variance in the biased ensemble. The analogy with thermodynamic phase transitions suggests that the critical point  $s = -s^*$  coincides with the point where  $|w'(s)|$  is maximal. Figures 3(b) and 3(d) show that appropriate changes of variable can be used to collapse the data for different  $\tilde{l}_p$  (details are given just below, in Sec. III C).

Figure 4 shows the dependence on system size. For  $-s^* < s < 0$ , the order parameter decreases with  $N$  as in Eq. (18), while for  $s < -s^*$  it depends weakly on  $N$  as in Eq. (19). Together with the large variance (20), this justifies our identification of  $s = -s^*$  as a critical point at which symmetry is spontaneously broken. The dependence of  $w(s)$  on  $N$  is weaker; this function should be continuous at a critical point, with a singularity in its derivative at  $-s^*$ ; this is consistent with the data.

### C. Enhancement of self-propulsion in biased ensembles

We now discuss the reasons for the data collapse observed in Figs. 3(b) and 3(d). Note that there are two nontrivial contributions to  $w_\tau$  in Eq. (7), which affect the biased ensemble (12) in different ways. To separate their effects, observe that the average (12) in the biased ensemble may be reformulated as

$$\langle \mathcal{A} \rangle_s = \frac{\langle \mathcal{A} e^{-sN\tau w_{f,\tau}} \rangle_{v_s^{\text{con}}}}{\langle e^{-sN\tau w_{f,\tau}} \rangle_{v_s^{\text{con}}}}, \quad (21)$$

where  $w_{f,\tau}$  was defined in Eq. (8); the averages on the right hand side are computed for the natural dynamics of a controlled ABP system in which the velocity  $v_0$  in Eq. (2) is replaced by

$$v_s^{\text{con}} = v_0 \left( 1 - \frac{2sD}{v_0^2} \right). \quad (22)$$

This result was noted previously in Refs. [37,42]. To highlight connections between optimal-control theory and large-deviation theory [49,50,59], we refer generically to systems with modified equations of motion as *controlled systems*

(see also Sec. V A). In this case, the only modification is the change in self-propulsion velocity, from  $v_0$  to  $v^{\text{con}}$ .

Equation (21) is an exact equality, even for finite  $\tau$ . This is explained in Appendix A by considering the biased time-evolution operators corresponding to (12) and (21). Comparing (12) and (21), the equations of motion of the system have been modified, and the contribution  $w_{\eta,\tau}$  has been removed from the exponential biasing factor.

Noninteracting ABPs have  $w_{f,\tau} = 0$ , in which case this construction allows a full solution of the large-deviation problem. The average (21) in the biased ensemble reduces to an unbiased steady state average for ABPs where only the self-propulsion velocity is modified, leading to an active work

$$w^{\text{free}}(s) = 1 - \frac{2sD}{v_0^2}. \quad (23)$$

The linearity of this function means that the SCGF and the rate function are quadratic: there are no collisions so the only fluctuations of  $w_\tau$  come from the (Gaussian) noise, via (9).

We now consider the effect of interactions in the modified system of (21), with propulsion velocity  $v^{\text{con}}$ . The normalized work  $w_{f,\tau}$  is not a particularly natural quantity in the modified system because it contains a normalization factor  $v_0$  from (8). It is more natural to rescale  $w_f$  by  $w^{\text{free}}$ : the quantity that appears in the exponent of (21) is then

$$s w_{f,\tau} = s^{\text{con}} \frac{w_{f,\tau}}{w^{\text{free}}} \quad (24)$$

with

$$s^{\text{con}} = s \left( 1 - \frac{2sD}{v_0^2} \right). \quad (25)$$

The field  $s^{\text{con}}$  is conjugate to the normalized work ( $w_{f,\tau}/w^{\text{free}}$ ): it is the natural biasing parameter for the controlled system. Both  $s$  and  $s^{\text{con}}$  have dimensions of inverse time.

The data collapse in Figs. 3(b) and 3(d) is obtained by plotting the normalized work  $\langle w_{f,\tau} \rangle_s / w^{\text{free}}(s)$  against its conjugate variable  $s^{\text{con}}$ , rescaled by  $D_r$ . The behavior of (21) can thus be captured by the single parameter  $s^{\text{con}}/D_r$  for  $\tilde{l}_p \gtrsim 2$ , but it also depends separately on  $\tilde{l}_p$  when that parameter is smaller, as discussed in Sec. III D. Hence, the decomposition of  $w_\tau$  as (7) allows its large deviations to be analyzed as a combination of two factors: the bias acts on  $w_{\eta,\tau}$  to increase the self-propulsion; it acts on  $w_{f,\tau}$  to generate alignment and collective motion. The reasons why  $s^{\text{con}}$  should be scaled by  $D_r$  and not by some other rate (for example,  $v_0/\sigma$ ) will be discussed in later sections.

#### D. The case $\tilde{l}_p = 1$

We now turn to cases where the self-propulsion is weaker, corresponding to smaller  $\tilde{l}_p$ . Figure 5 shows results for several system sizes, comparing  $\tilde{l}_p = 1$  with  $\tilde{l}_p = 5$ . The order parameter  $\langle \bar{v} \rangle_s$  no longer collapses perfectly but the qualitative behavior is the same as for larger  $\tilde{l}_p$ . In particular the collective motion transition is robust. However, the active work  $\langle w_\tau \rangle_s$  no longer collapses as a function of  $s^{\text{con}}/D_r$ : there are significant deviations. In other words, the dependence of (21) can no longer be captured by the single dimensionless parameter

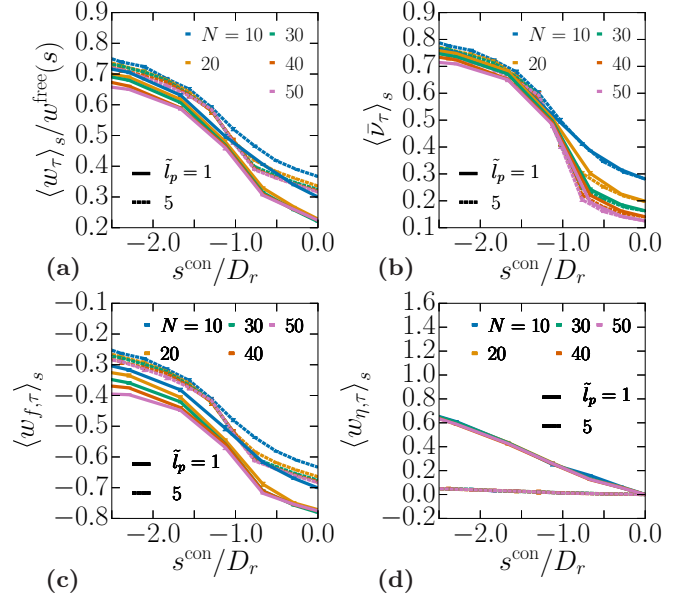


FIG. 5. (a) Biased average of the active work rescaled by the biased free particle active work [see Eq. (23)]. (b) Biased average of the polarization norm  $\langle \bar{v}_\tau \rangle_s$  [see Eq. (17)]. (c) Biased average of the force part of the active work. (d) Biased average of the noise part of the active work. Parameter values:  $\phi = 0.65$ ,  $n_c = 10^2$ ,  $t_{\text{max}} = 10^2$ .

$s^{\text{con}}/D_r$ , but it depends also on  $\tilde{l}_p$  when that parameter is (relatively) small.

Physically, this can be rationalized by considering two mechanisms for events with large  $w_{f,\tau}$ ; in practice, these are events where particles spend less time in contact. The first mechanism is collective motion: if particles all move with fixed speed  $v^{\text{con}}$  in the same direction, they never collide, as discussed in Sec. III A. A second mechanism is isotropic: particles move in random directions, but tend to avoid each other when they get close. When  $\tilde{l}_p$  is large, then the self-propulsion velocity dominates particles' relative velocity, and the collective motion mechanism dominates. This is the regime where the data collapses in Fig. 3. We argue that the isotropic mechanism is becoming important for smaller  $\tilde{l}_p$ , at which point the response to the bias  $s$  acquires a more complicated dependence on  $s/D_r$  and  $\tilde{l}_p$ .

Note that the isotropic mechanism does not rely on active self-propulsion and can be relevant in passive systems, when considering large deviations of quantities like  $w_{f,\tau}$ . On the other hand, the collective motion mechanism is inherent to active systems. It is therefore not surprising that the isotropic mechanism becomes more important for small  $\tilde{l}_p$ , where the system is behaving more like a passive fluid. The limit  $\tilde{l}_p \rightarrow 0$  would be an interesting direction for future study; our expectation is that collective motion still appears for  $s^{\text{con}} \lesssim -D_r$ , but a more detailed understanding of the isotropic mechanism would be required, in order to establish this.

In addition, Fig. 5(d) shows the noise contribution to the active work  $w_{\eta,\tau}$ . For  $\tilde{l}_p = 1$ , this contribution to  $w_\tau$  is comparable to the force part  $w_{f,\tau}$  while for larger  $\tilde{l}_p$ , the force contribution  $w_{f,\tau}$  is the larger contribution. From (23), the noise part is of order  $-sD/v_0^2$ ; taking  $s \sim -D_r$  then this scales

as  $\tilde{l}_p^{-2}$  and is indeed small in the active limit of large  $\tilde{l}_p$ . The force part  $w_{f,\tau}$  is of order unity in this limit.

#### IV. SPONTANEOUS ALIGNMENT OF TWO RTPS

As a preliminary step before describing a detailed theory of collective motion, we illustrate the mechanism for collective motion by an analytic computation of large deviations of the active work, for two RTPs on a one-dimensional periodic ring. Since the system is finite, there cannot be any spontaneous symmetry breaking, but we do find that the particles tend to align their orientations when biased to large active work. An outline of the large-deviation calculations can be found in Appendix C. Analysis of the unbiased behavior of the system can be found in Ref. [67].

The RTPs have positions  $r_i$  for  $i = 1, 2$ , with periodic boundaries, in a domain of size  $L$ . Their active self-propulsion velocities are  $\alpha_i v_0$  with  $\alpha_i = \pm 1$ . Particle  $i$  tumbles with rate  $\tau_p^{-1}$ , which corresponds to  $\alpha_i$  changing its sign, and we introduce  $l = v_0 \tau_p$  the persistence length. The particles interact by a pair potential  $V$ , and there is no thermal diffusion. Let the particle separation be  $r_{12} = |r_1 - r_2|$ . Hence, the equation of motion (between tumbles) is

$$\dot{r}_i = \alpha_i v_0 - \frac{\partial}{\partial r_i} V(r_{12}). \quad (26)$$

The potential  $V$  is short ranged with  $V(r_{12}) \rightarrow \infty$  as  $r_{12} \rightarrow 0$ , and a length scale  $\epsilon$  such that  $V(r_{12}) = 0$  for  $r_{12} > \epsilon$ . For particles in contact, there is a particular distance  $r^*$  (less than  $\epsilon$ ) such that  $V'(r^*) = -v_0$ , so that two particles with opposite orientations can have a force-balanced state with  $\dot{r}_i = 0$ . We focus on the limit of hard particles such that  $\epsilon \rightarrow 0$  and also  $r^* \rightarrow 0$ .

The unnormalized instantaneous rate of active work is defined analogous to (8) as

$$\dot{w}_f^{\text{RTP}} = v_0(\alpha_1 - \alpha_2) \frac{\partial}{\partial r_1} V(r_{12}), \quad (27)$$

$$w_f^{\text{RTP}} = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau \dot{w}_f^{\text{RTP}}(t) dt, \quad (28)$$

while there is no analog of the noise term  $w_\eta$  because there is no thermal noise. In the limit of hard particles one has simply that  $\dot{w}_f^{\text{RTP}} = -2v_0^2$  if the particles are in a force-balanced touching state (for example,  $r_{12} = 0$  with  $\alpha_1 = 1 = -\alpha_2$ ) and  $\dot{w}_f^{\text{RTP}} = 0$  otherwise. We denote by  $\lambda$  the biasing field conjugate to the active work, and introduce the cumulant generating function

$$\psi^{\text{RTP}}(\lambda) = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \ln \langle e^{-\lambda \int_0^\tau \dot{w}_f^{\text{RTP}}(t) dt} \rangle \quad (29)$$

such that  $\langle w_f^{\text{RTP}} \rangle_\lambda = -\partial_\lambda \psi^{\text{RTP}}(\lambda)$  as usual. Here,  $\lambda$  is playing the role of  $s$  in the ABP system.

This quantity can be obtained by solving an eigenproblem, as discussed in Appendix C. Figure 6(a) shows that  $\psi^{\text{RTP}}$  converges to a constant when  $\lambda \rightarrow -\infty$ , therefore,  $\lim_{\lambda \rightarrow -\infty} \langle w_f^{\text{RTP}} \rangle_\lambda = 0$ , indicating that collisions are completely suppressed in this regime.

The analog of (17) in this system is

$$\bar{v}_\tau^{\text{RTP}} = \frac{1}{\tau} \int_0^\tau \frac{1 + \alpha_1(t)\alpha_2(t)}{2} dt, \quad (30)$$

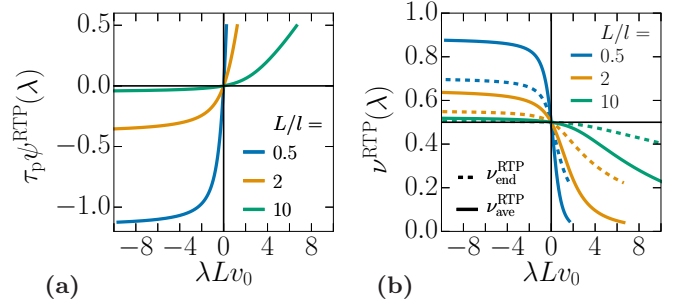


FIG. 6. (a) Rescaled CGF  $\tilde{\psi}^{\text{RTP}} = \tau_p \psi^{\text{RTP}}$  from (C35) as a function of the rescaled biasing parameter  $\tilde{\lambda} = \lambda l v_0$ , for different values of the rescaled ring length  $\tilde{L} = L/l$ . (b) Polarization  $v_{\text{ave}}^{\text{RTP}}$  as defined in Eq. (31), showing particle alignment for  $\tilde{\lambda} < 0$ . The dashed lines show  $v_{\text{end}}^{\text{RTP}}$ , which is evaluated at the end of the trajectory instead of as an average over all times [see also Eqs. (32) and (C64)].

which is between 0 and 1, with an average value of  $\frac{1}{2}$  in the unbiased state where the  $\alpha$ 's are independent. Its average value in the biased ensemble is

$$v_{\text{ave}}^{\text{RTP}}(\lambda) = \lim_{\tau \rightarrow \infty} \langle \bar{v}_\tau^{\text{RTP}} \rangle_\lambda. \quad (31)$$

Note that this quantity is averaged over the whole trajectory, and its determination requires one to consider both the left and right eigenvectors of the associated eigenproblem. The computation is described in Appendix C, which also considers the quantity

$$v_{\text{end}}^{\text{RTP}} = \left\langle \frac{1 + \alpha_1(\tau)\alpha_2(\tau)}{2} \right\rangle_\lambda \quad (32)$$

which measures degree of alignment at the final time  $\tau$ .

Figure 6(b) shows results. Starting from the zero-bias state where all configurations are equiprobable ( $v^{\text{RTP}} = \frac{1}{2}$ ), the polarization increases as  $\lambda$  is reduced from zero, corresponding to large positive fluctuations of the active work. For large negative  $\lambda$ , the polarization eventually reaches a plateau value. At fixed persistence length, the larger systems have weaker polarization. On the contrary, the polarization decreases for positive  $\lambda$ , indicating that antialigned states are more probable than aligned states, so particles spend more time in collision. As usual, the time-averaged measurement  $v_{\text{ave}}^{\text{RTP}}$  responds more strongly to the bias than the corresponding measurement  $v_{\text{end}}^{\text{RTP}}$  at the final time [52].

The main conclusion from this analysis is that a one-dimensional system of two self-propelled particles already shows that biasing towards fewer collisions promotes the alignment of the particles' orientations. We also describe in Appendix C 3 a scaling regime that is relevant when the system is very large, which allows some simplification of the resulting expressions.

#### V. LARGE-DEVIATION MECHANISM

We now present an analysis of the mechanism by which collective motion occurs in the system of many interacting ABPs, as in Figs. 2–5. For two RTPs, we have seen that alignment is a natural mechanism for suppressing collisions between particles. The same is true for ABPs. To characterize

the collective motion state, we compare the biased ensemble (12) with other ensembles that we define either by modifying the equations of motion of the system (via control forces, Secs. VA and VB), or by applying different kinds of bias to the system (Sec. VC). Numerical results are presented in Sec. VD.

We show that controlling the rotational dynamics of the system is enough to qualitatively reproduce the active work rate function in the collective motion state (Sec. VE). This suggests that the fluctuations of the active work are strongly coupled to those of the orientational order parameter in this regime. However, this control fails at capturing these fluctuations in the isotropic state (Sec. VF), calling for an alternative mechanism in this domain. Finally, we provide additional evidence that the transition to collective motion is at  $s = -s^* \approx -D_r$  (Sec. VG).

### A. Control forces

Within large deviation theory, it is often useful to compare biased ensembles like (12) with ensembles where the dynamics of the system is modified by control forces [49,50,59]. In principle, biased ensembles can be reproduced exactly by a suitable (optimally controlled) system. However, these optimal control forces cannot usually be derived exactly, except in idealized models.

As a generic controlled ABP system we take equations of motion

$$\begin{aligned} \dot{\mathbf{r}}_i &= v^{\text{con}} \mathbf{u}_i - D \nabla_i (U + U^{\text{con}}) + \sqrt{2D} \boldsymbol{\eta}_i, \\ \dot{\theta}_i &= -D_r \frac{\partial U^{\text{con}}}{\partial \theta_i} + \sqrt{2D_r} \xi_i, \end{aligned} \quad (33)$$

where  $U^{\text{con}}$  is a control potential (dependent on all particle positions and orientations) and  $v^{\text{con}}$  is a parameter (independent of positions and orientations).

For ABPs in the biased ensemble (12), we explain in Appendix A that the optimally controlled system has

$$\begin{aligned} v^{\text{con}} &= v_s^{\text{con}}, \\ U^{\text{con}} &= U_s^{\text{opt}}, \end{aligned} \quad (34)$$

where  $v_s^{\text{con}}$  was defined in Eq. (22) and  $U_s^{\text{opt}}$  must be determined from the solution to an eigenproblem [see (A2) and (A3)]. Note that the equations of motion for the optimally controlled process have exactly the same random noise terms as the original process (2); this is a general feature of large deviations in this class of system [68].

We consider several controlled systems where the orientational dynamics of ABPs is modified by long-ranged coupling that favors particle alignment. We test numerically how well they capture the properties of the biased ensemble (12), and hence the true large-deviation mechanism. This is achieved by deriving bounds on the rate function  $I(w)$ . The bounds would be exact equalities if our approximations for the optimally controlled dynamics were exact. In fact, the bounds are close but not exact; from this we conclude that our approximations capture important features of the large-deviation mechanism, but they also miss important parts of the physics. In contrast to this work, the bound considered in Ref. [42] is restricted

to  $U^{\text{con}} = 0$ , in which case the control forces only affect the self-propulsion velocity  $v^{\text{con}}$ .

To construct bounds, let  $P$  be the path probability measure for the ABPs, and let  $P^{\text{con}}$  be the corresponding measure for the system with control forces. Averages in the controlled system are denoted  $\langle \dots \rangle_{\text{con}}$ . Then, for any (ergodic) controlled system with  $\langle w_\tau \rangle_{\text{con}} = w$  we have

$$I(w) \leq \lim_{\tau \rightarrow \infty} \frac{1}{N\tau} \mathcal{D}_{\text{KL}}(P^{\text{con}} || P), \quad (35)$$

where  $\mathcal{D}_{\text{KL}}(Q || P)$  is the Kullback-Leibler (KL) divergence between distributions  $P$  and  $Q$  [see Eq. (A5)]. In the analogy between large-deviation theory and thermodynamics, (35) corresponds to the Gibbs-Bogoliubov inequality [50].

As a general controlled system we take (33). In this case, the KL divergence can be computed [see (A7) and (A8)]. The key point is that if the control forces are optimal, then (35) is an equality. We denote the path probability distribution for this optimally controlled system by  $P_s^{\text{opt}}$  so

$$I(w(s)) = \lim_{\tau \rightarrow \infty} \frac{1}{N\tau} \mathcal{D}_{\text{KL}}(P_s^{\text{opt}} || P). \quad (36)$$

Averages with respect to  $P_s^{\text{opt}}$  also match averages in the biased ensemble, that is,  $\langle \mathcal{A} \rangle_s \simeq \langle \mathcal{A} \rangle_{\text{opt}}$  [see Eq. (12)], as  $\tau \rightarrow \infty$ .

As a general rule, the closer is the controlled system to the true fluctuation mechanism, the more accurate will be the bound (35). The intuition is that choosing a controlled process corresponds to proposing a mechanism for the rare event, and this mechanism can occur with a particular probability of order  $\exp[-\mathcal{D}_{\text{KL}}(P^{\text{con}} || P)]$ . Hence, smaller values of  $\mathcal{D}_{\text{KL}}$  correspond to mechanisms that are exponentially more likely, and the mechanism that minimizes  $\mathcal{D}_{\text{KL}}$  is an accurate representation of the rare event.

### B. Coupling among ABP orientations and upper bound on rate function

As already discussed in Ref. [37], the spontaneous breaking of symmetry for  $s < 0$  leads to a natural comparison with systems where torques act on the ABPs, so that their orientations tend to align. This phenomenon is reminiscent of the flocking observed in systems of Vicsek particles [14], when the orientational coupling between neighboring particles exceeds a threshold set by the effective temperature of the rotational dynamics. To explore this effect, take as control potential an infinite-ranged (mean-field) coupling among the orientations, with strength  $g > 0$  corresponding to a ferromagnetic interaction. That is,

$$U_g^{\text{con}} = -\frac{gN}{D_r} |\mathbf{v}|^2 \quad (37)$$

independent of particle positions. The direction of the order parameter is determined by an angle  $\varphi$  through  $\mathbf{v} = |\mathbf{v}|(\cos \varphi, \sin \varphi)$ . The equation for the ABP orientation in the controlled system can then be written as

$$\dot{\theta}_i = -g|\mathbf{v}| \sin(\theta_i - \varphi) + \sqrt{2D_r} \xi_i. \quad (38)$$

Similar to the original ABPs, this orientational equation of motion is independent of all particle positions. Hence,



integrating out the particle positions leads to a mean-field system of interacting rotors, which is fully described by (38). For large  $N$ , Appendix D 1 shows that this system spontaneously breaks rotational symmetry at  $g = D_r$ . That is, for  $g > D_r$ , then  $\langle |\mathbf{v}| \rangle_{\text{con}} = O(1)$  as  $N \rightarrow \infty$ , but  $\langle |\mathbf{v}| \rangle_{\text{con}} = O(N^{-1/2})$  for  $g < D_r$ . The resulting situation is similar to (19) and (18).

It was argued in Ref. [37] that this very simple controlled model can already capture quite accurately the collective motion phase, and it can predict the rate function in this regime. To explore this idea in more detail, we write  $P_g^{\text{con}}$  for the path probability distribution for the ABP dynamics with control potential  $U_g^{\text{con}}$ . Using this distribution with Eq. (35) yields an upper bound

$$I(w_g^{\text{con}}) \leq \lim_{\tau \rightarrow \infty} \frac{1}{N\tau} \mathcal{D}_{\text{KL}}(P_g^{\text{con}} || P), \quad (39)$$

where  $w_g^{\text{con}} = \langle w_\tau \rangle_{\text{con}}$  with control potential  $U_g^{\text{con}}$ . This bound would be an equality if the controlled model fully captured the collective motion phase. It will be compared with the exact result in Sec. V D.

### C. Large deviations of order parameter and lower bound on rate function

We also derive a lower bound on the rate function, similar to [37]. To achieve this, we consider large deviations of the orientational order parameter. Specifically, we consider  $\bar{v}_\tau$ , defined in Eq. (17) as the time average of the modulus of the order parameter. The statistical properties of other similar quantities (for example, the modulus of the time average) have different dependence on  $N$ ,  $\tau$ , so some care is required in the following arguments.

The quantity  $\bar{v}_\tau$  obeys a large-deviation principle as  $\tau \rightarrow \infty$  which we write as

$$p(\bar{v}_\tau) \asymp \exp[-\tau N \mathcal{J}_1(\bar{v}_\tau)], \quad (40)$$

where  $\mathcal{J}_1$  is the rate function. For large  $N$ , the function  $\mathcal{J}_1$  can be obtained analytically (see Appendix D). In particular, one has for large  $N$  and small  $\bar{v}$  that

$$\lim_{N \rightarrow \infty} \mathcal{J}_1(\bar{v}) = \frac{1}{2} D_r \bar{v}^2 + O(\bar{v}^4). \quad (41)$$

Moreover, in this joint limit of large  $N$  and small  $\bar{v}$ , the optimally controlled dynamics associated with these large deviations can be captured exactly by (38). The large deviations of  $\bar{v}$  also satisfy a (general) bound analogous to (35), which is

$$\mathcal{J}_1(\bar{v}) \leq \lim_{\tau \rightarrow \infty} \frac{1}{N\tau} \mathcal{D}_{\text{KL}}(P^{\text{con}} || P) \quad (42)$$

which holds for any controlled dynamics with  $\langle |\mathbf{v}| \rangle_{\text{con}} = \bar{v}$ .

These results can be used to obtain a bound on  $I(w)$ . We write  $v(s) = \langle |\mathbf{v}| \rangle_s$ . Now consider (42) with  $P^{\text{con}} = P_s^{\text{opt}}$ , as the optimally controlled dynamics for large deviations of the active work, as in Eq. (36). This optimally controlled dynamics has  $\langle w_\tau \rangle_{\text{con}} = w(s)$  and  $\langle |\mathbf{v}| \rangle_{\text{con}} = v(s)$ . Combining (42) with (36), we obtain

$$I(w(s)) \geq \mathcal{J}_1(v(s)). \quad (43)$$

This is a lower bound on  $I(w)$ , which was derived in Ref. [37] by the contraction principle of large deviations. From (42), the bound (43) is exact if the optimally controlled dynamics

for large deviations of  $w$  (that is  $P_s^{\text{opt}}$ ) is also an optimally controlled dynamics for large deviations of  $\bar{v}$ .

Note that evaluation of (43) requires knowledge of  $v(s)$ , which is not available analytically; instead, one must perform cloning simulations. For this reason, (43) is not a predictive result. However, it is a useful result because the accuracy of the bound reveals the extent to which the (unknown) mechanism for large deviations of  $w$  is similar to the (known) mechanism for large deviations of  $\bar{v}$ , as we now discuss.

### D. Numerical evaluation of bounds

Figure 7 shows results for the rate function  $I(w)$  (Sec. II C), compared with the upper and lower bounds (39) and (43). We show data for  $\tilde{l}_p = 5$  as well as  $\tilde{l}_p = 40$ , which was the case considered in Ref. [37].

To evaluate the upper bound, we perform unbiased simulations of ABPs with the rotational equation of motion given by Eq. (38), over a range of torque parameters  $g$ . We compute  $\mathcal{D}_{\text{KL}}(P_g^{\text{con}} || P)$  using (A10) and the average active work  $\langle w_\tau \rangle_{\text{con}}$ . The upper bounds in Fig. 7 are parametric plots using these data (with  $g$  as the parameter).

For the lower bound, we compute  $\mathcal{J}_1(\bar{v})$  from cloning simulations of rotors (see Appendix D 3) and  $v(s)$  from cloning simulations of ABPs, then compose these functions to obtain the right hand side of Eq. (43). We stress that the number of particles  $N$  and the rotational diffusivity  $D_r$  have to be consistent between simulations for this comparison to hold.

The bounds of Fig. 7 capture the main features of the rate function but there are significant deviations, especially for smaller  $\tilde{l}_p$ . We note in particular that (i) the bounds (39) and (43) are not accurate for small values of  $w - \langle w \rangle_0$  but become accurate for larger  $w$ ; (ii) the bounds are more accurate for larger  $\tilde{l}_p$ . We now discuss these observations.

### E. Collective motion (symmetry-broken) state

From Fig. 3, the system is in the collective motion phase for  $w > w^*$  with  $w^* \approx 0.4$ , weakly dependent on  $\tilde{l}_p$ . For  $\tilde{l}_p = 5$ , the collective motion phase is  $w - \langle w_\tau \rangle \gtrsim 0.15$ . In this range, the inset of Fig. 7 shows that upper and lower bounds are accurate to around 20% of the rate function. For  $\tilde{l}_p = 40$ , the accuracy of the lower bound is even better.

In both cases, the lower bound is more accurate. Part of this effect may be attributed to the fact that the upper bound does not account for the enhancement of self-propulsion in the biased ensemble (Sec. III C). The lower bound does account (at least partly) for this effect since it uses the functions  $w(s)$  and  $v(s)$  as computed in cloning simulations. The upper bound could be improved by including the controlled velocity  $v^{\text{con}}$  as an additional variational parameter in Eq. (35) and optimizing it numerically, similar to [42,69]. However, it is sufficient for our argument to keep  $v^{\text{con}} = v_0$ .

The conclusion for this regime is that fluctuations of the active work are strongly coupled to those of the orientational order parameter. As a result, the bounds (39) and (43) can capture the behavior of the rate function almost quantitatively. As noted in Ref. [37], the log-probability of a large fluctuation of  $w_\tau$  is almost the same as that of the corresponding orientational fluctuation.

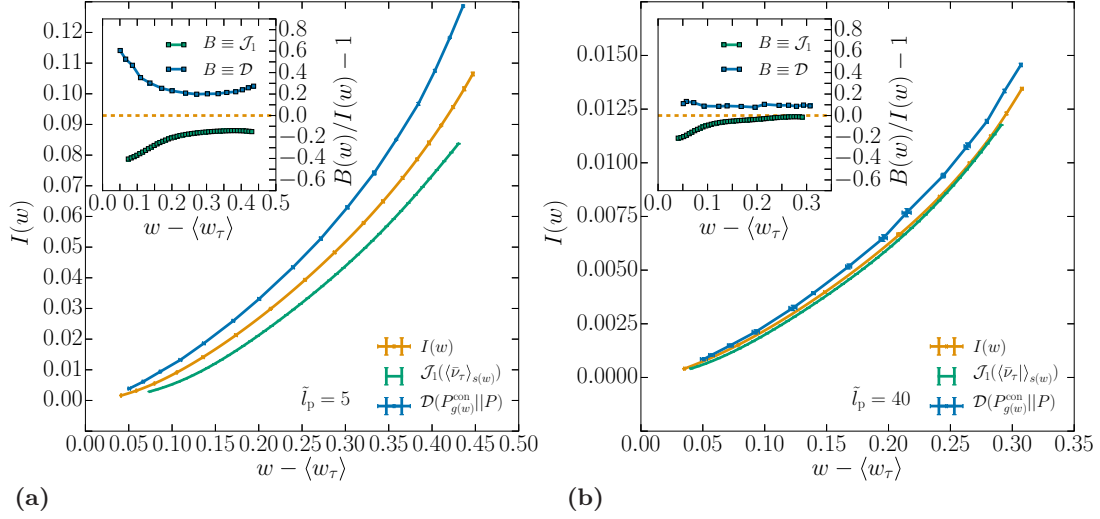


FIG. 7. (a) Rate function  $I(w)$  from the cloning algorithm at  $\tilde{l}_p = 5$ , compared with the upper and lower bounds (39) and (43). For the rate function, the solid line is a spline interpolation. To evaluate bounds, results for  $\langle \bar{v}_\tau \rangle_{s(w)}$  and  $\mathcal{J}_1(\bar{v})$  are obtained by the cloning algorithm, and  $\mathcal{D}(P_{g(w)}^{\text{con}} || P)$  is obtained by simulation of the controlled dynamics. Fourth order polynomial interpolation is then used to obtain values at the desired values of  $w$ . (Inset) Relative error between the bounds and the rate function. (b) Similar results as (a), for  $\tilde{l}_p = 40$ . General parameter values:  $N = 10$ ,  $\phi = 0.65$ . Cloning parameter values:  $n_c = 10^3$ ,  $t_{\text{max}} = 10^3$ ,  $N_{\text{runs}} = 10$ . Parameters for controlled dynamics:  $t_{\text{max}} = 10^4$ ,  $N_{\text{runs}} = 10$ .

### E. Isotropic state $\langle w_\tau \rangle < w < w^*$

When the active work is close to its average value, the system does not break symmetry (recall Fig. 3) and one also sees that the bounds in Fig. 7 do not capture the rate function in an accurate way. The symmetry-breaking transition happens at  $w = w^*$ . As  $N \rightarrow \infty$ , isotropic systems (with  $w < w^*$ ) have  $v(s) = 0$  so the lower bound (43) tends to zero. The rate function is not zero so the bound is not at all accurate in this range. The conclusion for this regime is that fluctuations where  $w_\tau$  is enhanced can also occur by an alternative fluctuation mechanism (not collective motion), and that this mechanism is dominant in the isotropic phase. As noted above, the modified self-propulsion  $v^{\text{con}} = v_s^{\text{con}}$  is relevant in this regime, and particles may also be repelled from each other without any alignment, as discussed in Sec. III D. We explain in Sec. VI that hydrodynamic density fluctuations are also relevant. In other words, several effects act to enhance the active work in this regime, there is no collective motion, and the bounds of Fig. 7 do not follow the rate function accurately.

### G. Comparison with controlled system

To conclude our discussion of orientational fluctuations, we consider Fig. 8 which further illustrates the relationship between controlled dynamics and the large-deviation mechanisms. Figures 8(a) and 8(c) show the dependence of the active work, as either  $s$  or  $g$  is varied. Also, Figs. 8(b), 8(d), and 8(f) show the orientational order parameter.

Note that  $s$  is a biasing field in the sense of large deviations while  $g$  is a physical coupling between ABP orientations; nevertheless, the response to both kinds of perturbation is similar, in that both  $\langle w_\tau \rangle_\lambda$  and  $\langle |\mathbf{v}| \rangle_\lambda$  increase. The responses are different when the system is close to its unbiased steady state, in that the biased ensemble responds mostly by a change

in  $\langle w_\tau \rangle_\lambda$  while the controlled system responds mostly by increasing  $\langle |\mathbf{v}| \rangle_\lambda$ . However, when the system is perturbed further from its steady state, both systems respond by entering a collective motion state, in which their behavior is similar, with long-ranged orientational order and enhanced active work.

Finally, Fig. 8(e) compares how large a coupling  $g$  is required in order to achieve an active work equal to  $w(s)$ . In the collective motion phase, one sees that  $g(w(s)) \approx -s$ . From (D5), rotational symmetry is spontaneously broken for the controlled system with  $g > D_r$ . Hence, Figs. 8(b), 8(d), and 8(f) are consistent with the observation from Fig. 3 that symmetry is broken in the biased ensemble for  $s < -s^*$  with  $s^* \approx D_r$ . A concrete theoretical explanation for this effect will be given in the next section.

## VI. LANDAU-GINZBURG THEORY

We have explained many aspects of the collective motion phase of ABPs in terms of mean-field interactions among their orientations. We now discuss how this mean-field picture can be embedded in a hydrodynamic theory (Sec. VI A), similar to macroscopic fluctuation theory [70]. We show that it predicts spontaneous symmetry breaking for  $s < -s^*$  with  $s^* \approx D_r$  (Sec. VI B), as well as a suppression of density fluctuations for  $s < 0$  (Sec. VI C).

### A. Theory

The minimal description that allows modeling of the symmetry-broken state is to consider a local density field  $\rho$  and a corresponding polarization field  $\mathbf{P}$ . These are defined on

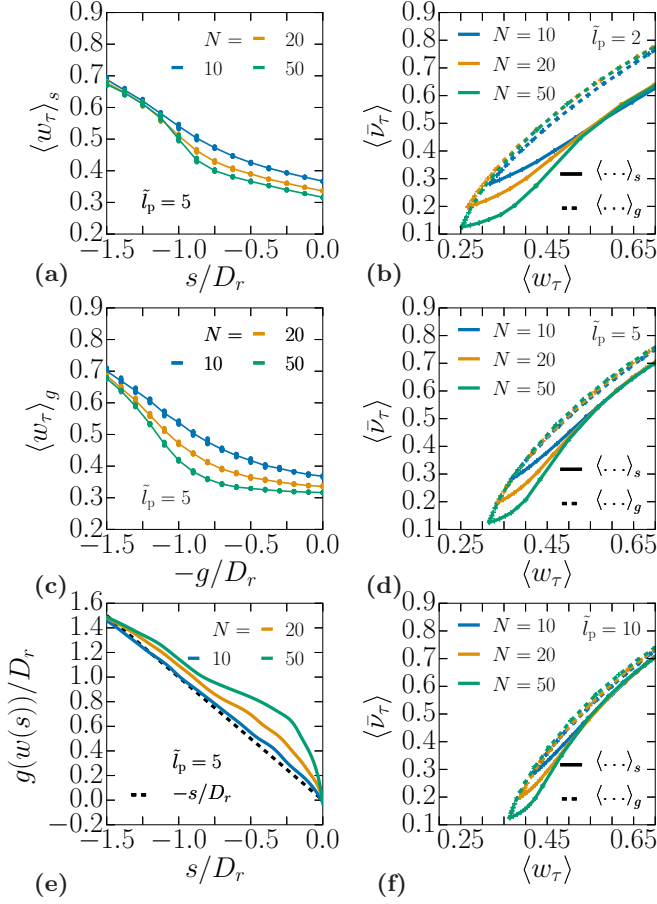


FIG. 8. (a) Active work as a function of the biasing parameter from cloning simulations. Solid lines are spline interpolations. (b) Parametric plot of  $\langle \mathbf{v} \rangle = \langle \frac{1}{\tau} \int_0^\tau dt |\mathbf{v}(t)| \rangle$  as a function of the active work. We show results for biased ensembles  $\langle \dots \rangle_s$  and for controlled dynamics  $\langle \dots \rangle_g$ . (c) Active work as a function of the torque parameter  $w$  in the controlled (modified) dynamics. Solid lines correspond to spline interpolations, points correspond to numerical data. (d) Similar to (b), with  $\tilde{l}_p = 5$ . (e) Composition of the relation  $g(w)$  from modified dynamics and  $w(s)$  from cloning. (f) Similar to (b) and (d), with  $\tilde{l}_p = 10$ . General parameter values:  $\phi = 0.65$ . Cloning parameter values:  $n_c = 10^3$ ,  $t_{\max} = 10^2$ ,  $N_{\text{runs}} = 10$ . Controlled dynamics parameter values:  $t_{\max} = 10^4$ ,  $N_{\text{runs}} = 10$ .

hydrodynamic length scales: given a mesoscopic region  $\Omega_r$  centered at  $\mathbf{r}$ , we define

$$\rho(\mathbf{r}) = \frac{1}{|\Omega_r|} \int_{\Omega_r} \sum_i \delta(\mathbf{r} - \mathbf{r}_i) d\mathbf{r}, \quad (44)$$

$$\mathbf{P}(\mathbf{r}) = \frac{1}{\rho(\mathbf{r})|\Omega_r|} \int_{\Omega_r} \sum_i \mathbf{u}_i \delta(\mathbf{r} - \mathbf{r}_i) d\mathbf{r}, \quad (45)$$

where  $|\Omega_r|$  denotes the volume of  $\Omega_r$ . The polarization  $\mathbf{P}$  is normalized as the average orientation (so  $|\mathbf{P}|^2 < 1$ ).

With this choice, it is notable that  $\rho$  is a slow hydrodynamic field, in the sense that density fluctuations on large length scales  $\ell$  relax on long timescales of order  $\ell^2$ . On the other hand,  $\mathbf{P}$  is a fast field, in that polarization fluctuations on all scales relax quickly (on a timescale of order  $D_r^{-1}$ ) to quasisteady states which depend in general on  $\rho$ .

As a minimal model for ABPs in biased ensembles, we propose (in a “top-down” coarse-grained approach) the following (Itô) equations of motion for  $(\rho, \mathbf{P})$ , similar to [71]:

$$\begin{aligned} \dot{\rho} &= -\nabla \cdot \mathbf{J}, \\ \mathbf{J} &= \mathbf{J}_d + \sqrt{2\sigma(\rho)}\boldsymbol{\eta}, \\ \dot{\mathbf{P}} &= -\gamma(\rho, \mathbf{P})\mathbf{f}(\mathbf{P}) + b(\rho, \mathbf{P})\nabla\rho + \sqrt{2\gamma(\rho, \mathbf{P})}\boldsymbol{\xi}, \end{aligned} \quad (46)$$

where  $\sigma$  and  $\gamma$  are noise strengths;  $b$  is a coupling between polarization and density gradients;  $\mathbf{f}$  is a thermodynamic force that acts on the polarization;  $\mathbf{J}_d$  is the deterministic part of the current; and  $\boldsymbol{\eta}$  is a Gaussian white noise with zero mean and variance  $\langle \eta^\alpha(t)\eta^\beta(t') \rangle = \delta_{\alpha\beta}\delta(t-t')$ . We take a deterministic current

$$\mathbf{J}_d = v_0\rho\mathbf{P} - D_c(\rho)\nabla\rho, \quad (47)$$

where the first term incorporates the effect of self-propulsion, while  $D_c(\rho)$  is the hydrodynamic (collective) diffusion constant, which depends on density. Our theory is restricted to states without MIPS; this requires that  $D_c(\rho) > 0$  for all  $\rho$ . (In fact, the diffusion constant is additionally renormalized by polarization fluctuations, as discussed in Appendix E, and it is the renormalized diffusion constant that should be positive.) In the absence of MIPS, it is consistent to assume that  $\nabla\rho$  is  $O(\ell^{-1})$  where  $\ell$  is the hydrodynamic length scale. This is the reason that higher gradients of  $\rho$  are neglected in Eq. (46).

It is also useful to compare (46) with [72], which is a rigorous hydrodynamic theory for a similar model, except that the polarization is a slow field in that case. In the ABP context, their model definition corresponds to reducing  $v_0$  and  $D_r$  as the system size increases, so that the polarization becomes a hydrodynamic variable. Then, Eqs. (5) and (6) from [72] can be compared with Eq. (46), on identifying their  $m$  as our  $\rho\mathbf{P}$ , and noting that they do not analyze fluctuations about the hydrodynamic limit, so all noise terms are absent. The details of the nonlinear terms differ between the models, but they have a similar structure. The model of [72] corresponds to an idealized limiting case, but it shows how such theories can be justified rigorously. Comparing (46) with the Toner-Tu theory [15] and other theories for collective motion such as [73], our minimal model (46) has fewer couplings, in particular, it lacks advective terms in the polarization equation. Such terms are not relevant at the level considered here, but might be required for a quantitative description of fluctuations in the collective motion phase.

On the hydrodynamic scale, it is consistent to approximate the (total) active work of a trajectory as a function of the density and polarization fields:

$$w_\tau N\tau = \int_0^\tau \int_{[0,L]^2} \omega(\rho, \mathbf{P}) d\mathbf{r} dt, \quad (48)$$

where  $\omega(\rho, \mathbf{P})$  is the typical (average) active work per unit volume, in a region with density  $\rho$  and polarization  $\mathbf{P}$ .

Now consider a large system of linear size  $L$ , and define hydrodynamic coordinates as

$$\tilde{\mathbf{r}} = \mathbf{r}/L, \quad \tilde{t} = Dt/L^2. \quad (49)$$

We also express the fields in these rescaled variables as  $\tilde{\rho}(\tilde{\mathbf{r}}, \tilde{t}) = \rho(\tilde{\mathbf{r}}L, \tilde{t}L^2/D)$  and  $\tilde{\mathbf{P}}(\tilde{\mathbf{r}}, \tilde{t}) = \mathbf{P}(\tilde{\mathbf{r}}L, \tilde{t}L^2/D)$ . [Note,

there is no rescaling of the fields themselves, hence,  $|\tilde{\mathbf{P}}(\tilde{\mathbf{r}}, \tilde{t})|^2 \leq 1$  is the modulus of the local polarization, and  $\tilde{\rho}(\tilde{\mathbf{r}}, \tilde{t})\sigma^2\pi/4$  is a local area fraction, of the same order as the global area fraction  $\phi = \bar{\rho}\sigma^2\pi/4$ .] At the expense of some heavy notation, we consistently use tildes to indicate that the hydrodynamic rescaling has been performed. Where derivatives act on fields with tildes, they are taken with respect to the rescaled variables, for example,  $\dot{\tilde{\rho}} = (\partial\tilde{\rho}/\partial\tilde{t})$ . We define  $\tilde{\mathbf{J}} = \mathbf{J}L/D$  so that the continuity equation retains its form

$$\dot{\tilde{\rho}} = -\nabla \cdot \tilde{\mathbf{J}}. \quad (50)$$

Using (46) and working in hydrodynamic coordinates, we obtain a formula for the probability of a trajectory in the biased ensemble,

$$\text{Prob}[\tilde{\rho}, \tilde{\mathbf{J}}, \tilde{\mathbf{P}}] \propto \exp\left(-L^2 \int_0^{\tilde{\tau}} \int_{[0,1]^2} \mathcal{S}[\tilde{\rho}, \tilde{\mathbf{J}}, \tilde{\mathbf{P}}] d\tilde{t} d\tilde{x}\right) \quad (51)$$

which is valid only if (50) holds (else the probability is zero); the Lagrangian is

$$\mathcal{S} = \frac{D}{4\sigma} |\tilde{\mathbf{J}} - \tilde{\mathbf{J}}_d|^2 + \frac{1}{4D\gamma} \left| \frac{D}{L} \dot{\tilde{\mathbf{P}}} + \gamma fL - b\nabla\tilde{\rho} \right|^2 + \frac{sL^2}{D} \omega, \quad (52)$$

where  $\tilde{\mathbf{J}}_d = \mathbf{J}_d L/D$  and we have omitted the dependence of  $\sigma, \gamma, \omega$  on  $(\tilde{\rho}, \tilde{\mathbf{P}})$ , for ease of writing.

The terms in Eq. (52) that involve  $\mathbf{J}$  and  $\omega$  are familiar from other analyses of large deviations in systems with hydrodynamic modes [60,74]. In particular, the fact that  $s$  enters through the combination  $sL^2$  is familiar from earlier studies [58], it reflects the fact that hydrodynamic degrees of freedom respond strongly to the bias because of their slow relaxation. However, the factors of  $L$  that appear in the term involving  $\tilde{\mathbf{P}}$  are not expected in hydrodynamic theories: they reflect the fact that  $\mathbf{P}$  is a fast field.

### B. Mean-field analysis

We first consider the case where  $(\tilde{\rho}, \tilde{\mathbf{J}}, \tilde{\mathbf{P}})$  are independent of space and time. Hence, we set  $\tilde{\rho} = \bar{\rho}$  where

$$\bar{\rho} = \frac{N}{L^2} \quad (53)$$

is the average density. The action is minimized by taking  $\tilde{\mathbf{J}} = (v_0 L \bar{\rho}/D) \tilde{\mathbf{P}}$ ; note that this is  $O(L)$  when expressed in these hydrodynamic variables because the self-propulsion leads to ballistic motion. Then,

$$\mathcal{S} = L^2 \left[ \frac{\gamma}{4D} |f|^2 + \frac{s\omega(\bar{\rho}, \mathbf{P})}{D} \right]. \quad (54)$$

The next step is to use properties of ABPs to express the remaining quantities in terms of microscopic parameters. Setting  $s = 0$  in Eq. (54) and combining with (51), one can read off the probability of a trajectory where the polarization is fixed at  $\mathbf{P}$  for all times between 0 and  $\tau$ . This same probability can be computed applying large-deviation theory to the microscopic ABP model: the relevant large-deviation principle is given in Eq. (D8), and Appendix D2 explains how the rate function  $\mathcal{J}$  can be computed as the Legendre transform of a

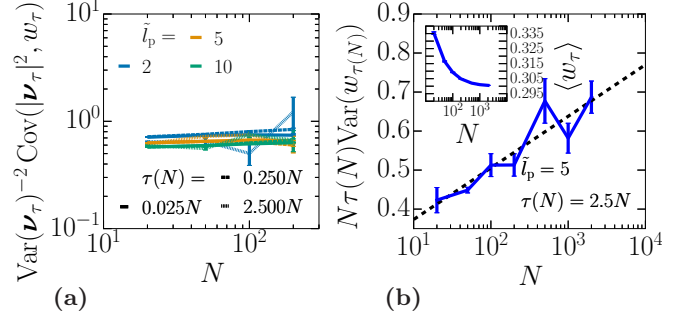


FIG. 9. (a) Normalized covariance of the active work  $w_\tau$  and the squared average polarization  $|\mathbf{v}_\tau|^2 = |\frac{1}{\tau} \int_0^\tau \mathbf{v}(t) dt|^2$ . (b) Normalized variance of the active work with a  $\ln N$  fit in dashed line. Parameter values:  $\phi = 0.65$ ,  $N_{\text{runs}} =$  (a)  $2 \times 10^2$ , (b)  $6 \times 10^2$ .

Mathieu function (see also [35]). Hence,

$$\mathcal{S} = \frac{L^2}{D} [\bar{\rho} \mathcal{J}(\mathbf{P}) + s\omega(\bar{\rho}, \mathbf{P})]. \quad (55)$$

Recalling that  $\omega$  is defined as the active work per unit volume for a system with prescribed density and polarization, we have  $\omega(\bar{\rho}, \mathbf{0}) = \bar{\rho} \langle w_\tau \rangle$ . For  $\mathbf{P} \neq \mathbf{0}$ , the consistent choice is to define  $\omega$  by considering an ensemble of trajectories that is biased by the polarization. Details are given in Appendix D2; we define  $\langle \cdot \rangle_h$  as an average analogous to (12), but with the bias acting on the polarization [see (D10)]. Then,

$$\omega(\bar{\rho}, \mathbf{P}) = \langle \bar{\rho} w_\tau \rangle_{h(\mathbf{P})}, \quad (56)$$

where  $h(\mathbf{P})$  is defined by  $\langle \bar{\mathbf{v}}_\tau \rangle_{h(\mathbf{P})} = \mathbf{P}$ . For small  $\mathbf{P}$ , we show in Appendix D4 that

$$\omega(\bar{\rho}, \mathbf{P}) = \bar{\rho} \left[ \langle w_\tau \rangle + \frac{c_\omega}{2} |\mathbf{P}|^2 + O(|\mathbf{P}|^4) \right], \quad (57)$$

$$c_\omega = \frac{\bar{\rho} \tau^2 L^4 D_r^2}{2} \text{Cov}(w_\tau, |\mathbf{v}_\tau|^2), \quad (58)$$

where  $\text{Cov}(x, y) = \langle xy \rangle - \langle x \rangle \langle y \rangle$  is the covariance in the natural (unbiased) ABP dynamics. Figure 9(a) shows that this covariance is positive. That is, trajectories with larger polarization also tend to have larger active work.

To analyze spontaneous symmetry breaking, we require a Taylor expansion of  $\mathcal{S}$  for small  $\mathbf{P}$ . The behavior of  $\mathcal{J}$  at small polarization can be obtained exactly: the result is (D18), which implies that

$$\mathcal{J}(\mathbf{P}) = \frac{1}{2} D_r |\mathbf{P}|^2 + O(|\mathbf{P}|^4). \quad (59)$$

Hence, (55) becomes

$$\mathcal{S} = \frac{\bar{\rho} L^2}{D} \left[ s \langle w_\tau \rangle + \frac{1}{2} |\mathbf{P}|^2 (D_r + s c_\omega) + O(|\mathbf{P}|^4) \right]. \quad (60)$$

Finally, minimizing the action, we predict spontaneous symmetry breaking for  $s < -s^*$  with  $s^* = D_r/c_\omega$ . From Fig. 9(a), one sees that  $c_\omega$  depends weakly on  $D_r$ , so this is consistent with the observation of Sec. III that  $s^* \propto D_r$ .

Physically,  $c_\omega > 0$  reflects the fact that breaking symmetry reduces collisions between particles and increases the active work. This effect is quadratic in  $\mathbf{P}$ , and so is the rate function  $\mathcal{J}$  associated with symmetry-breaking events. Hence, both contributions in Eq. (60) have the same scaling with  $\mathbf{P}$  (and



with  $L$ ), leading to a critical point  $s^*$  (independent of  $L$ ), for which the coefficients of the two relevant terms balance each other.

Since this result was obtained by a top-down coarse-grained approach, one can expect that it should be quite generic, in that the same theory would be a natural description of other active particles like RTPs. The key ingredient here is that the coefficient  $c_\omega$  in Eq. (57) should be positive, which means by (58) that a larger polarization is correlated (in the unbiased steady state) with a larger active work. In ABPs, there is a clear mechanism for this, that particles' relative velocities are reduced if they align because  $\mathbf{u}_i$  is a unit vector. In other active systems, this would have to be checked on a case-by-case basis.

### C. Fluctuations

This mean-field analysis predicts that symmetry is spontaneously broken for  $s < -s^*$  with  $s^* > 0$ . This means that the system remains isotropic in a finite range around  $s = 0$ . However, the fact that  $s$  enters (52) as  $sL^2$  indicates that the system can respond strongly to the bias already for very small  $s$ , via hydrodynamic fluctuations. As explained in Appendix E, the (fast) polarization field is not relevant on the hydrodynamic scale so it is sufficient for small bias to consider a scalar theory for the density. In this case, density fluctuations and fluctuations of the active work can be understood based on previous studies [60,74].

The behavior depends on the sign of  $\frac{\partial^2}{\partial \bar{\rho}^2} \omega(\bar{\rho}, \mathbf{0})$  which we abbreviate here by  $\bar{\omega}''_0$ . For ABPs then  $\bar{\omega}''_0 < 0$ . The first consequence of the hydrodynamic theory is that the variance of  $w_\tau$  (under the natural dynamics) has a hydrodynamic contribution. Appendix E shows that the asymptotic variance as  $L \rightarrow \infty$  behaves as

$$\bar{\rho} L^2 \lim_{\tau \rightarrow \infty} \tau \text{Var}(w_\tau) = \frac{(\bar{\omega}''_0 \sigma(\bar{\rho}))^2}{4\pi \bar{\rho} D_c(\bar{\rho})^3} [\ln L + O(1)]. \quad (61)$$

Figure 9(b) shows numerical data that are consistent with this prediction.

The hydrodynamic theory of Appendix E also predicts for  $\bar{\omega}''_0 < 0$  that a system described by (46) becomes inhomogeneous (phase separation) for  $s > s_c$  with [42,60,74]

$$s_c L^2 = -\frac{2\pi^2 D_c(\bar{\rho})^2}{\bar{\omega}''_0 \sigma(\bar{\rho})}. \quad (62)$$

Recall that  $\bar{\omega}''_0 < 0$ , which means that  $s_c > 0$ .

For any negative value of  $s$ , one expects the system to be hyperuniform [60,74]. Write  $\rho_q$  for a Fourier component of the density field and  $S_s(q) = \langle \rho_q \rho_{-q} \rangle_s$  for the structure factor in the biased ensemble, where  $q = |q|$ . The structure factor is derived in Eq. (E16) which shows that it behaves for small  $q$  as

$$S_s(q) = \langle \rho_q \rho_{-q} \rangle_s \simeq \begin{cases} \chi_0, & s = 0 \\ b_s q, & s < 0 \end{cases} \quad (63)$$

where  $\chi_0 = \sigma(\bar{\rho})/D_c(\bar{\rho})$  and  $b_s$  are constants [the behavior of  $b_s$  and can be read from (E16)].

Comparison of these theoretical predictions with numerical results requires some care because the numerical results are limited to small systems, while the hydrodynamic theory is

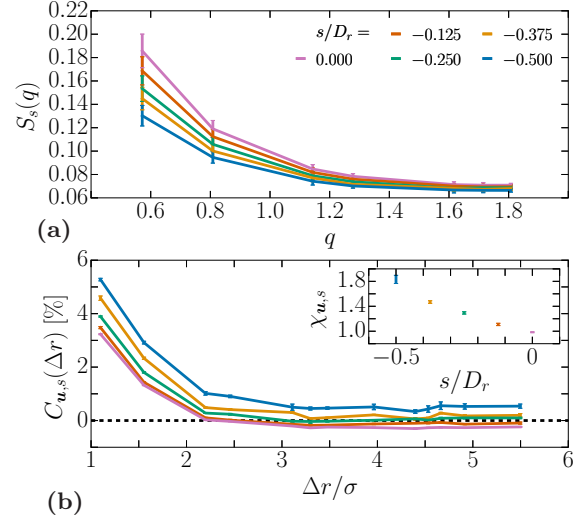


FIG. 10. (a) Structure factor as function of wave vector  $k = 2\pi/\lambda$ , zoomed on highest computed wavelengths, at biasing parameter  $s$ . (b) Main plot: Orientation correlation as a function of the distance at biasing parameter  $s$ . (b) Inset: Sum of the correlation function. Parameter values:  $N = 10^2$ ,  $\bar{l}_p = 5$ ,  $\phi = 0.65$ ,  $n_c = 10^3$ ,  $t_{\max} = 10^2$ .

valid only if the system size is much larger than all microscopic length scales.

The existence of inhomogeneous states for  $s > s_c$  was already discussed in Ref. [37]. For  $s \leq 0$ , Fig. 10(a) shows the behavior of the structure factor in the isotropic phase. For  $s = 0$ , the system is a homogeneous fluid so  $\lim_{q \rightarrow 0} S(q)$  must be a positive constant. However, the numerical results show  $S(q)$  increasing as  $q$  is reduced towards zero. The reason is that the system is not large enough to show the hydrodynamic behavior; in a much larger system, then smaller wave vectors would be accessible, and convergence of  $S(q)$  to its small- $q$  limit would be apparent. Similarly, for  $s < 0$ , one expects in very large systems to observe  $S(q) \propto q$  at small  $q$ , but this is not apparent from our numerical results in small systems. Preliminary data (not shown) indicate that systems with  $N = 10^4$  are sufficient to characterize the small- $q$  limit of  $S(q)$  at  $\bar{l}_p = 5$ , so hyperuniformity should be apparent in systems of that size, if computations were possible in the biased ensemble. However, a clear demonstration of hyperuniformity would likely require still larger systems. Despite these restrictions, the numerical results show that density fluctuations are suppressed in biased ensemble with  $s < 0$ , which is qualitatively consistent with the theory, and with the physical reasoning that reduced density fluctuations tend to suppress collisions and enhance the active work.

Density fluctuations also have consequences for the symmetry-breaking transition at  $s = -s^*$ . First, the transition takes place between a hyperuniform isotropic state and a symmetry-broken state (which may also be hyperuniform). In this case, the simple relationship (48) should be adjusted to account for the fact that the density fluctuations near  $s^*$  are different from those of the biased ensemble (D10). This effect presumably shifts the value of  $s^*$  but we expect qualitative predictions of mean-field theory to remain valid.

An additional question is how density fluctuations couple with those of the polarization, close to the critical point ( $s = -s^*$ ). If one ignores the coupling between  $\mathbf{P}$  and  $\rho$  in Eq. (52), one expects a transition in the universality class of an XY model in  $(2 + 1) = 3$  dimensions (two spatial dimensions and one of time). This situation is familiar in quantum phase transitions for rotors [75].

However, the coupling to a hydrodynamic density field may change this universality class. For example, in the Vicsek model, a locally aligning interaction leads to clustering of the particles. As a result, symmetry breaking for the orientations is coupled with collective motion and with phase separation [14,16]. By contrast, for the phase transition considered here, clustering is suppressed (hyperuniformity) [see Fig. 10(a)]. Also, Fig. 10 shows the spatial orientation correlation function

$$C_{u,s}(\Delta r) = \frac{\int_0^\tau \langle \sum_{i,j=1}^N \mathbf{u}(\theta_i(t)) \cdot \mathbf{u}(\theta_j(t)) \delta_{ij}(\Delta r, t) \rangle_s dt}{\int_0^\tau \langle \sum_{i,j=1}^N \delta_{ij}(\Delta r, t) \rangle_s dt},$$

$$\delta_{ij}(\Delta r, t) = \delta[|\mathbf{r}_i(t) - \mathbf{r}_j(t)| - \Delta r]. \quad (64)$$

This average is performed over trajectories extracted from the cloning algorithm [52]. In these (small) systems, the main effect of the bias on fluctuations in the isotropic phase seems to be a constant (infinite-ranged) additive contribution to  $C$  [see Fig. 10(b)]. This contribution is of order  $1/N$  so that

$$\chi_{u,s} = \int C_{u,s}(r) 2\pi r dr$$

$$= \frac{1}{N} \left\langle \left( \sum_{i=1}^N \mathbf{u}(\theta_i(t)) \right)^2 \right\rangle_{t,s} \quad (65)$$

is of order unity, and exactly equal to 1 in natural dynamics ( $s = 0$ ). Such behavior can be accounted for in mean-field theory: one might expect nontrivial spatial structure to emerge in larger systems, but analysis of such effects is beyond the scope of this work.

## VII. CONCLUSION

We have analyzed large deviations with enhanced active work in an ABP system. As found in Ref. [37], this results in spontaneous breaking of rotational symmetry and collective motion. We presented numerical evidence and theoretical arguments that this transition occurs at  $s = -s^*$  with  $s^* \simeq D_r$ , at least when  $D_r$  is small ( $\tilde{l}_p \gg 1$ ). This means that collective motion sets in above a threshold  $w^*$  for the active work, with  $w^* > \langle w_\tau \rangle$ .

We have compared the behavior of the collective motion state with that of a controlled system where the ABP orientations interact via an infinite-ranged (mean-field) coupling. This controlled system captures the large deviations semi-quantitatively. Based on a hydrodynamic theory, we have also explained that we expect hyperuniform behavior for values of the active work between  $\langle w_\tau \rangle$  and  $w^*$ . We discussed the extent to which the predictions of this theory should be generic in systems of self-propelled particles. As already anticipated in the Introduction, the result that  $w^* \neq \langle w_\tau \rangle$  (and hence  $s^* \neq 0$ ) resolves an open question from [37]. Hence,  $I(w^*) > 0$  too, so [by (10)] the log-probability cost of these large-deviation

events is proportional to  $N\tau$ , in contrast to (much smaller) cost of large-deviation events that originate in diffusive hydrodynamic modes [49,58,60]. It also means that the control forces (or torques) required to stabilize the long-ranged order are significant: that is, the control forces (torques) in Eq. (38) are applied to all particles and have typical magnitudes of order unity, in the thermodynamic limit. In the analogy of [37] between optimal control and evolutionary strategies, it means that alignment cannot be exploited without some cost, but the enhanced self-propulsion efficiency is still significant.

Despite the progress of this work for understanding collective motion in biased ensembles, several questions remain open, including the nature of the critical point where the symmetry is broken. A related point is whether a controlled system with a more complicated (distance-dependent) coupling of orientations would capture the collective motion phase more accurately. Another point of comparison is collective motion in aligning particles, such as the Vicsek model and its variants [14,73,76]. In that case, breaking of rotational symmetry is often accompanied by phase separation into dense (polar) and more dilute (apolar) regions, which leads to a first-order transition to collective motion [16].

For the systems considered here, the evidence [for example, Fig. 10(a)] is that density fluctuations are suppressed in the collective motion phase, contrary to the enhanced density fluctuations that one might expect in systems with polar clusters. However, these numerical results for small systems are not sufficient to settle the nature of the phase transition and its fluctuations. A detailed analysis of symmetry breaking within the Landau-Ginzburg theory of Sec. VI would be an interesting direction for future work.

## ACKNOWLEDGMENTS

The authors acknowledge insightful discussions with T. Nemoto, J. Tailleur, and T. Arnoult de Pirey. This work was funded in part by the European Research Council under the EU's Horizon 2020 Programme, Grant No. 740269. É.F. acknowledges support from an ATTRACT Investigator Grant of the Luxembourg National Research Fund, an Oppenheimer Research Fellowship from the University of Cambridge, and a Junior Research Fellowship from St Catharine's College. M.E.C. is funded by the Royal Society. F.v.W. and Y.E.K. acknowledge the support of THEMA funding from Agence Nationale de la Recherche (ANR). All the codes that were developed and used in this project are made freely available under the MIT license [77].

## APPENDIX A: LARGE DEVIATIONS FOR ABPs

### 1. Eigenvalue problem

Large deviations of  $w_\tau$  can be analyzed through the eigenvalues of an operator called the backwards generator. For compactness of notation, we define (only for this section)  $\tilde{s} = (s/v_0)$ . Using results of [68,78], the eigenvalue equation

is

$$\begin{aligned}\psi(s)\mathcal{F}_s &= D \sum_i (\nabla_i - \tilde{s}\mathbf{u}_i) \cdot (\nabla_i - \tilde{s}\mathbf{u}_i)\mathcal{F}_s \\ &+ \sum_i (v_0\mathbf{u}_i - D\nabla_i U) \cdot (\nabla_i - \tilde{s}\mathbf{u}_i)\mathcal{F}_s \\ &+ D_r \sum_i \frac{\partial^2 \mathcal{F}_s}{\partial \theta_i^2},\end{aligned}\quad (\text{A1})$$

where  $\mathcal{F}_s = \mathcal{F}_s(\mathbf{r}_1, \dots, \mathbf{r}_N, \theta_1, \dots, \theta_N)$  is the eigenvector and  $\psi(s)$  the eigenvalue. This may be simplified as

$$\begin{aligned}\psi(s)\mathcal{F}_s &= \sum_i [(v_0 - 2\tilde{s}D)\mathbf{u}_i - D\nabla_i U] \cdot \nabla_i \mathcal{F}_s \\ &+ \sum_i \left[ D\nabla_i^2 \mathcal{F}_s + D_r \frac{\partial^2 \mathcal{F}_s}{\partial \theta_i^2} \right] \\ &+ \sum_i [\tilde{s}D(\mathbf{u}_i \cdot \nabla_i U) + D\tilde{s}^2 - \tilde{s}/v_0]\mathcal{F}_s,\end{aligned}\quad (\text{A2})$$

where the first two lines correspond to the generator of an ABP system with a modified swim velocity  $v_s^{\text{con}}$  as defined in Eq. (22), and the last line corresponds to a bias in which the only nontrivial term is  $v_0 D(\mathbf{u}_i \cdot \nabla_i U)$ , which is the scalar product between the swim velocity and the interparticle force, as it appears in the  $w_f$  part of the active work.

The operator on the right hand side of (A2) is equal (up to an additive constant) to the operator that generates the biased ensemble on the right hand side of (21). The additive constant affects the eigenvalue but it does not affect the biased ensemble. This is sufficient to establish (21).

From the solution to the eigenproblem, one may construct a corresponding optimally controlled system, whose natural dynamics matches that of the biased ensemble. The corresponding optimal-control potential is

$$U_s^{\text{opt}} = -2 \ln \mathcal{F}_s. \quad (\text{A3})$$

To construct the controlled system, consider the controlled dynamics (33). The backwards generator for this model is  $\mathcal{L}^{\text{con}}$ , which acts on functions  $\mathcal{G} = \mathcal{G}(\mathbf{r}_1, \dots, \mathbf{r}_N, \theta_1, \dots, \theta_N)$  as

$$\begin{aligned}\mathcal{L}^{\text{con}}(\mathcal{G}) &= \sum_i \left[ D\nabla_i^2 \mathcal{G} + D_r \frac{\partial^2 \mathcal{G}}{\partial \theta_i^2} - D_r \frac{\partial U^{\text{con}}}{\partial \theta_i} \frac{\partial \mathcal{G}}{\partial \theta_i} \right] \\ &+ \sum_i [v^{\text{con}}\mathbf{u}_i - D\nabla_i(U + U^{\text{con}})] \cdot \nabla_i \mathcal{G}.\end{aligned}\quad (\text{A4})$$

Denoting the right hand side of (A2) by  $\mathcal{L}_s(\mathcal{F}_s)$ , the generator of the optimally controlled dynamics may be then derived using the general formula  $\mathcal{L}^{\text{con}}(\mathcal{G}) = \mathcal{F}_s^{-1} \mathcal{L}_s(\mathcal{F}_s \mathcal{G}) - \psi(s)\mathcal{G}$  (see, for example, [68]). Combining these ingredients, the optimally controlled dynamics is given by (33) and (34).

## 2. KL divergence between controlled and natural dynamics

To derive a formula for the KL divergence in Eq. (39), we recall its definition

$$\mathcal{D}_{\text{KL}}(Q||P) = \int Q(x) \ln \frac{Q(x)}{P(x)} dx, \quad (\text{A5})$$

where  $Q, P$  are probability densities. Using standard techniques in stochastic processes [79–81], the path probability distribution for ABPs can be derived. We work in Stratonovich calculus throughout. The path probability is  $P(X) \propto \exp[-\sum_i \int_0^\tau \mathcal{S}_i(X(t)) dt]$  where  $X = \{\mathbf{r}_i, \theta_i\}_0^\tau$  indicates a path and

$$\mathcal{S}_i = \frac{|\dot{\mathbf{r}}_i - v_0\mathbf{u}_i + D\nabla_i U|^2}{4D} + \frac{\dot{\theta}_i^2}{4D_r} - \frac{D}{2} \nabla_i^2 U. \quad (\text{A6})$$

[We have neglected here a contribution to  $P(X)$  from the initial condition; this does not cause a problem because we consider finally the long-time limit in Eq. (39).]

Defining the corresponding quantity  $\mathcal{S}_i^{\text{con}}$  for the controlled process one has

$$\mathcal{D}_{\text{KL}}(P^{\text{con}}||P) = N\tau \langle \mathcal{S}_i - \mathcal{S}_i^{\text{con}} \rangle_{\text{con}}, \quad (\text{A7})$$

where the average is computed in the steady state of the controlled process. One finds

$$\begin{aligned}\mathcal{S}_i - \mathcal{S}_i^{\text{con}} &= \frac{v^{\text{con}} - v_0}{2D} \mathbf{u}_i \cdot \dot{\mathbf{r}}_i + \Delta_i^{\text{con}} \\ &- \frac{|(v^{\text{con}} - v_0)\mathbf{u}_i - D\nabla_i U^{\text{con}}|^2}{4D} + \frac{D}{2} \nabla_i^2 U^{\text{con}} \\ &- \frac{1}{4D_r} \left( \frac{\partial U^{\text{con}}}{\partial \theta_i} \right)^2 + \frac{D_r}{2} \left( \frac{\partial^2 U^{\text{con}}}{\partial \theta_i^2} \right),\end{aligned}\quad (\text{A8})$$

where  $\Delta_i^{\text{con}}$  has the property that  $\sum_i \int_0^\tau \Delta_i^{\text{con}} d\tau = [U^{\text{con}}(0) - U^{\text{con}}(\tau)]/2$  which will be negligible on taking the limit in Eq. (39). The KL divergence is the average of this quantity with respect to the controlled dynamics; the fact that  $U^{\text{con}}$  appears in both the action and in the equation of motion can be used to simplify the formulas for  $\mathcal{D}_{\text{KL}}$  as, for example, in Eq. (H3) of [37]. Such simplifications are not essential for this work, so we omit them.

For the controlled dynamics of (37) and (38), it is useful to denote by  $\varphi$  the angle between  $\mathbf{v}$  and the  $x$  axis. This can be obtained by considering

$$|\mathbf{v}|e^{i\varphi} = \frac{1}{N} \sum_j e^{i\theta_j}. \quad (\text{A9})$$

Then, the KL divergence of (39) can be obtained from (A7) and (A8), using  $\nabla_i U^{\text{con}} = 0$  and  $v^{\text{con}} = v_0$ , as

$$\lim_{\tau \rightarrow \infty} \frac{1}{N\tau} \mathcal{D}_{\text{KL}}(P_g^{\text{con}}||P) = \left\langle g\mathcal{I}_{1,\tau} - \frac{g^2}{D_r} \mathcal{I}_{2,\tau} \right\rangle_{\text{con}} - \frac{g}{N}, \quad (\text{A10})$$

where

$$\mathcal{I}_{1,\tau} = \frac{1}{\tau} \int_0^\tau |\mathbf{v}(t)|^2 dt, \quad (\text{A11})$$

$$\mathcal{I}_{2,\tau} = \frac{1}{N\tau} \int_0^\tau |\mathbf{v}(t)|^2 \sum_i \sin[\theta_i(t) - \varphi(t)]^2 dt. \quad (\text{A12})$$

The steady state averages of these integrals are independent of  $\tau$ ; hence, the right hand side of (A10) is also independent of  $\tau$  and no limit is required there.

The KL divergence of (A10) was evaluated by numerical simulation of the controlled dynamics to obtain the upper bounds in Fig. 7. (Analytic results are available for this

quantity in the large- $N$  limit, but numerical comparison with the large-deviation rate function requires results for finite  $N$ .)

## APPENDIX B: CLONING ALGORITHM WITH MODIFIED DYNAMICS

### 1. Modified dynamics

The cloning algorithm is applied to ABPs similarly to [37]. The number of clones is denoted by  $n_c$ . We typically repeat each computation  $N_{\text{runs}}$  times: we take a simple average of the results and use the standard error for an error bar. Convergence with respect to  $n_c$  is discussed below. The computational cost of our calculations is controlled primarily by  $n_c$ .

Given that we want to work with  $n_c$  as small as possible, the low probabilities of large-deviation events can be the origin of large systematic errors [51,52,82]. To improve the convergence of the algorithm, we use a modification (or control) of the dynamics, informed by the typical behavior of the system for large fluctuations of the biasing observable. In particular, we take (33) with  $v^{\text{con}} = v_s^{\text{con}}$  from (22) and  $U^{\text{con}} = U_g$  from (37). The choice of the parameter  $g$  will be discussed below.

Following the same steps as (A8), we obtain the result of [37] that the path probability distributions for the controlled and original dynamics are related as

$$P[X] \exp[-sN\tau(w_\tau)] \propto P^{\text{con}}[X] \exp(-sN\tau w_\tau^{\text{mod}}), \quad (\text{B1})$$

where  $X$  denotes a trajectory and the modified active work obeys

$$s w_\tau^{\text{mod}} = s \left( 1 - \frac{sD}{v_0^2} + w_{f,\tau} \right) - g \left( \frac{1}{N} - \mathcal{I}_{1,\tau} + \frac{g}{D_r} \mathcal{I}_{2,\tau} \right), \quad (\text{B2})$$

where  $w_{f,\tau}$  is the force part of the active work from (8), and the integrals  $\mathcal{I}_{1,\tau}$  and  $\mathcal{I}_{2,\tau}$  are defined in Eqs. (A11) and (A12).

Physically, this means that the biased ensemble (12) for the ABP model (2) can be formulated alternatively as a biased ensemble for the controlled ABP model, with a modified bias  $s w^{\text{mod}}$ . Hence, the cloning algorithm is valid as a method for sampling large deviations for any choice of the parameter  $g$  (including  $g = 0$ ).

The role of the parameter  $g$  is to improve the numerical efficiency, and hence to obtain accurate results with smaller numbers of clones. Several methods have been proposed for determining suitable values of such parameters [52,82,83]. Here, we adopt the following method, similar to [37], where the value is chosen ‘‘on the fly’’ within the algorithm. We note that for the controlled dynamics with potential  $U_g$ , we have for large  $N$  (and  $g < D_r$ ) that

$$\langle |\mathbf{v}|^2 \rangle_{\text{con}} = \frac{1}{N} \frac{1}{1 - gD_r^{-1}}. \quad (\text{B3})$$

The derivation is given in Appendix D 1. We obtained good results from the cloning algorithm by inverting this equation for  $g$  and taking

$$g = D_r \left( 1 - \frac{1}{N t^{-1} \langle \int_0^t dt' |\mathbf{v}(t')|^2 \rangle_{\text{clo}}} \right), \quad (\text{B4})$$

where  $\langle \dots \rangle_{\text{clo}}$  designates an average over the clones within the algorithm. Figure 11 shows examples of such torque param-

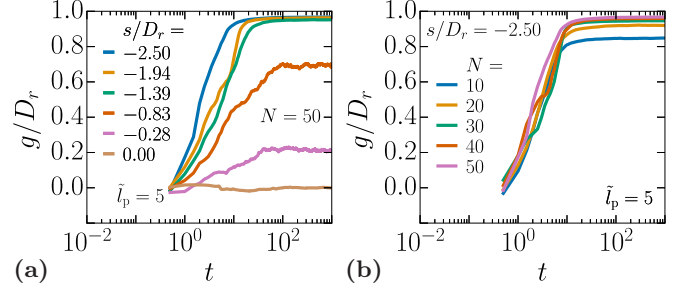


FIG. 11. Torque parameter  $g$  as a function of time using the relation (B4) and setting  $g(t=0) = 0$ , for different (a) biasing parameters  $s$  and (b) number of particles  $N$ . Parameter values:  $\phi = 0.65$ ,  $t_{\text{max}} = 10^3$ ,  $n_c = 10^3$ .

eters as functions of the simulation time. Note, this always gives  $g < D_r$ , the natural dynamics of the controlled system is always in the paramagnetic phase. However, as the biased system moves into the ferromagnetic phase we find that  $g$  gets close to  $D_r$ , and the controlled system approaches the mean-field critical point where fluctuations of  $\mathbf{v}$  are large. These large fluctuations are helpful for the algorithm, in that they generate a wide range of trajectories, from which the cloning part of the algorithm can select those with large values of  $w_\tau$ .

### 2. Convergence with respect to the number of clones

As stated in Ref. [52], the accuracy of the cloning algorithm is limited by the number  $n_c$  of copies of our system which we simultaneously evolve. Most of the results presented in the main text were computed for  $n_c = 10^3$  which is significantly lower than what was used, for example, in Ref. [37] (see Appendix D of that work).

We show in Fig. 12 relative errors on the values of the active work and the order parameter for two values of the persistence length and for  $N = 50$  particles, when varying the number of clones from  $n_c = 10^3$  to  $n_c = 5 \times 10^3$ , with respect to their value for the former number of clones. We have that the behavior of the relative error on both these quantities, which is at most of the order of 1%, indicates that the qualitative conclusions of the main text are robust.

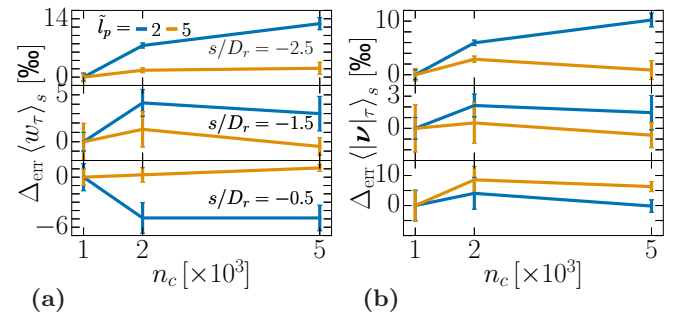


FIG. 12. Relative error of the value active work (left) and order parameter (right) to their value at  $n_c = 10^3$  (leftmost value),  $\Delta_{\text{err}}(\bullet)_s = [(\bullet)_s - (\bullet)_s(n_c = 10^3)] / (\bullet)_s(n_c = 10^3)$ . Parameter values:  $N = 50$ ,  $\phi = 0.65$ ,  $t_{\text{max}} = 10^3$ ,  $N_{\text{runs}} = 10$ .



### APPENDIX C: BIASED TRAJECTORIES OF TWO RTPs ON A RING

We consider the system of two RTPs defined in Sec. IV. The SCGF of (29) is the largest eigenvalue  $\psi^{\text{RTP}}(\lambda)$  that solves

$$\psi^{\text{RTP}}(\lambda)\mathbf{P}_\lambda = (\mathcal{L} - \lambda\dot{w}_f^{\text{RTP}}\mathbf{I})\mathbf{P}_\lambda, \quad (\text{C1})$$

where  $\mathcal{L}$  is the Fokker-Planck operator acting on a probability distribution vector  $\mathbf{P}_\lambda \equiv (P_\lambda^{++}, P_\lambda^{--}, P_\lambda^{+-}, P_\lambda^{-+})$ . The vector  $\mathbf{P}_\lambda$  depends on the distance  $r$  between the particles, which is measured clockwise around the circular system, starting at particle 1. Hence,  $0 \leq r \leq L$ . Since  $\mathbf{P}$  is a vector, the operator  $\mathcal{L}$  is a  $4 \times 4$  matrix, and the Fokker-Planck equation is

$$\dot{P}^{\alpha_1\alpha_2} = \sum_{\alpha'_1, \alpha'_2 = \pm 1} \mathcal{L}^{\alpha_1\alpha_2, \alpha'_1\alpha'_2} P^{\alpha'_1\alpha'_2}, \quad (\text{C2})$$

where the matrix elements of  $\mathcal{L}$  can be read off from the Fokker-Planck equation that corresponds to (26), which is

$$\begin{aligned} \dot{P}^{\alpha_1\alpha_2} &= v_0(\alpha_1 - \alpha_2) \frac{\partial}{\partial r} P^{\alpha_1\alpha_2} + 2 \frac{\partial}{\partial r} \left( P^{\alpha_1\alpha_2} \frac{\partial}{\partial r} V \right) \\ &+ \tau_p^{-1} (P^{\bar{\alpha}_1\alpha_2} + P^{\alpha_1\bar{\alpha}_2} - 2P^{\alpha_1\alpha_2}) \end{aligned} \quad (\text{C3})$$

in which  $\bar{\alpha}_i = -\alpha_i$ . In the hard-core limit, we expect the probability density functions to take the form [67,84]

$$P_\lambda^{\alpha_1\alpha_2}(r) = \varepsilon_\lambda^{\alpha_1\alpha_2}(r) + \gamma_\lambda^{\alpha_1\alpha_2,1} \delta(r) + \gamma_\lambda^{\alpha_1\alpha_2,r} \delta(L-r). \quad (\text{C4})$$

Here,  $\varepsilon_\lambda^{\alpha_1\alpha_2}$  is a smooth function of  $r$  that describes the probability to find the particles with the given orientations, at a separation  $r$ . Also,  $\gamma_\lambda^{\alpha_1\alpha_2,1}$  and  $\gamma_\lambda^{\alpha_1\alpha_2,r}$  indicate the probability that the particles are touching with particle 1 to the left (l) or right (r) of particle 2, with the prescribed orientations.

Particles never remain touching if their velocities point away from each other so  $\gamma_\lambda^{+-,r} = \gamma_\lambda^{-+,1} = 0$ . Note, however, that particles may be touching but moving parallel to each other,  $\gamma_\lambda^{+,+,r} > 0$  in general. In the hard-core limit, we have from (27) that  $\dot{w}_f^{\text{RTP}}$  is nonzero only when particles are touching and have opposite orientation, so

$$\begin{aligned} \dot{w}_f^{\text{RTP}} P_\lambda^{+-} &= -2v_0^2 \gamma_\lambda^{+-,1} \delta(r), \\ \dot{w}_f^{\text{RTP}} P_\lambda^{-+} &= -2v_0^2 \gamma_\lambda^{-+,r} \delta(L-r) \end{aligned} \quad (\text{C5})$$

with  $\dot{w}_f^{\text{RTP}} P_\lambda^{++} = 0 = \dot{w}_f^{\text{RTP}} P_\lambda^{--}$ .

The dominant eigenvector for (C1) obeys symmetry relations stemming from particle interchangeability (C6) and parity symmetry (C7):

$$P_\lambda^{\alpha_1\alpha_2}(r) = P_\lambda^{\alpha_2\alpha_1}(L-r), \quad (\text{C6})$$

$$P_\lambda^{\alpha_1\alpha_2}(r) = P_\lambda^{\bar{\alpha}_1\bar{\alpha}_2}(L-r). \quad (\text{C7})$$

We take as normalization condition

$$\int_0^L \sum_{\alpha_1, \alpha_2 = \pm 1} P_\lambda^{\alpha_1\alpha_2}(r) dr = 1. \quad (\text{C8})$$

The symmetry relations (C6) and (C7) imply for the probability density function of aligned particles that

$$\varepsilon_\lambda^{++}(r) = \varepsilon_\lambda^{--}(r) = \varepsilon_\lambda^{\alpha\alpha}(r), \quad (\text{C9})$$

$$\varepsilon_\lambda^{\alpha\alpha}(r) = \varepsilon_\lambda^{\alpha\alpha}(L-r). \quad (\text{C10})$$

Similarly, for the ‘‘sticking’’ terms

$$\gamma_\lambda^{+-,r} = \gamma_\lambda^{-+,1} = \gamma_\lambda^{\alpha\bar{\alpha}}, \quad (\text{C11})$$

$$\gamma_\lambda^{\alpha\alpha,1} = \gamma_\lambda^{\alpha\alpha,r} = \gamma_\lambda^{\alpha\alpha}, \quad (\text{C12})$$

which are to be interpreted as definitions of  $\gamma_\lambda^{\alpha\bar{\alpha}}$ ,  $\gamma_\lambda^{\alpha\alpha}$ . These symmetry relations greatly simplify our calculations.

We highlight that the introduction of thermal diffusion in Eq. (26) would smooth out the  $\delta$  functions associated to sticking in Eq. (C4), which will be replaced by exponentials decaying away from contact on a length scale set by the diffusivity [85].

#### 1. Unbiased steady state distribution

We first solve the case without bias ( $\lambda = 0$ ) in steady state ( $\dot{P} = 0$ ). Evaluating Eq. (C3) for  $r \neq 0, L$  gives

$$0 = 2v_0 \frac{\partial}{\partial r} \varepsilon_0^{\alpha\bar{\alpha}}(r) + \tau_p^{-1} [2\varepsilon_0^{\alpha\alpha}(r) - 2\varepsilon_0^{\alpha\bar{\alpha}}(r)]. \quad (\text{C13})$$

Using the symmetry properties (C6) and (C7) leads to  $\partial_r \varepsilon_0^{\alpha\bar{\alpha}}(r) = 0$  and hence  $\varepsilon_0^{\alpha\bar{\alpha}}(r) = \varepsilon_0^{\alpha\alpha}(r) = \varepsilon_0$ . In other words,  $P$  is independent of  $r$  and of the orientations, except when the particles are touching.

Relations between  $\gamma_0^{\alpha\alpha}$  and  $\gamma_0^{\alpha\bar{\alpha}}$  follow from integrating Eq. (C3) from  $r = 0^-$  to  $\varepsilon$  and then taking the hard-core limit:

$$\tau_p^{-1} (\gamma_0^{\alpha\bar{\alpha}} - 2\gamma_0^{\alpha\alpha}) = 0, \quad (\text{C14})$$

$$-2v_0\varepsilon_0 + \tau_p^{-1} 2\gamma_0^{\alpha\alpha} = 0. \quad (\text{C15})$$

Finally, using the normalization condition (C8),

$$4L\varepsilon_0 + 4\gamma_0^{\alpha\alpha} + 2\gamma_0^{\alpha\bar{\alpha}} = 1. \quad (\text{C16})$$

One can then solve to obtain

$$P_0^{++}(r) = P_0^{--}(r) = a + la\delta(r) + la\delta(L-r), \quad (\text{C17})$$

$$P_0^{+-}(r) = a + 2la\delta(r), \quad (\text{C18})$$

$$P_0^{-+}(r) = a + 2la\delta(L-r), \quad (\text{C19})$$

where  $a = [4(L+2l)]^{-1}$ , and the persistence length  $l = v_0\tau_p$ , as in the main text. This exactly recovers the continuous space and time results of Ref. [67], Eqs. (11) and (12). Also,

$$\langle \dot{w}_f^{\text{RTP}} \rangle = -\frac{2lv_0^2}{L+2l} \quad (\text{C20})$$

by (C5).

Note that a steady state distribution that is nonuniform also for particles not in contact, as observed for RTPs on a discrete lattice [67], can be introduced by considering finite-time tumbles [86] or thermal diffusion [85].

#### 2. Biased steady state distribution

We now solve the general biased case ( $\lambda \in \mathbb{R}$ ) using (C1). The method follows the unbiased case. For particles not in

contact we have

$$\begin{aligned} \psi^{\text{RTP}}(\lambda)\varepsilon_{\lambda}^{\alpha\alpha}(r) \\ = \tau_p^{-1}[\varepsilon_{\lambda}^{\alpha\bar{\alpha}}(r) + \varepsilon_{\lambda}^{\alpha\bar{\alpha}}(L-r) - 2\varepsilon_{\lambda}^{\alpha\alpha}(r)], \end{aligned} \quad (\text{C21})$$

$$\begin{aligned} \psi^{\text{RTP}}(\lambda)\varepsilon_{\lambda}^{\alpha\bar{\alpha}}(r) = 2\alpha v_0 \frac{\partial}{\partial r} \varepsilon_{\lambda}^{\alpha\bar{\alpha}}(r) \\ + \tau_p^{-1}[2\varepsilon_{\lambda}^{\alpha\alpha}(r) - 2\varepsilon_{\lambda}^{\alpha\bar{\alpha}}(r)]. \end{aligned} \quad (\text{C22})$$

Hence,

$$\frac{d^2}{dr^2}(\varepsilon_{\lambda}^{\alpha\bar{\alpha}} + \varepsilon_{\lambda}^{\bar{\alpha}\alpha}) = k_{\lambda}^2(\varepsilon_{\lambda}^{\alpha\bar{\alpha}} + \varepsilon_{\lambda}^{\bar{\alpha}\alpha}) \quad (\text{C23})$$

with

$$k_{\lambda}^2 l^2 = \frac{\tau_p \psi^{\text{RTP}}(\lambda)}{4} [4 + \tau_p \psi^{\text{RTP}}(\lambda)]. \quad (\text{C24})$$

Note that  $k_{\lambda}$  may be either real or imaginary. The solutions for  $\varepsilon$  (in both cases) can then be expressed as

$$\varepsilon_{\lambda}^{\alpha\alpha}(r) = \frac{1}{2 + \tau_p \psi^{\text{RTP}}(\lambda)} A_{\lambda} (e^{-k_{\lambda} r} + e^{-k_{\lambda}(L-r)}), \quad (\text{C25})$$

$$\begin{aligned} 2\varepsilon_{\lambda}^{\alpha\bar{\alpha}}(r) = \left(1 - \frac{2\alpha k_{\lambda} l}{2 + \tau_p \psi^{\text{RTP}}(\lambda)}\right) A_{\lambda} e^{-k_{\lambda} r} \\ + \left(1 + \frac{2\alpha k_{\lambda} l}{2 + \tau_p \psi^{\text{RTP}}(\lambda)}\right) A_{\lambda} e^{-k_{\lambda}(L-r)}, \end{aligned} \quad (\text{C26})$$

where there is a single constant of integration  $A_{\lambda}$  because we have enforced the symmetry (C10).

We now derive four equations that can be used to express  $(\gamma_{\lambda}^{\alpha\alpha}, \gamma_{\lambda}^{\alpha\bar{\alpha}}, A_{\lambda}, \lambda)$  as functions of  $\psi^{\text{RTP}}(\lambda)$ , which enables a full solution of this problem. Integrating Eq. (C1) from  $0^-$  to  $\epsilon$ , as in the unbiased case, and taking the  $++$  component of the vector  $\mathbf{P}$ , one obtains

$$\psi^{\text{RTP}}(\lambda)\gamma_{\lambda}^{\alpha\alpha} = \tau_p^{-1}(\gamma_{\lambda}^{\alpha\bar{\alpha}} - 2\gamma_{\lambda}^{\alpha\alpha}). \quad (\text{C27})$$

Similarly, taking the  $P^-$  component gives

$$0 = -2v_0 \varepsilon_{\lambda}^{\alpha\bar{\alpha}}(0^+) + \tau_p^{-1} 2\gamma_{\lambda}^{\alpha\alpha}, \quad (\text{C28})$$

in which  $\varepsilon_{\lambda}^{\alpha\bar{\alpha}}(0^+)$  may be substituted using (C26) to obtain

$$\gamma_{\lambda}^{\alpha\alpha} = \frac{\tau_p v_0 A_{\lambda}}{2} \left[ (1 + e^{-k_{\lambda} L}) + \frac{2k_{\lambda} l}{2 + \tau_p \psi^{\text{RTP}}(\lambda)} (1 - e^{-k_{\lambda} L}) \right]. \quad (\text{C29})$$

In addition, using (C25) and (C26) in the normalization condition (C8) leads to

$$\begin{aligned} 1 = 2\gamma_{\lambda}^{\alpha\alpha} [4 + \tau_p \psi^{\text{RTP}}(\lambda)] \\ + \frac{2A_{\lambda}}{k_{\lambda}} \frac{(1 - e^{-k_{\lambda} L}) [4 + \tau_p \psi^{\text{RTP}}(\lambda)]}{2 + \tau_p \psi^{\text{RTP}}(\lambda)}, \end{aligned} \quad (\text{C30})$$

where we have also used Eq. (C27) to eliminate  $\gamma^{\alpha\bar{\alpha}}$ .

Now, since the Fokker-Planck equation (C3) preserves the normalization of  $\mathbf{P}$ , one may integrate (C1) over  $r$  and sum over  $\alpha_1, \alpha_2$ , then apply (C8) to obtain  $\psi^{\text{RTP}}(\lambda) = 4\lambda v_0^2 \gamma_{\lambda}^{\alpha\bar{\alpha}}$ . Then, use (C27) to obtain

$$\psi^{\text{RTP}}(\lambda) = 4\gamma_{\lambda}^{\alpha\alpha} \lambda v_0^2 [2 + \tau_p \psi^{\text{RTP}}(\lambda)]. \quad (\text{C31})$$

Equations (C27) and (C29)–(C31) are the promised four equations for  $(\gamma_{\lambda}^{\alpha\alpha}, \gamma_{\lambda}^{\alpha\bar{\alpha}}, A_{\lambda}, \lambda)$ , in terms of  $\psi^{\text{RTP}}(\lambda)$ . The problem

is now solved by computing the inverse of  $\psi^{\text{RTP}}(\lambda)$ , which amounts to treating  $\psi^{\text{RTP}}$  as a parameter and solving for  $\lambda$  and the other variables. Note that  $k_{\lambda}$  is fully determined by the value of  $\psi^{\text{RTP}}$  [see (C24)].

To simplify the computation, we introduce dimensionless variables that treat the persistence length  $l$  as the unit of length:

$$\begin{aligned} \tilde{\lambda} &= \lambda l v_0, \\ \tilde{A}_{\tilde{\lambda}} &= A_{\lambda} l, \\ \tilde{\psi}^{\text{RTP}} &= \tau_p \psi^{\text{RTP}}, \\ \tilde{k}_{\tilde{\lambda}} &= k_{\lambda} l, \\ \tilde{L} &= \frac{L}{l}. \end{aligned} \quad (\text{C32})$$

The ratio  $\gamma_{\tilde{\lambda}}^{\alpha\alpha} / \tilde{A}_{\tilde{\lambda}}$  determines the relative probabilities of the particles being in contact or separated. One has from (C29) that this ratio can be expressed in terms of  $\tilde{\psi}^{\text{RTP}}$  as

$$\frac{\gamma_{\tilde{\lambda}}^{\alpha\alpha}}{\tilde{A}_{\tilde{\lambda}}} = \frac{1 + e^{-\tilde{k}_{\tilde{\lambda}} \tilde{L}}}{2} + \frac{\tilde{k}_{\tilde{\lambda}} (1 - e^{-\tilde{k}_{\tilde{\lambda}} \tilde{L}})}{2 + \tilde{\psi}^{\text{RTP}}} \quad (\text{C33})$$

and dividing Eq. (C30) by Eq. (C31) yields (after some rearrangements)

$$\tilde{\lambda} = \frac{\tilde{\psi}^{\text{RTP}} (\tilde{\psi}^{\text{RTP}} + 4)}{2(\tilde{\psi}^{\text{RTP}} + 2)} \left[ 1 + \frac{\tilde{A}_{\tilde{\lambda}}}{\gamma_{\tilde{\lambda}}^{\alpha\alpha} \tilde{k}_{\tilde{\lambda}} (\tilde{\psi}^{\text{RTP}} + 2)} (1 - e^{-\tilde{k}_{\tilde{\lambda}} \tilde{L}}) \right] \quad (\text{C34})$$

yielding  $\tilde{\psi}^{\text{RTP}}(\tilde{\lambda}) \sim 2\tilde{\lambda}$  when  $\lambda \rightarrow \infty$ . Combining (C33) and (C34) gives the promised inverse of  $\tilde{\psi}^{\text{RTP}}(\tilde{\lambda})$ , as

$$\tilde{\lambda} = \frac{\tilde{\psi}^{\text{RTP}} (\tilde{\psi}^{\text{RTP}} + 4)}{(\tilde{\psi}^{\text{RTP}} + 2)} \left[ \frac{1}{2} + \frac{1}{\Omega_{\tilde{L}}(\tilde{\psi}^{\text{RTP}})} \right] \quad (\text{C35})$$

with

$$\begin{aligned} \Omega_{\tilde{L}}(\tilde{\psi}^{\text{RTP}}) &= \frac{\tilde{\psi}^{\text{RTP}}}{2} (\tilde{\psi}^{\text{RTP}} + 4) \\ &+ \tilde{k}_{\tilde{\lambda}} (\tilde{\psi}^{\text{RTP}} + 2) \frac{1 + e^{-\tilde{k}_{\tilde{\lambda}} \tilde{L}}}{1 - e^{-\tilde{k}_{\tilde{\lambda}} \tilde{L}}}, \end{aligned} \quad (\text{C36})$$

where we used also (C24).

Recall from (C24) that  $\tilde{k}$  is an imaginary number for  $-4 < \tilde{\psi}^{\text{RTP}} < 0$ . In this case, it is useful to rewrite

$$\tilde{k}_{\tilde{\lambda}} \frac{1 + e^{-\tilde{k}_{\tilde{\lambda}} \tilde{L}}}{1 - e^{-\tilde{k}_{\tilde{\lambda}} \tilde{L}}} = |\tilde{k}_{\tilde{\lambda}}| \cot\left(\frac{|\tilde{k}_{\tilde{\lambda}}| \tilde{L}}{2}\right). \quad (\text{C37})$$

With this result in hand, a careful analysis shows that  $\Omega_{\tilde{L}}(\psi)$  has at least one zero for  $-2 < \psi < 0$ , at which point  $\tilde{\lambda}$  diverges. This implies that  $\tilde{\psi}^{\text{RTP}}(\tilde{\lambda})$  has a horizontal tangent for  $\tilde{\lambda} \rightarrow -\infty$ . The location of the (largest) zero sets the smallest possible value for  $\tilde{\psi}^{\text{RTP}}(\tilde{\lambda})$ , which is achieved as  $\tilde{\lambda} \rightarrow -\infty$  [see Fig. 6(a)].

### 3. Scaling regime

It is instructive to consider two particles in a very large system,  $L \rightarrow \infty$ . The system has an associated scaling limit whose behavior is shown in Fig 13.

Since the persistence length of the run-and-tumble motion is much less than the system size, the particle motion on large

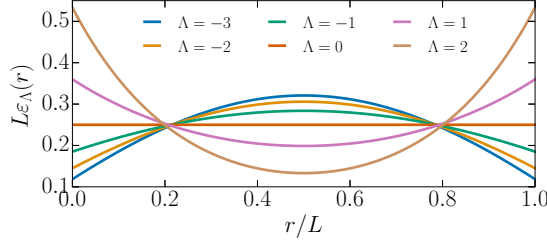


FIG. 13. Regular part of the probability density function  $\varepsilon_\Lambda(r)$  from (C48) and (C49) scaled by the ring length  $L$ .

scales can be characterized as (athermal) diffusion, and the particle explores the system on a timescale  $O(L^2)$ . Since  $\psi^{\text{RTP}}$  is an inverse timescale, it is expected to be  $O(L^{-2})$ . Moreover, it follows from Eq. (C20) that typical values of  $w_f^{\text{RTP}}$  are of order  $L^{-1}$  in this regime. Physically, this small value arises because of the small fraction of time that the particles spend in contact, when  $L$  is large. Hence,

$$\psi^{\text{RTP}}(\lambda) = \frac{2lv_0^2}{L+2l}\lambda + O(\lambda^2). \quad (\text{C38})$$

One then expects a scaling form as  $L \rightarrow \infty$ :

$$\psi^{\text{RTP}}(\lambda) \simeq L^{-2}\varphi(\lambda L), \quad (\text{C39})$$

which will be verified below. The corresponding form of the rate function for  $w_f^{\text{RTP}}$  is obtained from (14) as

$$I(w_f) \simeq L^{-2}\mathcal{I}(w_f L). \quad (\text{C40})$$

The natural dimensionless quantities in this regime are

$$\Lambda = \frac{\lambda Lv_0}{2}, \quad (\text{C41})$$

$$\Psi(\Lambda) = \frac{\tau_p L^2}{4l^2} \psi^{\text{RTP}}(\lambda) \quad (\text{C42})$$

so that  $\varphi(\lambda L) = 2l^2\Psi(\Lambda)/\tau_p$ . The quantities  $\Lambda$ ,  $\Psi(\Lambda)$  have same sign.

Since  $\psi^{\text{RTP}}$  is  $O(L^{-2})$ , then Eq. (C35) shows that  $k = O(L^{-1})$ . At the lowest order in  $L^{-1}$  we then infer from Eq. (C35) that for  $\Lambda > 0$ , then

$$\Lambda = \sqrt{\Psi} \tanh(\sqrt{\Psi}), \quad (\text{C43})$$

while for  $\Lambda < 0$  we have

$$\Lambda = -\sqrt{|\Psi|} \tan(\sqrt{|\Psi|}) \quad (\text{C44})$$

yielding  $\lim_{\Lambda \rightarrow -\infty} \Psi(\Lambda) = -\pi^2/4$ .

Moreover, at leading order in  $L^{-1}$ , it follows from Eqs. (C25) and (C26) that

$$\begin{aligned} \varepsilon_\Lambda^{\alpha\alpha}(r) &= \varepsilon_\Lambda^{\alpha\bar{\alpha}}(r) = \varepsilon_\Lambda(r) \\ &= \frac{1}{2}A_\Lambda(e^{-k_\Lambda r} + e^{-k_\Lambda(L-r)}) \end{aligned} \quad (\text{C45})$$

and from Eqs. (C31) and (C33) that

$$\gamma_\Lambda = \frac{l}{4L} \frac{\Psi(\Lambda)}{\Lambda}, \quad (\text{C46})$$

$$A_\Lambda = \frac{2\gamma_\Lambda}{l} \frac{1}{1 + e^{-k_\Lambda L}}. \quad (\text{C47})$$

Using Eq. (C24) leads for  $\Lambda > 0$  to

$$\varepsilon_\Lambda(r) = \frac{1}{4L} \frac{\Psi(\Lambda)}{\Lambda} \frac{\cosh\left[\sqrt{\Psi(\Lambda)}\left(1 - \frac{2r}{L}\right)\right]}{\cosh(\sqrt{\Psi(\Lambda)})} \quad (\text{C48})$$

and for  $\Lambda < 0$  to

$$\varepsilon_\Lambda(r) = \frac{1}{4L} \frac{\Psi(\Lambda)}{\Lambda} \frac{\cos\left[\sqrt{|\Psi(\Lambda)|}\left(1 - \frac{2r}{L}\right)\right]}{\cos(\sqrt{|\Psi(\Lambda)|})} \quad (\text{C49})$$

which we plot in Fig. 13.

The physical picture emerging from Fig. 13 is as follows. In a large system, a very weak bias  $\lambda = O(L^{-1})$  is sufficient to change qualitatively the separation of the particles. The resulting probability distributions are independent of particle orientation but depend on the particle separation through the scaling variable  $r/L$ . For  $\Lambda > 0$  (corresponding to reduced active work), the particles are more likely to approach each other, which favors collisions. For  $\Lambda < 0$  (enhanced work), they are more likely to be far apart, suppressing collisions. One sees from (C46) that the probability to find the particles in contact vanishes as  $L^{-1}$ ; this holds throughout the scaling regime  $\lambda = O(L^{-1})$ .

#### 4. Distribution over the infinite-time interval

The probability density vector  $\mathbf{P}_\lambda$  which satisfies (C1) and (C8) indicates the fraction of trajectories for which the particles have final orientations  $\alpha_i$  and separation  $r$  in the  $\lambda$  ensemble [52,68]. We are also interested in the fraction of time spent with given orientations  $\alpha_i$  and separation  $r$  in the  $\lambda$  ensemble, which we will denote  $\hat{\mathbf{P}}_\lambda$ . We expect these probability density functions to take the same form and respect the same symmetries as their final time counterparts (C4), (C6), (C7), so that

$$\hat{\mathbf{P}}_\lambda^{\alpha\alpha}(r) = \hat{\varepsilon}_\lambda^{\alpha\alpha}(r) + \hat{\gamma}_\lambda^{\alpha\alpha}\delta(r) + \hat{\gamma}_\lambda^{\alpha\alpha}\delta(L-r), \quad (\text{C50})$$

$$\hat{\mathbf{P}}_\lambda^{+-}(r) = \hat{\varepsilon}_{+-}(r) + \hat{\gamma}_{\alpha\bar{\alpha}}\delta(r), \quad (\text{C51})$$

$$\hat{\mathbf{P}}_\lambda^{-+}(r) = \hat{\varepsilon}_{-+}(r) + \hat{\gamma}_{\alpha\bar{\alpha}}\delta(L-r), \quad (\text{C52})$$

in addition to being normalized accordingly to (C8). In order to compute them, it is necessary to solve the eigenproblem adjoint to (C1):

$$\psi^{\text{RTP}}(\lambda)\mathbf{Q}_\lambda = (\mathcal{L}^\dagger - \lambda\hat{w}_f^{\text{RTP}}\mathbf{I})\mathbf{Q}_\lambda, \quad (\text{C53})$$

where the same  $\psi^{\text{RTP}}(\lambda)$  is the largest eigenvalue. It then follows that

$$\hat{\mathbf{P}}_\lambda^{\alpha_1\alpha_2}(r) = P_\lambda^{\alpha_1\alpha_2}(r)Q_\lambda^{\alpha_1\alpha_2}(r) \quad (\text{C54})$$

according to Ref. [68].

The matrix elements of  $\mathcal{L}^\dagger$  can be read off from the backward Fokker-Planck equation that corresponds to (C3):

$$\begin{aligned} \dot{Q}^{\alpha_1\alpha_2} &= -v_0(\alpha_1 - \alpha_2) \frac{\partial}{\partial r} Q^{\alpha_1\alpha_2} - 2 \frac{\partial}{\partial r} Q^{\alpha_1\alpha_2} \frac{\partial}{\partial r} V \\ &+ \tau_p^{-1}(Q^{\bar{\alpha}_1\alpha_2} + Q^{\alpha_1\bar{\alpha}_2} - 2Q^{\alpha_1\alpha_2}) \end{aligned} \quad (\text{C55})$$

and indicate that the eigenvector of this operator corresponding to the eigenvalue 0, which is also  $\mathbf{Q}_0$  in Eq. (C53), is constant. Since bias introduces in Eq. (C53) terms which are at least as regular as those in Eq. (C55), we expect  $Q_\lambda^{\alpha_1\alpha_2}$

to remain smooth and continuous for  $\lambda \neq 0$ . For particles not in contact,  $Q_{\lambda}^{\alpha_1\alpha_2}$  satisfies the same equations as  $P_{\lambda}^{\alpha_2\alpha_1}$  [Eqs. (C21) and (C22)], with the replacement  $v_0 \rightarrow -v_0$ . We can then finally conclude

$$\hat{\varepsilon}_{\lambda}^{\alpha\alpha}(r) = \frac{\hat{A}_{\lambda}}{A_{\lambda}} \varepsilon_{\lambda}^{\alpha\alpha}(r)^2, \quad (\text{C56})$$

$$\hat{\varepsilon}_{\lambda}^{\alpha\bar{\alpha}}(r) = \frac{\hat{A}_{\lambda}}{A_{\lambda}} \varepsilon_{\lambda}^{\alpha\bar{\alpha}}(r) \varepsilon_{\lambda}^{\bar{\alpha}\alpha}(r), \quad (\text{C57})$$

$$\hat{\gamma}_{\lambda}^{\alpha\alpha} = \frac{\hat{A}_{\lambda}}{A_{\lambda}} \gamma_{\lambda}^{\alpha\alpha} \varepsilon_{\lambda}^{\alpha\alpha}(0^+), \quad (\text{C58})$$

$$\hat{\gamma}_{\lambda}^{\alpha\bar{\alpha}} = \frac{\hat{A}_{\lambda}}{A_{\lambda}} \gamma_{\lambda}^{\alpha\bar{\alpha}} \varepsilon_{\lambda}^{+-}(L^-), \quad (\text{C59})$$

where  $\hat{A}_{\lambda}$  is a normalization constant for  $\hat{P}_{\lambda}$ .

### 5. Polarization

From (31) and (32) we identify

$$v_{\text{ave}}^{\text{RTP}}(\lambda) = \int_0^L [\hat{P}_{\lambda}^{++}(r) + \hat{P}_{\lambda}^{--}(r)] dr \quad (\text{C60})$$

$$v_{\text{ave}}^{\text{RTP}}(\lambda) = \frac{4\gamma_{\lambda}^{\alpha\alpha} \varepsilon_{\lambda}^{\alpha\alpha}(0^+) + 2 \int_0^L \varepsilon_{\lambda}^{\alpha\alpha}(r)^2 dr}{4\gamma_{\lambda}^{\alpha\alpha} \varepsilon_{\lambda}^{\alpha\alpha}(0^+) + 2\gamma_{\lambda}^{\alpha\alpha} (\tilde{\psi}^{\text{RTP}} + 2) \varepsilon_{\lambda}^{+-}(L^-) + 2 \int_0^L \varepsilon_{\lambda}^{\alpha\alpha}(r)^2 dr + 2 \int_0^L \varepsilon_{\lambda}^{\alpha\bar{\alpha}}(r) \varepsilon_{\lambda}^{\bar{\alpha}\alpha}(r) dr} \quad (\text{C65})$$

which avoids the need to determine  $\hat{A}_{\lambda}$ , and where we have used (C56)–(C59). This allows  $v_{\text{ave}}^{\text{RTP}}$  to be determined in full via (C25), (C26), and (C33). It is easily verified that  $v_{\text{end}}^{\text{RTP}} = v_{\text{ave}}^{\text{RTP}} = \frac{1}{2}$  for  $\lambda = 0$ , corresponding to a state where aligned and antialigned states are equiprobable.

### APPENDIX D: STATISTICS OF ORIENTATIONAL ORDER PARAMETER(S)

This Appendix analyzes (controlled) ABPs in situations where their orientations evolve independently of their positions. In these cases, the statistics of the order parameter  $\mathbf{v}$  can be computed.

#### 1. Mean-field analysis

We consider the dynamics of the particle orientations alone, for the controlled system (37). We have  $U_g^{\text{con}} = -g/(ND_r) \sum_{ij} \cos(\theta_i - \theta_j)$  so the controlled equation of motion for the orientation is

$$\dot{\theta}_i = -g/N \sum_j \sin(\theta_i - \theta_j) + \sqrt{2D_r} \xi_i. \quad (\text{D1})$$

Writing  $\sin(\theta_i - \theta_j) = \sin \theta_i \cos \theta_j - \cos \theta_i \sin \theta_j$ , and using (16) with  $\mathbf{v} = |\mathbf{v}|(\cos \varphi, \sin \varphi)$  yields (38) of the main text.

For  $N \gg 1$ , mean-field theory is valid because the individual  $\theta_i$  relax much faster than the global  $\varphi$ , and fluctuations of  $\mathbf{v}$  are also negligible. Without loss of generality, we take  $\varphi = 0$ ,

and (replacing  $\hat{P}$  with  $P$ )

$$v_{\text{end}}^{\text{RTP}}(\lambda) = \int_0^L [P_{\lambda}^{++}(r) + P_{\lambda}^{--}(r)] dr. \quad (\text{C61})$$

An exact expression of the polarization  $v_{\text{end}}^{\text{RTP}}$  [Eq. (C61)] can be computed by noting that

$$v_{\text{end}}^{\text{RTP}}(\lambda) = \frac{4\gamma_{\lambda}^{\alpha\alpha} + 2 \int_0^L \varepsilon_{\lambda}^{\alpha\alpha}(r) dr}{4\gamma_{\lambda}^{\alpha\alpha} + 2\gamma_{\lambda}^{\alpha\bar{\alpha}} + 2 \int_0^L [\varepsilon_{\lambda}^{\alpha\alpha}(r) + \varepsilon_{\lambda}^{\alpha\bar{\alpha}}(r)] dr}, \quad (\text{C62})$$

where the denominator is 1 by (C8) and we used that  $\int_0^L [\varepsilon_{\lambda}^{\alpha\bar{\alpha}}(r) - \varepsilon_{\lambda}^{\bar{\alpha}\alpha}(r)] dr = 0$ . Moreover, Eq. (C27) yields  $\gamma_{\lambda}^{\alpha\bar{\alpha}} = (\tilde{\psi}^{\text{RTP}} + 2)\gamma_{\lambda}^{\alpha\alpha}$  and Eqs. (C25) and (C26) yield

$$2 \int_0^L \varepsilon_{\lambda}^{\alpha\bar{\alpha}}(r) dr = (\tilde{\psi}^{\text{RTP}} + 2) \int_0^L \varepsilon_{\lambda}^{\alpha\alpha}(r) dr, \quad (\text{C63})$$

therefore,

$$v_{\text{end}}^{\text{RTP}} = \frac{2}{\tilde{\psi}^{\text{RTP}} + 4} \quad (\text{C64})$$

whose dependence on  $\tilde{\lambda}$  can then be obtained parametrically via (C35).

An exact expression of the polarization  $v_{\text{ave}}^{\text{RTP}}$  [Eq. (C60)] is also available by writing

hence, it is consistent to set

$$|\mathbf{v}| = \langle \cos \theta_i \rangle_{\text{con}} \quad (\text{D2})$$

in Eq. (38). Treating this quantity as a fixed number, the steady state of the system obeys a Boltzmann distribution where each  $\theta_i$  is independent with distribution

$$p_{|\mathbf{v}|}^{\text{con}}(\theta_i) \propto \exp\left(\frac{2g|\mathbf{v}|}{D_r} \cos \theta_i\right), \quad (\text{D3})$$

where the constant of proportionality is fixed by normalization. Combining (D2) and (D3) leads to a self-consistency relationship

$$|\mathbf{v}| = \int p_{|\mathbf{v}|}^{\text{con}}(\theta) \cos \theta d\theta. \quad (\text{D4})$$

The integral can be expressed in terms of a Bessel function. However, the relevant question is for which values of  $g$  nontrivial solutions exist (excluding  $|\mathbf{v}| = 0$ ). For that purpose, one may expand for small values of the parameter  $2g|\mathbf{v}|/D_r$  which yields  $\langle \cos \theta_i \rangle_{\text{con}} = (g/D_r)|\mathbf{v}| + O(|\mathbf{v}|^3)$ . The correction term is negative and hence the nontrivial (ferromagnetic) solution appears for

$$g > D_r. \quad (\text{D5})$$

To analyze the paramagnetic phase we have by the central theorem of Sec. III B that for  $g = 0$  then  $p_0(\mathbf{v}) \propto e^{-N|\mathbf{v}|^2}$ . The Boltzmann distribution for the steady state of the controlled



system is obtained by multiplying by  $e^{-U_s^{\text{con}}}$ , yielding

$$p_g(\mathbf{v}) \propto \exp[-N|\mathbf{v}|^2(1 - gD_r^{-1})] \quad (\text{D6})$$

from which we obtain Eq. (B3). Consistent with the previous argument, that fluctuation diverges at the critical point  $g = D_r$ . For larger  $g$ , estimation of  $p_0$  by the central limit theorem is too simplistic and a more detailed analysis is required, for example, as in Eqs. (D2) and (D3).

## 2. Large deviations of the time-averaged order parameter

In this section we consider large deviations of the time-averaged (vectorial) order parameter

$$\bar{\mathbf{v}}_\tau = \frac{1}{\tau} \int_0^\tau \mathbf{v}(t) dt. \quad (\text{D7})$$

Recall that  $\mathbf{v}(t)$  is defined in Eq. (16) as a simple average of individual orientations which evolve independently by (2). Hence, the statistics of  $\bar{\mathbf{v}}_\tau$  can be analyzed exactly. As  $\tau \rightarrow \infty$  there is a large-deviation principle

$$p(\bar{\mathbf{v}}_\tau) \sim \exp[-\tau N \mathcal{J}(|\bar{\mathbf{v}}_\tau|)], \quad (\text{D8})$$

where  $\mathcal{J}$  is the rate function, which only depends on the modulus of  $\bar{\mathbf{v}}_\tau$ , by symmetry. Note that the function  $\mathcal{J}$  is distinct from  $\mathcal{J}_1$  in Eq. (40), which describes large deviations of the time-integrated modulus of  $\mathbf{v}$ . See, however, (D22) below. There is an associated SCGF

$$\psi^{\text{OP}}(h) = \lim_{\tau \rightarrow \infty} \frac{1}{N\tau} \ln \langle \exp(-\tau N \mathbf{h} \cdot \bar{\mathbf{v}}_\tau) \rangle, \quad (\text{D9})$$

where  $h = |\mathbf{h}|$ ; the right hand side only depends on the modulus of  $\mathbf{h}$ , by symmetry. There is a corresponding biased ensemble of trajectories, in which a generic observable  $\mathcal{A}$  has average value

$$\langle \mathcal{A} \rangle_h = \frac{\langle \mathcal{A} \exp(-\tau N \mathbf{h} \cdot \bar{\mathbf{v}}_\tau) \rangle}{\langle \exp(-\tau N \mathbf{h} \cdot \bar{\mathbf{v}}_\tau) \rangle}. \quad (\text{D10})$$

Since the rotors are independent under the ABP dynamics, the expectation value in Eq. (D9) reduces to a product of expectation values for single rotors. Hence,  $\psi^{\text{OP}}$  solves the eigenvalue problem

$$\psi^{\text{OP}}(h) \mathcal{F}_h(\theta) = D_r \mathcal{F}_h''(\theta) - h \mathcal{F}_h(\theta) \cos \theta. \quad (\text{D11})$$

As noted in Ref. [35], this problem is related to Mathieu's equation. Let  $\tilde{\theta} = \theta/2$  and define  $h$ -dependent quantities  $a = -4\psi^{\text{OP}}(h)/D_r$  and  $q = 2h/D_r$ . Defining also  $\tilde{\mathcal{F}}(\tilde{\theta}) = \mathcal{F}_h(2\tilde{\theta})$  we have

$$\tilde{\mathcal{F}}''(\tilde{\theta}) + (a - 2q \cos 2\tilde{\theta}) \tilde{\mathcal{F}}(\tilde{\theta}) = 0. \quad (\text{D12})$$

We recognize Eq. (D12) as Mathieu's differential equation [87]. For any real number  $q$  there is a countable infinity of possible values of  $a$  and associated solutions  $\tilde{\mathcal{F}}$ . We are interested in functions  $\mathcal{F}_h$  that are even and  $2\pi$  periodic. Hence,  $\tilde{\mathcal{F}}$  must be  $\pi$  periodic in  $\tilde{\theta}$ . We therefore introduce  $\mathcal{M}^{(0)}(\tilde{\theta}, q)$  which is the zeroth even and  $\pi$ -periodic Mathieu function and

$a_{\mathcal{M}}^{(0)}(q)$  its characteristic value [35], such that

$$\psi^{\text{OP}}(h) = -\frac{D_r}{4} a_{\mathcal{M}}^{(0)}(2h/D_r), \quad (\text{D13})$$

$$\mathcal{F}_h(\theta) = \mathcal{M}^{(0)}\left(\frac{\theta}{2}, \frac{2h}{D_r}\right). \quad (\text{D14})$$

Hence, by Legendre transform the rate function in Eq. (D8) is

$$\mathcal{J}(\bar{v}) = \sup_h [-h\bar{v} - \psi^{\text{OP}}(h)]. \quad (\text{D15})$$

This result is exact for all  $N$  and all  $\bar{v}$ .

To obtain additional physical insight, we obtain the quadratic behavior of the rate function at small  $\bar{v}$ . This requires that we solve (D12) for small  $q$ , which is a computation in perturbation theory [88]. Since  $\tilde{\mathcal{F}}$  is even and  $\pi$  periodic in  $\tilde{\theta}$  it can be expanded as  $\tilde{\mathcal{F}}(\tilde{\theta}) = 1 + \sum_{n=1}^{\infty} \beta_n \cos 2n\tilde{\theta}$ . The coefficients  $\beta_n$  and the eigenvalue  $a$  can then be expanded in powers of  $q$ . To leading order,

$$a = -q^2/2 + O(q^4), \quad \beta_1 = -q/2 + O(q^3) \quad (\text{D16})$$

and  $\beta_n = O(q^n)$  for  $n \geq 2$ . Hence, by (D13) and (D14)

$$\psi^{\text{OP}}(h) = \frac{1}{2D_r} h^2 + O(h^4), \quad (\text{D17})$$

and so for small  $\bar{v}$ ,

$$\mathcal{J}(\bar{v}) = \frac{1}{2} D_r \bar{v}^2 + O(\bar{v}^4). \quad (\text{D18})$$

The corresponding eigenfunction is

$$\mathcal{F}_h(\theta) = 1 - \frac{h}{D_r} \cos \theta + O(h^2). \quad (\text{D19})$$

Hence, the optimal control potential (for this single orientation vector) is  $U_{\text{opt}}^{\text{OP}} = -2 \ln \mathcal{F}_h$ , so

$$U_{\text{opt}}^{\text{OP}}(\theta) = \frac{2h}{D_r} \cos \theta + O(h^2). \quad (\text{D20})$$

Recall, we are considering here large deviations where the order parameter is aligned parallel (or antiparallel) to the  $x$  axis. For  $h > 0$ , the control potential acts to align the orientation vectors antiparallel to this axis, as expected from (D10).

## 3. Time-averaged modulus of the order parameter

The discussion of the large-deviation principle (D8) of the previous section is useful as a way to characterize the large-deviation principle (40) of the main text. In this case the relevant SCGF is

$$\psi_1(\lambda) = \lim_{\tau \rightarrow \infty} \frac{1}{N\tau} \ln \langle \exp(-\tau N \lambda \bar{v}_\tau) \rangle \quad (\text{D21})$$

and  $\mathcal{J}_1(\bar{v}) = \sup_\lambda [-\lambda \bar{v} - \psi_1(\lambda)]$ . In contrast to the previous case, this problem cannot (to our knowledge) be solved exactly for finite  $N$ . However, as  $N \rightarrow \infty$  the problem is of mean-field type. In this case, the intuitive result is that the large-deviation mechanism for  $\bar{v}_\tau$  should be the same as that of  $\bar{\mathbf{v}}_\tau$ , so that

$$\lim_{N \rightarrow \infty} \mathcal{J}_1(\bar{v}) = \mathcal{J}(\bar{v}) \quad (\text{D22})$$

as illustrated by Fig. 14(a).

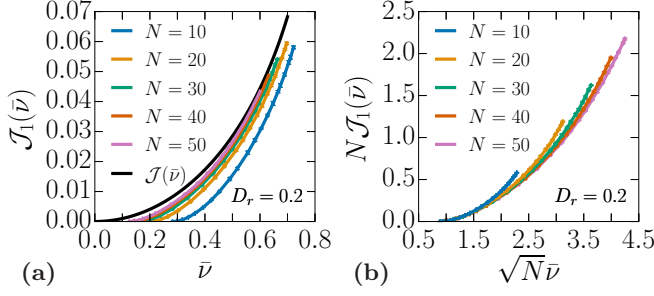


FIG. 14. (a), (b) Rate function of the time-averaged modulus of the order parameter  $\mathcal{J}_1(\bar{v})$  computed from cloning simulations of rotors (parameter values:  $n_c = 10^3$ ,  $t_{\max} = 10^3$ ) (a) and compared to the rate function of the time-averaged order parameter  $\mathcal{J}(\bar{v})$  [see Eq. (D15)].

We note from this figure that  $\mathcal{J}$  has its zero at the origin  $\langle \bar{v}_\tau \rangle = 0$ . On the other hand, (18) shows that the unique zero of  $\mathcal{J}_1$  is at  $\langle \bar{v}_\tau \rangle = \sqrt{\pi/(4N)}$  and in fact  $\mathcal{J}_1$  increases rapidly for smaller values of  $\bar{v}_\tau$ . Hence, for any given  $N$ , there is a region between 0 and  $\langle \bar{v}_\tau \rangle$  where  $\mathcal{J}_1$  deviates strongly from  $\mathcal{J}$ . However, this region vanishes as  $N \rightarrow \infty$  so (D22) holds for all  $\bar{v} > 0$ .

Moreover, Appendix F in Ref. [37] shows that

$$\tilde{\psi}_1(k) = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \ln \left\langle e^{-k \int_0^\tau \sqrt{N} |\mathbf{v}(t)| dt} \right\rangle \quad (\text{D23})$$

is a well-defined smooth function for  $N \gg 1$ . Therefore, with  $N\psi_1(\lambda) = \tilde{\psi}_1(\sqrt{N}\lambda)$ , we obtain

$$N\mathcal{J}_1(\bar{v}) \simeq B(\sqrt{N}\bar{v} - \sqrt{\pi/4})^2 \quad (\text{D24})$$

in the regime where  $\sqrt{N}\bar{v} = O(1)$ , and where  $B$  is a constant independent of  $N$ , as illustrated by Fig. 14(b).

The next step is to outline a derivation of (D22), which also yields the optimally controlled dynamics for this problem. The SCGF can be obtained by solving an eigenproblem

$$N\psi_1(\lambda)\mathcal{F}_\lambda^{\text{pol}} = D_r \sum_i \frac{\partial^2}{\partial \theta_i^2} \mathcal{F}_\lambda^{\text{pol}} - \lambda N |\mathbf{v}| \mathcal{F}_\lambda^{\text{pol}}, \quad (\text{D25})$$

where  $\mathcal{F}_\lambda^{\text{pol}} = \mathcal{F}_\lambda^{\text{pol}}(\theta_1, \dots, \theta_N)$  depends on all angles. Since this is a mean-field problem, the solution for large  $N$  can be approximated as

$$\mathcal{F}_\lambda^{\text{pol}} = f(\mathbf{v}) \prod_i \zeta_\lambda(\theta_i | \mathbf{v}), \quad (\text{D26})$$

where the orientation vectors interact only through their average, which is  $\mathbf{v}$ . Note also  $|\mathbf{v}| = N^{-1} \sum_i \cos(\theta_i - \varphi)$  where  $\varphi$  is the angle between  $\mathbf{v}$  and the  $x$  axis. We assume that  $\zeta$  is normalized as  $\int \zeta(\theta | \mathbf{v})^2 d\theta = 1$ . The function  $f$  is assumed to depend only on  $|\mathbf{v}|$ , and for large  $N$  it is sharply peaked in this variable, at  $|\mathbf{v}| = v^*$ . In order that  $\mathcal{F}^{\text{pol}}$  is also sharply peaked at  $v^*$ , we require a self-consistency condition

$$\int \cos(\theta - \varphi) \zeta(\theta | \mathbf{v})^2 d\theta = v^*. \quad (\text{D27})$$

The eigenproblem (D25) is Hermitian (self-adjoint) so there is a variational (Rayleigh-Ritz) formula for its largest

eigenvalue. Using  $\mathcal{F}^{\text{pol}}$  as ansatz yields

$$\psi_1(\lambda) \gtrsim -\lambda v^* - \frac{1}{2\pi N} \sum_i \int \Psi_\zeta(\theta_i, v^*) d\varphi, \quad (\text{D28})$$

where we used (D27) as well as  $\mathbf{v}^* = v^*(\cos \varphi, \sin \varphi)$  and

$$\Psi_\zeta(\theta, \mathbf{v}) = - \int D_r \zeta(\theta | \mathbf{v}) \frac{\partial^2}{\partial \theta^2} \zeta(\theta | \mathbf{v}) d\theta_i. \quad (\text{D29})$$

We have neglected terms in Eq. (D28) which arise from the action of the derivatives on  $\mathbf{v}$ ; these are negligible as  $N \rightarrow \infty$ . In fact, the mean-field structure of the problem means that the bound (D28) will become an equality as  $N \rightarrow \infty$ , if the optimal choice is made for  $\zeta$ .

It is convenient to work in terms of the rate function, using  $\mathcal{J}_1(\bar{v}) = \sup_\lambda [-\lambda \bar{v} - \psi_1(\lambda)]$  to see that  $\mathcal{J}_1(\bar{v}) \simeq \inf_\zeta \Psi_\zeta(\theta, \bar{v})$  where the maximization is again subject to (D27). Implementing this constraint and the normalization constraint on  $\zeta$  by Lagrange multipliers  $\mu_1, \mu_2$ , we find

$$D_r \frac{\partial^2 \zeta}{\partial \theta^2} + [\mu_1 - \mu_2 \cos(\theta - \varphi)] \zeta = 0. \quad (\text{D30})$$

Hence, we have recovered Mathieu's equation. Proceeding similar to Appendix D2 and using that extremization of the Lagrange multiplier  $\mu_2$  yields a maximum in this case, one obtains

$$\mathcal{J}_1(\bar{v}) \simeq \sup_{\mu_2} [-\mu_2 \bar{v} - \psi^{\text{OP}}(\mu_2)]. \quad (\text{D31})$$

The notation  $\simeq$  indicates that this relation becomes exact as  $N \rightarrow \infty$ . Equation (D22) follows on comparing with (D15). It follows that the optimal control potential for the large-deviation principle of (40) and (D21) is

$$U_{\text{con}}^{\text{pol}}(\theta_1, \dots, \theta_N) = -2 \sum_i \ln \mathcal{F}_\lambda(\theta_i - \varphi), \quad (\text{D32})$$

where the notation  $\mathcal{F}_\lambda$  indicates the function defined in Eq. (D14), evaluated at  $h = \lambda$ . This is indeed a mean-field-type interaction among orientations. By the same argument as (D20), it reduces for small  $\lambda$  to

$$U_{\text{con}}^{\text{pol}}(\theta_1, \dots, \theta_N) \simeq \frac{2\lambda}{ND_r} \sum_{ij} \cos(\theta_i - \theta_j). \quad (\text{D33})$$

This is nothing but  $U_g^{\text{con}}$  from (37) with  $g = -2\lambda$ . The result is that for small values of  $\bar{v}$ , Eq. (37) is an optimal control potential for large deviations of the orientation.

#### 4. Expansion of $\omega$

This section derives Equ. (57) of the main text. We assume that  $\omega(\bar{\rho}, \mathbf{P})$  can be inferred from the ensemble of trajectories biased with respect to the polarization  $\mathbf{P}$  [Eq. (56)]:

$$\omega(\bar{\rho}, \mathbf{P}) = \frac{\langle \bar{\rho} w_\tau e^{-\mathbf{h}(\mathbf{P}) \cdot \tau N \bar{v}_\tau} \rangle}{\langle e^{-\mathbf{h}(\mathbf{P}) \cdot \tau N \bar{v}_\tau} \rangle} \Big|_{\mathbf{h}(\mathbf{P}), \langle \bar{v}_\tau \rangle_{\mathbf{h}(\mathbf{P})} = \mathbf{P}}, \quad (\text{D34})$$

where we have used the averaged polarization from Eq. (D7).

We have in the limit  $\mathbf{h}(\mathbf{P}) \rightarrow \mathbf{0}$ ,

$$\begin{aligned} & \langle w_\tau e^{-\mathbf{h}(\mathbf{P}) \cdot \tau N \bar{v}_\tau} \rangle \\ &= \langle w_\tau \rangle + \frac{1}{4} \tau^2 N^2 |\mathbf{h}(\mathbf{P})|^2 \langle w_\tau | \bar{v}_\tau|^2 \rangle, \end{aligned} \quad (\text{D35})$$

$$\langle e^{-\mathbf{h}(\mathbf{P}) \cdot \tau N \bar{v}_\tau} \rangle = 1 + \frac{1}{4} \tau^2 N^2 |\mathbf{h}(\mathbf{P})|^2 \langle |\bar{v}_\tau|^2 \rangle \quad (\text{D36})$$

up to  $O[\mathbf{h}(\mathbf{P})^2]$  terms and where we have discarded linear terms in  $\mathbf{h}(\mathbf{P})$  by symmetry, therefore,

$$\bar{\rho}^{-1}\omega(\bar{\rho}, \mathbf{P}) = \langle w_\tau \rangle + \frac{1}{4}\tau^2 N^2 |\mathbf{h}(\mathbf{P})|^2 \text{Cov}(w_\tau, |\bar{\mathbf{v}}_\tau|^2) \quad (\text{D37})$$

linking  $\omega(\bar{\rho}, \mathbf{P})$  and the covariance of the active work and the squared average polarization.

We note that

$$\langle \bar{\mathbf{v}}_\tau \rangle_{\mathbf{h}(\mathbf{P})} = -\frac{1}{2}\tau N \mathbf{h}(\mathbf{P}) \text{Var}(\bar{\mathbf{v}}_\tau) \quad (\text{D38})$$

using  $\langle \bar{\mathbf{v}}_\tau \rangle = \mathbf{0}$  in the limit  $\tau \rightarrow \infty$  such that  $\text{Var}(\bar{\mathbf{v}}_\tau) = \langle |\bar{\mathbf{v}}_\tau|^2 \rangle$ , and

$$\begin{aligned} \langle |\bar{\mathbf{v}}_\tau|^2 \rangle &= \frac{2}{D_r \tau N} \left[ 1 - \frac{1}{D_r \tau} (1 - e^{-D_r \tau}) \right] \\ &= \frac{2}{D_r \tau N}, \quad \tau \rightarrow \infty \end{aligned} \quad (\text{D39})$$

using  $\langle \mathbf{u}(\theta_i(t)) \cdot \mathbf{u}(\theta_j(t')) \rangle = \delta_{ij} e^{-D_r |t-t'|}$  from Eq. (2), therefore, with  $\langle \bar{\mathbf{v}}_\tau \rangle_{\mathbf{h}(\mathbf{P})} = \mathbf{P}$  we may write

$$\mathbf{h}(\mathbf{P}) = -\frac{2}{\tau N \text{Var}(\bar{\mathbf{v}}_\tau)} \mathbf{P} \underset{\tau \rightarrow \infty}{=} -D_r \mathbf{P}, \quad (\text{D40})$$

linking the biasing parameter and the polarization.

We then have

$$\begin{aligned} \bar{\rho}^{-1}\omega(\bar{\rho}, \mathbf{P}) - \langle w_\tau \rangle &= |\mathbf{P}|^2 \frac{1}{\text{Var}(\bar{\mathbf{v}}_\tau)^2} \text{Cov}(w_\tau, |\bar{\mathbf{v}}_\tau|^2) \\ &= \frac{1}{4} |\mathbf{P}|^2 \tau^2 N^2 D_r^2 \text{Cov}(w_\tau, |\bar{\mathbf{v}}_\tau|^2), \\ &\quad \tau \rightarrow \infty \end{aligned} \quad (\text{D41})$$

at leading order in  $|\mathbf{P}|^2$ .

## APPENDIX E: FLUCTUATIONS OF THE ACTIVE WORK IN THE HYDRODYNAMIC THEORY

This Appendix describes density fluctuations of ABPs at hydrodynamic level, including large deviations. We follow Ref. [60], which draws on earlier results including [58,74]. At hydrodynamic level, we are restricted to small biasing fields  $s = O(1/L^2)$ . In this case, the (fast) polarization field is unaffected by the bias and can be safely integrated out. At the level of (46), this leads to renormalization of the diffusion constant  $D_c$  but we do not distinguish the bare and renormalized values of  $D_c$ , for simplicity. Since the polarization has been integrated away, this analysis of density fluctuations is restricted to states with  $\langle \mathbf{P} \rangle = 0$ , but this is sufficient to cover homogeneous phases for  $s > 0$  (small enough that the system remains homogeneous) and for  $s < 0$  (small enough that there is no collective motion).

### 1. Quadratic theory

As in Ref. [60] (Sec. 5.1 and Appendix B), we consider a perturbation around the homogeneous profile

$$\rho(\mathbf{r}, t) = \bar{\rho} + \delta\rho(\mathbf{r}, t) \quad (\text{E1})$$

with  $\delta\rho \ll \bar{\rho}$  and  $\int \delta\rho(\mathbf{r}, t) d\mathbf{r} = 0$ . Since we consider an isotropic system  $\mathbf{P} = \mathbf{0}$ , we define

$$\bar{\omega}_0 = \omega(\bar{\rho}, \mathbf{P} = \mathbf{0}), \quad (\text{E2})$$

$$\bar{\omega}_0'' = \frac{\partial^2}{\partial \rho^2} \omega(\bar{\rho}, \mathbf{P} = \mathbf{0}) \quad (\text{E3})$$

so that we can Taylor expand the active work over  $\rho$ :

$$\begin{aligned} N\tau w_\tau &= L^2 \tau \bar{\omega}_0 + \frac{1}{2} \bar{\omega}_0'' \int_{[0, L]^2} \int_0^\tau (\delta\rho)^2 d^2\mathbf{r} dt \\ &\quad + O(\delta\rho^3). \end{aligned} \quad (\text{E4})$$

At the consistent level of expansion, the stochastic equation for the density (46) and (47) is

$$\frac{\partial}{\partial t} \delta\rho = D_c(\bar{\rho}) \nabla^2 \delta\rho - \sqrt{2\sigma(\bar{\rho})} \nabla \cdot \boldsymbol{\eta}. \quad (\text{E5})$$

We introduce the Fourier modes of the density

$$\tilde{\rho}_\mathbf{q} = \frac{1}{L^2} \int_{[0, L]^2} \delta\rho(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r} \quad (\text{E6})$$

so that

$$\delta\rho = \sum_{\mathbf{q} \neq (0,0)} \tilde{\rho}_\mathbf{q} e^{i\mathbf{q}\cdot\mathbf{r}}. \quad (\text{E7})$$

Hence,

$$N\tau w_\tau = L^2 \tau \bar{\omega}_0 + L^2 \bar{\omega}_0'' \sum_{\substack{q_x \geq 0, q_y \\ \mathbf{q} \neq (0,0)}} \int_0^\tau \tilde{\rho}_\mathbf{q} \tilde{\rho}_{-\mathbf{q}} dt, \quad (\text{E8})$$

where the sum runs over nonzero modes  $\mathbf{q} = 2\pi(n_x, n_y)/L$ , with  $q_x \geq 0$ , and where we have used

$$\begin{aligned} \int_{[0, L]^2} \delta\rho^2 d^2\mathbf{r} &= L^2 \sum_{\mathbf{q} \neq (0,0)} \tilde{\rho}_\mathbf{q} \tilde{\rho}_{-\mathbf{q}} \\ &= 2L^2 \sum_{\substack{q_x \geq 0, q_y \\ \mathbf{q} \neq (0,0)}} \tilde{\rho}_\mathbf{q} \tilde{\rho}_{-\mathbf{q}} \end{aligned} \quad (\text{E9})$$

according to Parseval's theorem. Since the theory is defined on the mesoscopic scale, sums over  $\mathbf{q}$  are restricted to  $|\mathbf{q}| < \Lambda$  where  $\Lambda$  is an upper cutoff of order unity [to be precise, it is of order  $|\Omega_r|^{-1/d}$ , for consistency with (44)].

From Eqs. (E5) and (E6), we derive the stochastic equation satisfied by the nonzero Fourier modes

$$\frac{\partial}{\partial t} \tilde{\rho}_\mathbf{q} = -D(\bar{\rho}) |\mathbf{q}|^2 \tilde{\rho}_\mathbf{q} + \sqrt{2\sigma(\bar{\rho})} |\mathbf{q}|^2 \tilde{\eta}_\mathbf{q}, \quad (\text{E10})$$

where the longitudinal part of the noise term  $\tilde{\eta}_\mathbf{q}$  is a complex Gaussian white noise with zero mean and variance

$$\begin{aligned} \langle \tilde{\eta}_\mathbf{q}(t) \tilde{\eta}_\mathbf{q}^*(t') \rangle &= \frac{1}{L^4 |\mathbf{q}|^2} \\ &\quad \times \iint_{[0, L]^2} \langle [(-i\mathbf{q}) \cdot \boldsymbol{\eta}(t)] [(i\mathbf{q}) \cdot \boldsymbol{\eta}(t')] \rangle d^2\mathbf{r} d^2\mathbf{r}' \\ &= \delta(t - t'), \end{aligned} \quad (\text{E11})$$

and also  $\langle \text{Re}(\tilde{\eta}_q(t)) \text{Im}(\tilde{\eta}_q(t')) \rangle = 0$ . Noises with different wave vectors  $\mathbf{q}$  are independent. The key point is that (E10) is diagonal in  $\mathbf{q}$  so every wave vector can be analyzed separately.

## 2. Biased ensemble of trajectories

The next step is to consider a biased ensemble defined by the (linear) equations of motion (E10) and the reweighting factor  $e^{-s\tau N w_\tau}$ , where the exponential factor (E8) is quadratic (and diagonal) in the density fluctuations.

Following Ref. [60] (Appendix B), we consider the complex Ornstein-Uhlenbeck process

$$z = -\zeta z + \sqrt{2\gamma} \eta, \quad (\text{E12})$$

where  $\eta$  is a complex Gaussian white noise with the same statistics as  $\tilde{\eta}_q$ . For a biased ensemble with exponential biasing factor of  $e^{-sK_\tau}$  where  $K_\tau = \alpha \int_0^\tau |z(t)|^2 dt$ , the scaled cumulant generating function for  $K_\tau$  can be computed. Identifying  $(z, \zeta, \gamma, \alpha)$  with  $(\rho_q, D_c(\bar{\rho})q^2, \sigma(\bar{\rho})q^2, \bar{\omega}_0'')$ , the resulting SCGF for  $w_\tau$  is obtained by summing over the modes, to find

$$\begin{aligned} \psi(s) &= \lim_{\tau \rightarrow \infty} \frac{1}{N\tau} \ln \langle e^{-sN\tau w_\tau} \rangle \\ &= -s \langle w_\tau \rangle - \frac{1}{N} \sum_{\substack{q_x \geq 0, q_y \\ \mathbf{q} \neq (0,0)}} \left( \sqrt{D_c(\bar{\rho})^2 |\mathbf{q}|^4 + 2s\bar{\omega}_0'' \sigma(\bar{\rho}) |\mathbf{q}|^2} - D_c(\bar{\rho}) |\mathbf{q}|^2 - \frac{s\bar{\omega}_0'' \sigma(\bar{\rho})}{D_c(\bar{\rho})} \right) \end{aligned} \quad (\text{E13})$$

consistent with [58,74]. We emphasize that this result is valid only on the hydrodynamic scale, which means very small bias  $s = O(1/L^2)$ .

Several results are available from this formula. We first compute

$$-w'(s) = \psi''(s) = \frac{[\bar{\omega}_0'' \sigma(\bar{\rho})]^2}{D(\bar{\rho})^3} \frac{1}{N} \sum_{\substack{q_x \geq 0, q_y \\ \mathbf{q} \neq (0,0)}} \frac{1}{q^2} \quad (\text{E14})$$

which is related to the variance of the active work from Eq. (20). The sum in this last expression can be approximated as

$$\begin{aligned} \sum_{\substack{q_x \geq 0, q_y \\ \mathbf{q} \neq (0,0)}} \frac{1}{q^2} &\simeq \frac{L^2}{(2\pi)^2} \int_{2\pi/L}^\Lambda \frac{\pi q dq}{q^2} \\ &= \frac{L^2}{4\pi} [\ln L + O(1)], \end{aligned} \quad (\text{E15})$$

where  $\Lambda$  is the upper cutoff on  $\mathbf{q}$ . Hence, (61) follows, by combining these results with (20) and (53). The origin of this

diverging variance is the presence of a slow (hydrodynamic) timescale, diverging as  $L^2$ .

The second result that is available from (E13) is that the argument of the square root will become negative if  $s\bar{\omega}_0''$  is sufficiently negative, indicating that the system becomes inhomogeneous. The instability is in the lowest mode, which has  $|\mathbf{q}| = 2\pi/L$ . Noting that  $\bar{\omega}_0'' < 0$  we obtain Eq. (62), which is the point at which the system becomes unstable to phase separation.

Finally, observe that since  $\psi$  in Eq. (E13) is the scaled cumulant generating function for squared density fluctuations, the structure factor of the biased ensemble can also be obtained by taking a derivative, leading to

$$\langle |\rho_q|^2 \rangle_s = \frac{\sigma(\bar{\rho}) |\mathbf{q}|^2}{\sqrt{D_c(\bar{\rho})^2 |\mathbf{q}|^4 + 2s\bar{\omega}_0'' \sigma(\bar{\rho}) |\mathbf{q}|^2}} \quad (\text{E16})$$

similar to [74]. Observe that the limiting behavior of this function at  $q \rightarrow 0$  is different according to whether  $s\bar{\omega}_0''$  is zero or positive. In the latter case then  $\langle |\rho_q|^2 \rangle_s \rightarrow 0$  as  $q \rightarrow 0$ , corresponding to hyperuniformity. Hence, (63) follows. If  $s\bar{\omega}_0'' < 0$ , then it is not permissible to take  $q \rightarrow 0$  in Eq. (E16), as the argument of the square root would be negative at small  $q$ , which signals phase separation, as noted above.

- 
- [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] C. Bechinger, R. Di Leonardo, H. Löwen, C. Reichhardt, G. Volpe, and G. Volpe, Active particles in complex and crowded environments, *Rev. Mod. Phys.* **88**, 045006 (2016).
- [3] É. Fodor and M. C. Marchetti, The statistical physics of active matter: From self-catalytic colloids to living cells, *Phys. A (Amsterdam)* **504**, 106 (2018).
- [4] X.-L. Wu and A. Libchaber, Particle Diffusion in a Quasi-Two-Dimensional Bacterial Bath, *Phys. Rev. Lett.* **84**, 3017 (2000).
- [5] J. Elgeti, R. G. Winkler, and G. Gompper, Physics of microswimmers—single particle motion and collective behavior: a review, *Rep. Prog. Phys.* **78**, 056601 (2015).
- [6] A. Cavagna, A. Cimarelli, I. Giardina, G. Parisi, R. Santagati, F. Stefanini, and M. Viale, Scale-free correlations in starling flocks, *Proc. Natl. Acad. Sci. USA* **107**, 11865 (2010).
- [7] A. Cavagna and I. Giardina, Bird flocks as condensed matter, *Annu. Rev. Condens. Matter Phys.* **5**, 183 (2014).
- [8] A. Bottinelli, D. T. J. Sumpter, and J. L. Silverberg, Emergent Structural Mechanisms for High-Density Collective Motion Inspired by Human Crowds, *Phys. Rev. Lett.* **117**, 228301 (2016).



- [9] N. Bain and D. Bartolo, Dynamic response and hydrodynamics of polarized crowds, *Science* **363**, 46 (2019).
- [10] J. Deseigne, O. Dauchot, and H. Chaté, Collective Motion of Vibrated Polar Disks, *Phys. Rev. Lett.* **105**, 098001 (2010).
- [11] N. Kumar, H. Soni, S. Ramaswamy, and A. K. Sood, Flocking at a distance in active granular matter, *Nat. Commun.* **5**, 4688 (2014).
- [12] J. Palacci, S. Sacanna, A. P. Steinberg, D. J. Pine, and P. M. Chaikin, Living crystals of light-activated colloidal surfers, *Science* **339**, 936 (2013).
- [13] I. Buttinoni, J. Bialké, F. Kümmel, H. Löwen, C. Bechinger, and T. Speck, Dynamical Clustering and Phase Separation in Suspensions of Self-Propelled Colloidal Particles, *Phys. Rev. Lett.* **110**, 238301 (2013).
- [14] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, Novel Type of Phase Transition in A System of Self-Driven Particles, *Phys. Rev. Lett.* **75**, 1226 (1995).
- [15] J. Toner and Y. Tu, Long-Range Order in A Two-Dimensional Dynamical XY Model: How Birds Fly Together, *Phys. Rev. Lett.* **75**, 4326 (1995).
- [16] H. Chaté, Dry aligning dilute active matter, *Annu. Rev. Condens. Matter Phys.* **11**, 189 (2020).
- [17] Y. Fily and M. C. Marchetti, Athermal Phase Separation of Self-Propelled Particles with No Alignment, *Phys. Rev. Lett.* **108**, 235702 (2012).
- [18] G. S. Redner, M. F. Hagan, and A. Baskaran, Structure and Dynamics of A Phase-Separating Active Colloidal Fluid, *Phys. Rev. Lett.* **110**, 055701 (2013).
- [19] M. E. Cates and J. Tailleur, Motility-induced phase separation, *Annu. Rev. Condens. Matter Phys.* **6**, 219 (2015).
- [20] X. Yang, M. L. Manning, and M. C. Marchetti, Aggregation and segregation of confined active particles, *Soft Matter* **10**, 6477 (2014).
- [21] S. C. Takatori, W. Yan, and J. F. Brady, Swim Pressure: Stress Generation in Active Matter, *Phys. Rev. Lett.* **113**, 028103 (2014).
- [22] A. P. Solon, J. Stenhammar, R. Wittkowski, M. Kardar, Y. Kafri, M. E. Cates, and J. Tailleur, Pressure and Phase Equilibria in Interacting Active Brownian Spheres, *Phys. Rev. Lett.* **114**, 198301 (2015).
- [23] É. Fodor, C. Nardini, M. E. Cates, J. Tailleur, P. Visco, and F. van Wijland, How Far from Equilibrium is Active Matter? *Phys. Rev. Lett.* **117**, 038103 (2016).
- [24] D. Mandal, K. Klymko, and M. R. DeWeese, Entropy Production and Fluctuation Theorems for Active Matter, *Phys. Rev. Lett.* **119**, 258001 (2017).
- [25] P. Pietzonka and U. Seifert, Entropy production of active particles and for particles in active baths, *J. Phys. A: Math. Theor.* **51**, 01LT01 (2017).
- [26] C. Nardini, É. Fodor, E. Tjhung, F. van Wijland, J. Tailleur, and M. E. Cates, Entropy Production in Field Theories without Time-Reversal Symmetry: Quantifying the Non-Equilibrium Character of Active Matter, *Phys. Rev. X* **7**, 021007 (2017).
- [27] S. Shankar and M. C. Marchetti, Hidden entropy production and work fluctuations in an ideal active gas, *Phys. Rev. E* **98**, 020604(R) (2018).
- [28] L. Dabelow, S. Bo, and R. Eichhorn, Irreversibility in Active Matter Systems: Fluctuation Theorem and Mutual Information, *Phys. Rev. X* **9**, 021009 (2019).
- [29] S. Toyabe, T. Okamoto, T. Watanabe-Nakayama, H. Taketani, S. Kudo, and E. Muneyuki, Nonequilibrium Energetics of A Single  $f_1$ -Atpase Molecule, *Phys. Rev. Lett.* **104**, 198103 (2010).
- [30] T. Speck, Stochastic thermodynamics for active matter, *Europhys. Lett.* **114**, 30006 (2016).
- [31] É. Fodor, W. W. Ahmed, M. Almonacid, M. Bussonnier, N. S. Gov, M.-H. Verlhac, T. Betz, P. Visco, and F. van Wijland, Nonequilibrium dissipation in living oocytes, *Europhys. Lett.* **116**, 30008 (2016).
- [32] L. Tociu, É. Fodor, T. Nemoto, and S. Vaikuntanathan, How Dissipation Constrains Fluctuations in Nonequilibrium Liquids: Diffusion, Structure, and Biased Interactions, *Phys. Rev. X* **9**, 041026 (2019).
- [33] É. Fodor, T. Nemoto, and S. Vaikuntanathan, Dissipation controls transport and phase transitions in active fluids: Mobility, diffusion and biased ensembles, *New J. Phys.* **22**, 013052 (2020).
- [34] F. Cagnetta, F. Corberi, G. Gonnella, and A. Suma, Large Fluctuations and Dynamic Phase Transition in A System of Self-Propelled Particles, *Phys. Rev. Lett.* **119**, 158002 (2017).
- [35] T. GrandPre and D. T. Limmer, Current fluctuations of interacting active brownian particles, *Phys. Rev. E* **98**, 060601(R) (2018).
- [36] S. Whitelam, K. Klymko, and D. Mandal, Phase separation and large deviations of lattice active matter, *J. Chem. Phys.* **148**, 154902 (2018).
- [37] T. Nemoto, É. Fodor, M. E. Cates, R. L. Jack, and J. Tailleur, Optimizing active work: Dynamical phase transitions, collective motion, and jamming, *Phys. Rev. E* **99**, 022605 (2019).
- [38] G. Gradenigo and S. N. Majumdar, A first-order dynamical transition in the displacement distribution of a driven run-and-tumble particle, *J. Stat. Mech.* (2019) 053206.
- [39] E. Mallmin, R. A. Blythe, and M. R. Evans, A comparison of dynamical fluctuations of biased diffusion and run-and-tumble dynamics in one dimension, *J. Phys. A: Math. Theor.* **52**, 425002 (2019).
- [40] F. Cagnetta and E. Mallmin, Efficiency of one-dimensional active transport conditioned on motility, *Phys. Rev. E* **101**, 022130 (2020).
- [41] P. Chiarantoni, F. Cagnetta, F. Corberi, G. Gonnella, and A. Suma, Work fluctuations of self-propelled particles in the phase separated state, *J. Phys. A: Math. Theor.* **53**, 36LT02 (2020).
- [42] T. GrandPre, K. Klymko, K. K. Mandadapu, and D. T. Limmer, Entropy production fluctuations encode collective behavior in active matter, *Phys. Rev. E* **103**, 012613 (2021).
- [43] G. Gallavotti and E. G. D. Cohen, Dynamical ensembles in stationary states, *J. Stat. Phys.* **80**, 931 (1995).
- [44] J. L. Lebowitz and H. Spohn, A gallavotti-cohen-type symmetry in the large deviation functional for stochastic dynamics, *J. Stat. Phys.* **95**, 333 (1999).
- [45] U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines, *Rep. Prog. Phys.* **75**, 126001 (2012).
- [46] V. Lecomte, C. Appert-Rolland, and F. van Wijland, Thermodynamic formalism for systems with markov dynamics, *J. Stat. Phys.* **127**, 51 (2007).
- [47] H. Touchette, The large deviation approach to statistical mechanics, *Phys. Rep.* **478**, 1 (2009).
- [48] R. L. Jack and P. Sollich, Large deviations and ensembles of trajectories in stochastic models, *Prog. Theor. Phys. Suppl.* **184**, 304 (2010).

- [49] R. L. Jack and P. Sollich, Effective interactions and large deviations in stochastic processes, *Eur. Phys. J.: Spec. Top.* **224**, 2351 (2015).
- [50] R. L. Jack, Ergodicity and large deviations in physical systems with stochastic dynamics, *Eur. Phys. J. B* **93**, 74 (2020).
- [51] C. Giardinà, J. Kurchan, and L. Peliti, Direct Evaluation of Large-Deviation Functions, *Phys. Rev. Lett.* **96**, 120603 (2006).
- [52] T. Nemoto, F. Bouchet, R. L. Jack, and V. Lecomte, Population-dynamics method with a multicanonical feedback control, *Phys. Rev. E* **93**, 062123 (2016).
- [53] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, Dynamical First-Order Phase Transition in Kinetically Constrained Models of Glasses, *Phys. Rev. Lett.* **98**, 195702 (2007).
- [54] L. O. Hedges, R. L. Jack, J. P. Garrahan, and D. Chandler, Dynamic order-disorder in atomistic models of structural glass formers, *Science* **323**, 1309 (2009).
- [55] T. Speck, A. Malins, and C. P. Royall, First-Order Phase Transition in A Model Glass Former: Coupling of Local Structure and Dynamics, *Phys. Rev. Lett.* **109**, 195703 (2012).
- [56] J. Tailleur and J. Kurchan, Probing rare physical trajectories with lyapunov weighted dynamics, *Nat. Phys.* **3**, 203 (2007).
- [57] T. Laffargue, K.-D. Nguyen Thu Lam, J. Kurchan, and J. Tailleur, Large deviations of lyapunov exponents, *J. Phys. A: Math. Theor.* **46**, 254002 (2013).
- [58] C. Appert-Rolland, B. Derrida, V. Lecomte, and F. van Wijland, Universal cumulants of the current in diffusive systems on a ring, *Phys. Rev. E* **78**, 021122 (2008).
- [59] R. Chetrite and H. Touchette, Variational and optimal control representations of conditioned and driven processes, *J. Stat. Mech.* (2015) P12001.
- [60] J. Dolezal and R. L. Jack, Large deviations and optimal control forces for hard particles in one dimension, *J. Stat. Mech.* (2019) 123208.
- [61] S. Torquato and F. H. Stillinger, Local density fluctuations, hyperuniformity, and order metrics, *Phys. Rev. E* **68**, 041113 (2003).
- [62] K. Sekimoto, Langevin equation and thermodynamics, *Prog. Theor. Phys. Suppl.* **130**, 17 (1998).
- [63] A. Puglisi and U. Marini Bettolo Marconi, Clausius relation for active particles: What can we learn from fluctuations, *Entropy* **19**, 356 (2017).
- [64] U. M. B. Marconi, A. Puglisi, and C. Maggi, Heat, temperature and clausius inequality in a model for active brownian particles, *Sci. Rep.* **7**, 46496 (2017).
- [65] V. Lecomte and J. Tailleur, A numerical approach to large deviations in continuous time, *J. Stat. Mech.* (2007) P03004.
- [66] T. Brewer, S. R. Clark, R. Bradford, and R. L. Jack, Efficient characterisation of large deviations using population dynamics, *J. Stat. Mech.* (2018) 053204.
- [67] A. B. Slowman, M. R. Evans, and R. A. Blythe, Jamming and Attraction of Interacting Run-and-Tumble Random Walkers, *Phys. Rev. Lett.* **116**, 218101 (2016).
- [68] R. Chetrite and H. Touchette, Nonequilibrium markov processes conditioned on large deviations, in *Annales Henri Poincaré* (Springer, Berlin, 2015), Vol. 16, pp. 2005–2057.
- [69] D. Jacobson and S. Whitelam, Direct evaluation of dynamical large-deviation rate functions using a variational ansatz, *Phys. Rev. E* **100**, 052139 (2019).
- [70] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, Macroscopic fluctuation theory, *Rev. Mod. Phys.* **87**, 593 (2015).
- [71] M. E. Cates and J. Tailleur, When are active brownian particles and run-and-tumble particles equivalent? consequences for motility-induced phase separation, *Europhys. Lett.* **101**, 20010 (2013).
- [72] M. Kourbane-Houssene, C. Erignoux, T. Bodineau, and J. Tailleur, Exact Hydrodynamic Description of Active Lattice Gases, *Phys. Rev. Lett.* **120**, 268003 (2018).
- [73] F. D. C. Farrell, M. C. Marchetti, D. Marenduzzo, and J. Tailleur, Pattern Formation in Self-Propelled Particles with Density-Dependent Motility, *Phys. Rev. Lett.* **108**, 248101 (2012).
- [74] R. L. Jack, I. R. Thompson, and P. Sollich, Hyperuniformity and Phase Separation in Biased Ensembles of Trajectories for Diffusive Systems, *Phys. Rev. Lett.* **114**, 060601 (2015).
- [75] S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press, Cambridge, 1999).
- [76] A. P. Solon, H. Chaté, and J. Tailleur, From Phase to Microphase Separation in Flocking Models: The Essential Role of Nonequilibrium Fluctuations, *Phys. Rev. Lett.* **114**, 068101 (2015).
- [77] Y.-E. Keta, [github.com/yketa/active\\_work](https://github.com/yketa/active_work), GitHub repository (2019), MIT licensed.
- [78] H. Touchette, Introduction to dynamical large deviations of markov processes, *Phys. A (Amsterdam)* **504**, 5 (2018).
- [79] L. Onsager and S. Machlup, Fluctuations and irreversible processes, *Phys. Rev.* **91**, 1505 (1953).
- [80] L. F. Cugliandolo and V. Lecomte, Rules of calculus in the path integral representation of white noise langevin equations: the onsager–machlup approach, *J. Phys. A: Math. Theor.* **50**, 345001 (2017).
- [81] L. F. Cugliandolo, V. Lecomte, and F. van Wijland, Building a path-integral calculus: a covariant discretization approach, *J. Phys. A: Math. Theor.* **52**, 50LT01 (2019).
- [82] A. Das and D. T. Limmer, Variational control forces for enhanced sampling of nonequilibrium molecular dynamics simulations, *J. Chem. Phys.* **151**, 244123 (2019).
- [83] U. Ray, G. K.-L. Chan, and D. T. Limmer, Exact Fluctuations of Nonequilibrium Steady States from Approximate Auxiliary Dynamics, *Phys. Rev. Lett.* **120**, 210602 (2018).
- [84] T. Arnoult de Pirey, G. Lozano, and F. van Wijland, Active Hard Spheres in Infinitely Many Dimensions, *Phys. Rev. Lett.* **123**, 260602 (2019).
- [85] A. Das, A. Dhar, and A. Kundu, Gap statistics of two interacting run and tumble particles in one dimension, *J. Phys. A: Math. Theor.* (to be published).
- [86] A. B. Slowman, M. R. Evans, and R. A. Blythe, Exact solution of two interacting run-and-tumble random walkers with finite tumble duration, *J. Phys. A: Math. Theor.* **50**, 375601 (2017).
- [87] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Table* (Dover, New York, 1970), Chap. 20, pp. 722–750.
- [88] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists*, 7th ed. (Elsevier, Amsterdam, 2013), Chap. 31.