Diversity of self-propulsion speeds reduces motility-induced clustering in confined active matter

5 authors, including:

Pablo de Castro
São Paulo State University
19 PUBLICATIONS   **56** CITATIONS

SEE PROFILE

F. M. Rocha
Aix-Marseille Université
8 PUBLICATIONS   **61** CITATIONS

SEE PROFILE

Saulo de Mesquita Diles
Federal University of Pará
18 PUBLICATIONS   **118** CITATIONS

SEE PROFILE

Rodrigo Soto
University of Chile
146 PUBLICATIONS   **1,961** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project BRS Model View project

Project Rheology and flows of magnetotactic bacteria View project

Diversity of self-propulsion speeds reduces motility-induced clustering in confined active matter

Pablo de Castro,[1, *] Francisco M. Rocha,[2] Saulo Diles,[3] Rodrigo Soto,[1] and Peter Sollich[4, 5]

[1]*Departamento de Física, FCFM, Universidad de Chile, Santiago, Chile*
[2]*Aix Marseille University, CNRS, IUSTI, 13453 Marseille, France*
[3]*Faculdade de Física, Universidade Federal do Pará, Campus Salinópolis,*
Rua Raimundo Santana Cruz S/N, 68721-000, Salinópolis, Pará, Brazil
[4]*Disordered Systems Group, Department of Mathematics,*
King's College London, London, United Kingdom
[5]*Institut für Theoretische Physik, Georg-August-Universität Göttingen, 37077 Göttingen, Germany*
(Dated: October 18, 2021)

Self-propelled swimmers such as bacteria agglomerate into clusters as a result of their persistent motion. In 1D, those clusters do not coalesce macroscopically and the stationary cluster size distribution (CSD) takes an exponential form. We develop a minimal lattice model for active particles in narrow channels to study how clustering is affected by the interplay between *self-propulsion speed diversity* and confinement. A mixture of run-and-tumble particles with a distribution of self-propulsion speeds is simulated in 1D. Particles can swap positions at rates proportional to their relative self-propulsion speed. Without swapping, we find that the average cluster size L_c *decreases* with diversity and follows a non-arithmetic power mean of the single-component L_c's, unlike the case of tumbling-rate diversity previously studied. Effectively, the mixture is thus equivalent to a system of identical particles whose self-propulsion speed is the *harmonic* mean self-propulsion speed of the mixture. With swapping, particles escape more quickly from clusters. As a consequence, L_c decreases with swapping rates and depends less strongly on diversity. We derive a dynamical equilibrium theory for the CSDs of binary and fully polydisperse systems. Similarly to the clustering behaviour of one-component models, our qualitative results for mixtures are expected to be universal across active matter. Using literature experimental values for the self-propulsion speed diversity of unicellular swimmers known as choanoflagellates, which naturally differentiate into slower and faster cells, we predict that the error in estimating their L_c via one-component models which use the conventional arithmetic mean self-propulsion speed is around 30%.

I. INTRODUCTION

A collection of self-propelled particles can spontaneously separate into dense and dilute regions even without attractive forces. This process, known as motility-induced phase separation (MIPS) [1–4], occurs if the propulsion direction is sufficiently persistent against stochasticity, in which case there is enough time for the particles to trap each other and form large clusters [5, 6]. For one-dimensional (1D) systems, such as fertilizing bacteria living in long narrow soil pores [7–10], MIPS generates an exponentially decaying stationary cluster size distribution (CSD) as shown for run-and-tumble (RT), active Brownian, and active Ornstein-Uhlenbeck particles [11–19]. When the self-propellers are sufficiently strong to push each other, the clusters themselves can also move persistently [19, 20], producing deviations from a simple exponential CSD [13]. In two-dimensional (2D) or quasi-two-dimensional systems, an exponential CSD modulated by a power law arises as observed in experiments with bacteria [21–24], in

experiments and simulations with active colloids [25, 26], and in lattice models [2]. The CSDs can also be affected by solvent-mediated hydrodynamic interactions between the particles [27].

In a typical bacteria population there is a broad dispersion of motility parameters, i.e. the bacteria are not identical swimmers [28–30]. However, for simplicity, effects of *motility diversity* on CSDs are usually overlooked [31]. For RT bacteria, one can consider that the tumbling rate or the self-propulsion speed (or both) is not the same for all particles, that is, the system has a distribution of motility parameters. The fact that in a population of bacteria different self-propulsion speeds are found is a result of their different ages, reproduction stages, shapes, sizes, running modes, etc [29, 32–34]. For both passive and active fluids, 'diversity' of some particle attribute can fundamentally change phase behaviour [35–50]. In Ref. [30], of which the present work can be regarded as a companion paper, we considered a multi-component mixture of RT particles on a 1D discrete lattice (as a simple model for active particles in narrow channels [9, 16, 51–64]) interacting only via excluded volume, i.e. they do not push each other. The distinct particle types were characterized by their own tumbling rates

* pdecastro@ing.uchile.cl

as inspired by experimental observations with *Escherichia coli*; these experiments show that the tumbling rate of each bacterium changes stochastically but slowly, leading approximately to a log-normal distribution of constant tumbling rates for the system [28, 65]. Particles moving directly towards each other were allowed to cross at a constant rate, therefore mimicking the effects of a "soft" confinement where particles can swap their positions along the quasi-1D channel. Interactions via biochemical signalling were assumed negligible or absent. We observed an exponential CSD with an average cluster size L_c that *increases* with tumbling-rate diversity. This clustering amplification phenomenon is induced solely by tumbling-rate diversity as the global average tumbling rate remains fixed in the analysis. On the other hand, by relaxing the confinement, large cluster sizes are reduced and tumbling-rate diversity becomes less important. Furthermore, tumbling-rate diversity generates an average cluster size L_c that is given by an arithmetic average of the L_c's that the single-component systems would have at the same global density.

The self-propulsion speeds in Ref. [30] were set the same for all particles. The motility diversity was therefore entirely encoded into the tumbling-rate distribution. In the present work we examine the complementary and fundamentally distinct case of self-propulsion speed diversity and its effects on the stationary cluster size distributions of active particles in narrow environments: particles differ only in their self-propulsion speed but are assigned the same tumbling rate. To isolate the effects of self-propulsion speed diversity, we consider a distribution of self-propulsion speeds whose system average is kept fixed while only its variance is tuned. We find that the average cluster size *decreases* with the variance of the self-propulsion speed distribution if one keeps everything else fixed, including the distribution average; see Fig. 1. Also, here we consider the clustering effects of particle "overtaking", that is, faster particles can overtake slower ones—if moving in the same direction—with a rate proportional to their relative self-propulsion speed.

Since in the monodisperse case (i.e., without diversity) L_c depends on the motility parameters self-propulsion speed and tumbling rate (denoted by v and α, respectively) as $L_c \sim \sqrt{v/\alpha}$, one could naïvely think that knowing L_c for tumbling-rate diversity automatically provides a prescription for obtaining L_c in the case of self-propulsion speed diversity (hereafter referred to just as *speed diversity*). We show here that this is not true. By employing an arithmetic average of the L_c's of the single-component systems, one would still get the

correct qualitative result that L_c *decreases* with speed diversity, but the simulation values for L_c presented below are much lower than the arithmetic average, implying that this type of average is quantitatively inadequate to describe speed diversity.
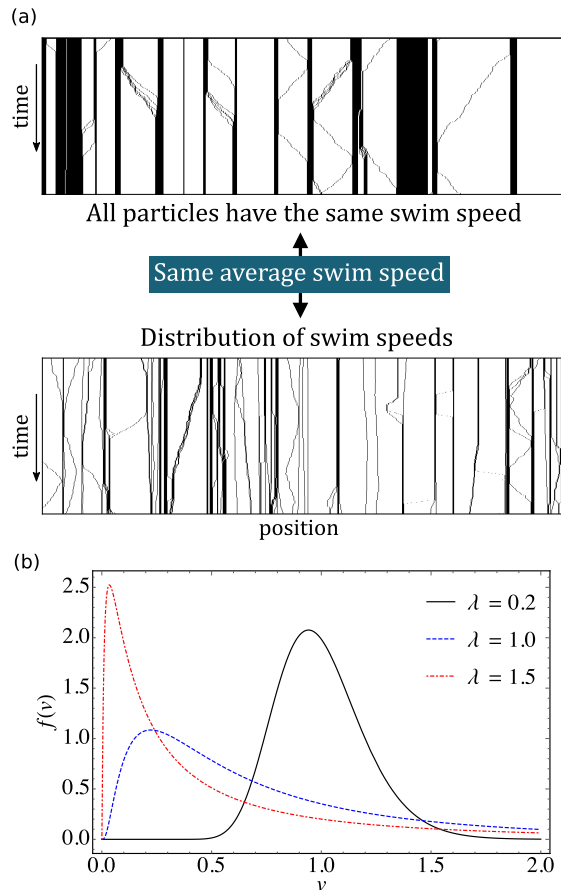

FIG. 1. (a) Steady state of a monodisperse system (top) of run-and-tumble particles on a 1D lattice compared against its mixture counterpart with polydispersity in the self-propulsion speeds (bottom); particles cannot swap ($p_s = 0$). Only 1000 of $N = 10^4$ simulated lattice sites are shown. Considering the entire simulated system, the global concentration and the average self-propulsion speed in both cases are the same: dimensionless concentration $\phi = 0.2$, tumbling rate $\alpha = 0.005$, and average self-propulsion speed $\langle v \rangle = 1$. In the mixture case (bottom) the polydisperse distribution parameter [see eqn (1)] is $\lambda = 2$. Time flows downwards along the vertical axis for 300 time steps. Positions are on the horizontal axis. Vacancies are in white and particles in black. (b) Polydisperse distribution of self-propulsion speeds, eqn (1), for three different distribution parameters λ at fixed average self-propulsion speed $\langle v \rangle = v_0 \exp\{(\lambda^2/2)\} = 1$.

We notice that Ref. [66] considers off-lattice active particles in 2D whose self-propulsion speeds vary both in time and between particles, follow-

ing a Gaussian colored noise. In a certain way, this also corresponds to having diversity of self-propulsion speeds, but not "quenched" in time as we do below. Furthermore, the parameter used by the authors of Ref. [66] to increase the variance of the self-propulsion speed distribution also increases the average self-propulsion speed, i.e., these two distribution parameters cannot be decoupled in their analysis. Finally, in Ref. [66] the cluster sizes are not calculated. Similarly, we mention that, for off-lattice 2D binary mixtures of fast and slow active Brownian particles, motility-induced stationary average cluster sizes have been briefly discussed in Ref. [41]. In particular, the authors showed numerically that by reducing the *ratio* between the slower and faster self-propulsion speeds, the average cluster sizes decrease. This could be regarded as a somewhat expected result since by reducing any active speed of the problem without increasing the rest of them, activity-induced clustering should indeed be reduced. This is crucially different from our work as here we decrease the slower speeds by the same amount that we increase the faster speeds. In our analysis we (i) consider RT particles on a lattice, (ii) analytically calculate cluster sizes, (iii) obtain new insights for confined systems, and (iv) probe quantitatively how clustering depends solely on self-propulsion speed diversity.

This paper is organized as follows. In Section II our lattice model is presented. Sections III and IV contain our main numerical and analytical results, respectively, for each model variant: (i) binary mixture, (ii) binary mixture with particle swapping, and (iii) fully polydisperse mixture. Section V gives our conclusions and discussion.

II. MODEL

Let us review the RT model presented in Ref. [2], where all particles are identical. Consider a 1D discrete lattice with N sites and periodic boundary conditions. The maximum occupancy per site is one. Each particle has a propulsion director, which can be left or right. The total number of particles is $M = \phi N$, where ϕ is the dimensionless global particle concentration. The initial positions are all distinct and otherwise random. Each particle is also given an initial random director. In each time step, M individual particle updates are performed. The update algorithm is as follows. A particle is selected at random and a new director for this particle is chosen at random, with probability α (hereafter we refer to α either as a probability or as a "rate", leaving the multiplication by the fixed time step implied, and similarly for the probabilities of

other elementary processes occurring at each time step). Thus, the probability to have a *different* director is $\alpha/2$. A tumble event occurs if the particle changes its director. Otherwise, the particle preserves its previous director. Next, if the propulsion director points towards a neighbouring empty site, then the particle moves to this new position. A particle is then chosen anew. The updates are sequential. Our units are such that the lattice spacing and the time step are fixed to unity. A particle can be chosen more than once in a single time step, so that the swim speed (i.e. the free-space speed acquired by the self-propelled particle in the run mode) of the mobile particles fluctuates around unity. To consider a nominal swim speed $v \neq 1$, all one needs to do is to perform vM particle updates, that is, an average of v updates per particle. In this case, we divide α by v to prevent the tumbling rate from scaling with the self-propulsion speed. In this model, there is neither imposed Vicsek-like velocity alignment [67] nor the spontaneous velocity alignment shown to arise due to interparticle forces in off-lattice systems of active particles [68] in narrow circular channels, where, for high persistence, particles can start rotating collectively [56]. Also, although active particles in a channel may push large clusters for long distances [69], it is not completely clear to which extent this is reflected in experimental CSDs [13, 21]. Such effects are not considered here. Finally, a cluster is defined as a contiguous group of occupied sites.

The above model is monodisperse: all particles have the same motility properties, i.e. they have identical swim speed and tumbling rate. Here we consider a more realistic scenario where the particles have distinct swim speeds, while their tumbling rates continue to be monodisperse. To simulate a system of particles of different speeds, we use the following approach. A selected particle that has been assigned a fixed speed v will be updated if v is larger than a random variable between 0 and $v_{\max}$, a simulation cutoff parameter. Otherwise, the selection is discarded. Once the total amount of individual particle updates is $\langle v \rangle M$, where $\langle v \rangle$ is the average particle speed in the sense of an arithmetic average, the time step ends. By using this procedure, the number of updates per particle per time step is $\langle v \rangle$, and faster (slower) particles are more (less) likely to be updated (and therefore to move), proportionally to their nominal speeds. Here, too, we divide α by the selected particle's v so that the tumbling rate does not scale up with the self-propulsion speed. For a binary mixture, we consider half the particles with swim speed $v_A = v_0(1 + \delta)$ and the other half with speed $v_B = v_0(1 - \delta)$. As we vary the degree of speed diversity, δ, the average speed $\langle v \rangle = v_0$ remains

fixed. We use specifically $v_{\max} = 2v_0$, as that is the largest value the larger speed v_A can have (for $\delta = 1$), and set $v_0 = 1$. The speeds are assigned so that the initial state is randomly homogeneous and well-mixed. For simplicity, we do not present simulation data for mixing proportions other than 50-50%. This is because the continuous distribution case described below already covers a more general situation, although the binary theory in Section IV was also validated via simulations with different global proportions.

We also consider a fully polydisperse system, i.e. with a continuous distribution of speeds. We choose a log-normal distribution, which is adequate for non-negative variables. Also, it corresponds to the same shape as the distribution of tumbling rates in *E. coli* bacteria[30] and is visually similar (see Fig. 1b) to the experimental speed distributions of many swimming microorganisms, *E. coli* included [29, 32–34]. Our normalized fully-polydisperse distribution thus reads

$$f(v) = \frac{1}{\sqrt{2\pi}\lambda v} \exp\left\{ \left(-\frac{[\log(v/v_0)]^2}{2\lambda^2} \right) \right\}, \quad (1)$$

where v_0 and λ are the distribution parameters. We keep $\langle v \rangle \equiv \int_0^\infty v f(v)\, \mathrm{d}v = v_0 \exp\{(\lambda^2/2)\}$ fixed while the polydispersity degree is changed by varying λ. The monodisperse case corresponds to the limit $\lambda \to 0$. For the simulations, the cut-off $v_{\max} = 4\langle v \rangle$ was found to be sufficiently large. In any case, according to the theory in Section IV, the particular functional form or parameters of the speed distribution do not affect our main qualitative results as described below.

To mimic the effects of a narrow channel whose width allows for neighbouring bacteria to swap positions, we proceed as follows. Consider a particle of type i after it has potentially tumbled and moved but before a new particle selection occurs. At this stage of the algorithm, we allow for *head-to-head crossing*, i.e. the particle will exchange positions with its neighbour with a probability p_s if, and only if, their directors point towards each other. Here p_s is a constant rate such that an increase in channel width corresponds to an increase in p_s. With this algorithm, out of all possible head-to-head crossing events involving particles of types i and j within a given time step, a fraction approximately equal to $p_s(v_i + v_j)$ will indeed occur, on average. The scaling with the speeds occurs because the number of particle updates is proportional to the speed. Here, we do not divide p_s by speed: it is physically reasonable that the effective swapping rates are proportional to the relative speed since the self-propulsion forces of bacteria are typically proportional to their self-propulsion

speeds. Similarly, we allow for *overtaking*, i.e. particle i will swap positions with its neighbouring particle j with a probability proportional to $v_i - v_j$ if, and only if, their directors point in the same directions and i is behind j with $v_i > v_j$. Since in our algorithm the possibility of overtaking is considered only when the faster particle is selected, we set the nominal overtaking swapping rate to be $p_s(v_i - v_j)/v_i$: otherwise the effective swapping rate would scale with the square of the speeds. In summary, faster particles can overtake slower ones at an effective rate proportional to their relative speed.

III. SIMULATIONS

Our stationary numerical results were obtained from simulations with periodic boundary conditions and $N = 2000$ sites, except where a proper sampling of the swim speed distribution requires a larger system, in which case we used $N = 10^4$. For visualization, we recorded snapshots of the steady-state system for 300 successive time steps after $t = 10^7$. The CSD and other similarly averaged quantities were calculated from 9000 uncorrelated configurations recorded every 10^4 time steps from $t = 10^7$ onwards, within the same simulation.

A. Binary mixture

We start by analysing the binary mixture model for 50-50% global composition with $p_s = 0$, i.e. no particle swapping. Fig. 2 shows snapshots of a section of the 1D simulated system at successive times within the stationary state for different degrees of bidispersity δ. From a homogeneous initial state, particles start to trap each other and form clusters, reaching a steady state characterized by the CSD. For $\delta = 0$, all particles are identical and thus have the same tendency to form clusters. In this case, the average cluster size is given by $L_c \approx \sqrt{2v\phi/[\alpha(1 - \phi)]}$ and the CSD is proportional to $\exp\{(-l/L_c)\}$ [2]. For $\delta > 0$, there are two groups of particles, each with a different tendency to form clusters. Distinct typical slopes in the space-time plot indicate distinct swim speeds. Since particle swapping is not yet allowed, the random sequence of particle types does not change in time. The probability to have many particles of the same type at successive positions is vanishingly small, except for small clusters, which are dominated by slower particles as they are more likely to tumble before being absorbed by a cluster.

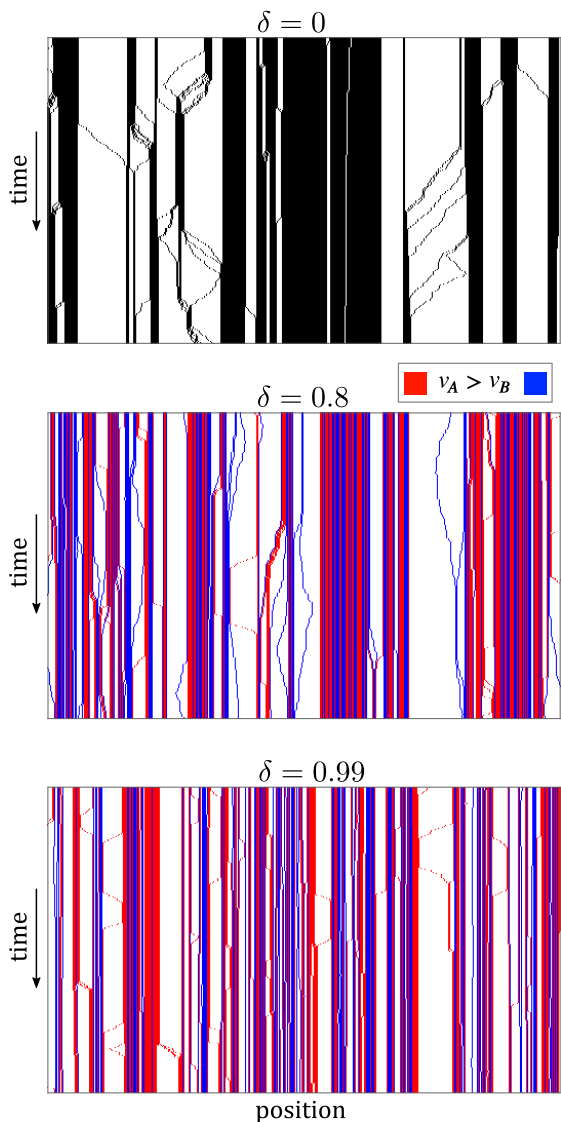The CSD is defined as the average number of clusters of size l and denoted by $F_c(l)$. Fig. 3a

FIG. 2. As Fig. 1a but for the steady state of a binary system of run-and-tumble particles on a 1D lattice for different bidispersities δ. Only 500 of $N = 2000$ simulated sites are shown. The monodisperse case ($\delta = 0$) is shown for comparison. In the bidisperse cases ($\delta > 0$), particles with higher (lower) swim speed are in red (blue). The global average speed, tumbling rate, concentration, and composition of the entire simulated system are $\langle v \rangle = v_0 = 1$, $\alpha = 0.01$, $\phi = 0.5$, and 50-50%, respectively.

shows that as δ increases the CSD moves towards smaller clusters, even though the global average speed $\langle v \rangle = v_0$ is fixed. The exponential shape of the CSD remains preserved (which allows us to map onto a monodisperse system; see Section IV). Fig. 3b shows that L_c indeed decreases with δ. The inset of Fig. 3b has the ratio between the bidisperse and monodisperse values of L_c. This ratio quantifies the clustering reduction by speed diversity.

It does not depend on ϕ (for $\delta \approx 1$, a small dependence on ϕ is found, indicating that additional simulation statistics would be required). Similarly, it does not depend on α (data not shown).
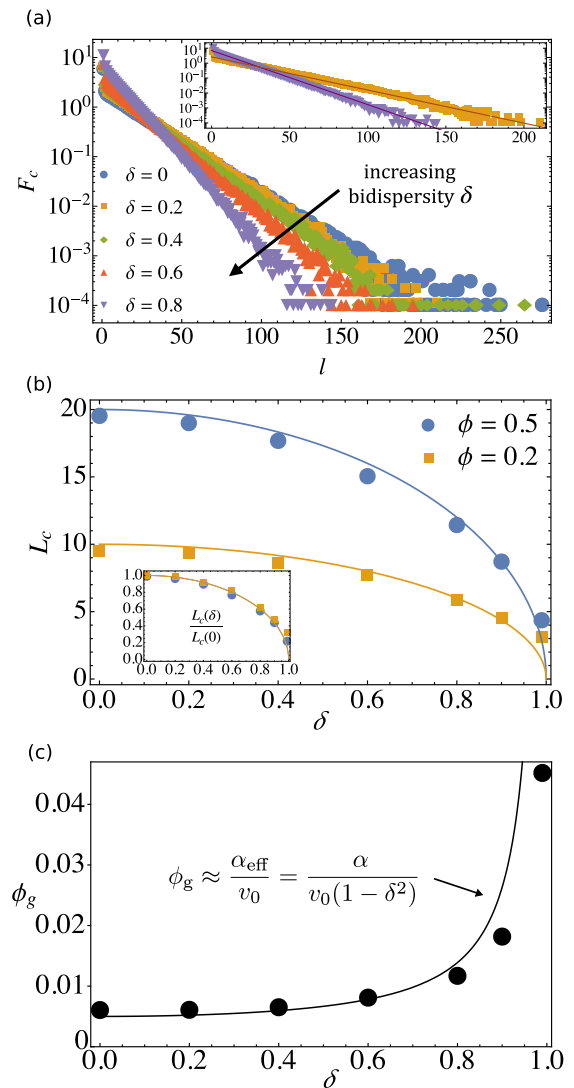

FIG. 3. (a) Cluster size distribution (log scale) from simulations for various bidispersities δ, with $\langle v \rangle = 1$, $\alpha = 0.005$, $\phi = 0.5$, $N = 2000$, and $p_s = 0$. Global composition: 50-50%. The inset shows the cases $\delta = 0.2$ and $\delta = 0.8$ (with the same plot markers as in the main figure) compared against the corresponding theoretical results [eqn (11)]. (b) Average cluster size L_c versus bidispersity δ at fixed average swim speed, with other parameters as in (a). The points show the simulation results and the lines are the theoretical predictions from Section IV. The inset shows the ratio to the corresponding monodisperse case, which is independent of ϕ. (c) Gas concentration ϕ_g vs. δ with parameters as in (a). Points are the simulation results and the solid line is the theory as obtained by identifying an effective tumbling rate for the mixture.

Notice that clusters of size $l = 1$ correspond to isolated particles and therefore should, in principle, be considered as part of the "gas", i.e., not clusters. However, our theory in Section IV [2] relies on integrating quantities across all positive l. Thus, $l = 1$ is included in calculating L_c for a more appropriate comparison. In any case, since at low tumbling rates the gas density is typically small, the contribution to L_c from isolated particles is likewise very small, as confirmed numerically. The gas concentration ϕ_g, which is defined as the average particle concentration in the regions containing only isolated particles (i.e., the regions between clusters of size ≥ 2), increases with δ (see Fig. 3c) since at higher speed diversity fewer particles participate in clusters of size $l > 1$, as seen in Fig. 3a.

B. Binary mixture with particle swapping

With $p_s \neq 0$, particles have a higher chance to escape from clusters as now this can occur by either tumbling or swapping, where the latter includes both head-to-head crossing and overtaking. Thus, the tumbling rate is effectively increased (see Section IV) and therefore cluster formation is further reduced. Fig. 4 shows snapshots for different p_s values. The higher the swapping rate, the more the cluster sizes fluctuate in time. At high p_s, the clusters have been mostly destroyed.

With increasing p_s the CSDs recede towards low cluster sizes while approximately maintaining a purely exponential functional form. Fig. 5a shows L_c versus δ for various values of p_s, whereas Fig. 5b shows L_c versus p_s for fixed values of δ. For high p_s, the cluster size dependence on δ is negligible. In this scenario, the particles end up leaving the cluster sooner by swapping than by tumbling. Since tumbling is no longer important, the diversity in the ratio of speed to tumbling rate becomes irrelevant. As a result, the clustering process becomes controlled by swapping events. Fig. 5b shows that if overtaking is turned off but head-to-head crossing is kept on, one obtains higher values of L_c, but the overall behaviour with p_s is similar. The inset of Fig. 5b shows that the difference between these two swapping scenarios peaks at small values of p_s. This is because at high p_s bidispersity becomes irrelevant and therefore overtaking should not be important since its rate is zero without diversity.

C. Fully polydisperse mixture

For the fully polydisperse distribution of swim speeds in eqn (1), we keep $p_s = 0$ as particle swap-
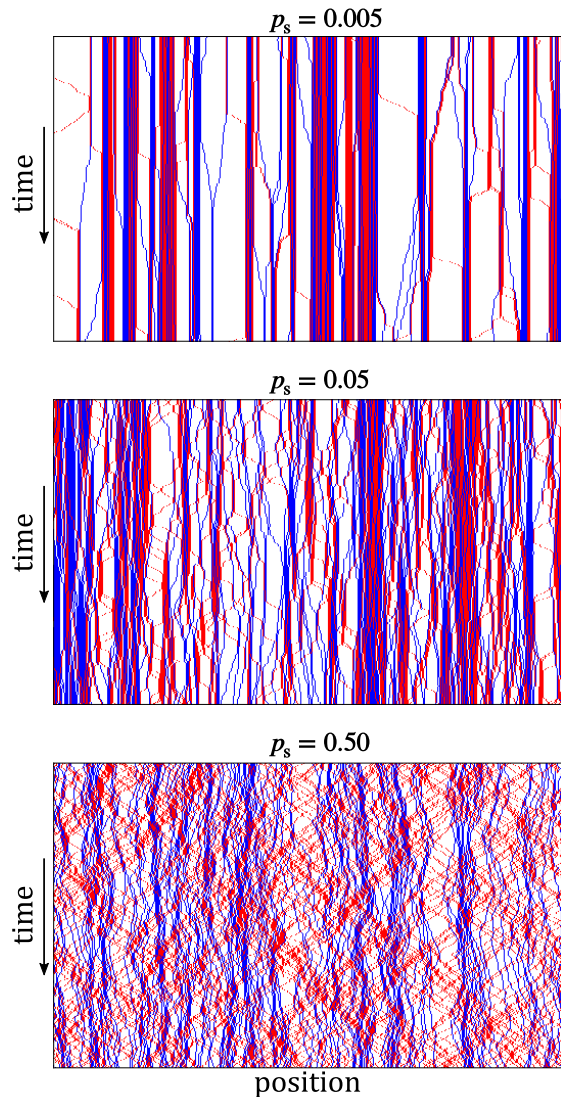

FIG. 4. As Fig. 2 but for fixed $\delta = 0.9$ and swapping rate $p_s > 0$ as indicated, with both particle head-to-head crossing and particle overtaking mechanisms turned on.

ping effects are analogous to those in the binary mixture. Fig. 6 shows the average cluster size. The higher the λ, the smaller are the clusters, at fixed $\langle v \rangle$. The CSD maintains an approximately purely exponential form becoming almost horizontal at sufficiently high λ (data not shown). Additional details for this case are discussed in Section IV.

IV. THEORY

This section is dedicated to deriving theoretical expressions for the average cluster size L_c which includes the effect of speed diversity and particle swapping, to be compared to our numerical results
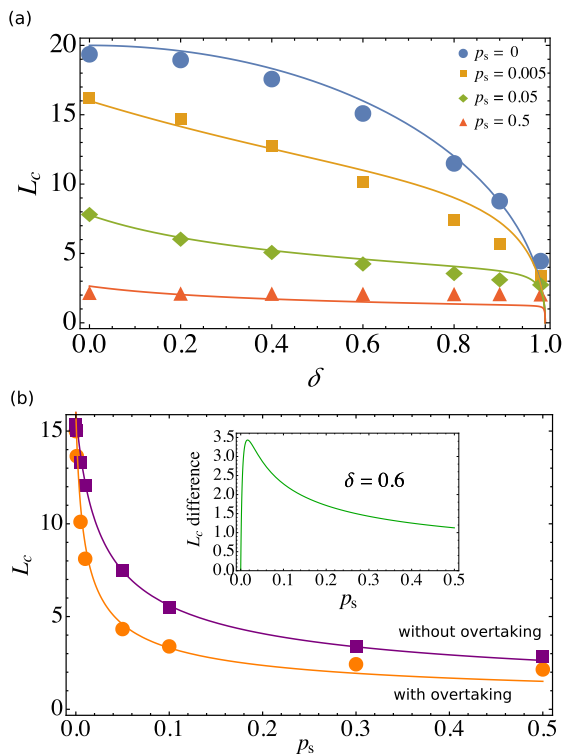
FIG. 5. (a) Average cluster size L_c versus bidispersity δ at fixed average speed for various particle swapping rates p_s, from simulations (points) and theoretical predictions for the CSD length scale L_c, eqn.(11) (lines). (b) Same as in (a) but as a function of $p_s \in \{0, 0.001, 0.005, 0.01, 0.05, 0.1, 0.3, 0.5\}$ and comparing the cases with and without particle overtaking. The inset shows the difference between the swapping scenarios. Other parameters as in Fig. 3a.

presented in Section III. We start by reviewing the developments of Ref. [2] (providing previously omitted details) for a lattice model with identical RT particles. We explicitly consider an arbitrary self-propulsion speed v instead of the case $v = 1$ of the original derivation. By doing so, we will be able to extend the results to speed-diverse systems. We conclude the section by incorporating particle swapping into the theory, considering both head-to-head crossing and overtaking.

A. Monodisperse systems

Following Ref. [2], we assume that the positions of the borders of a cluster, as well as the cluster itself, are independent random walkers, and that interactions between clusters are weak. Within this approximation (which corresponds to $\alpha/v \ll \phi$ and is used throughout the present work), the authors of Ref. [2] have calculated the distributions of sizes
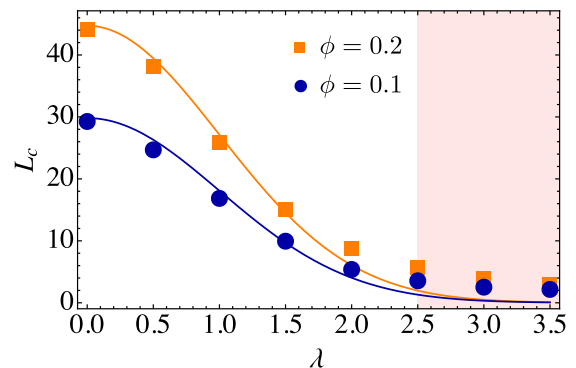

FIG. 6. Average cluster size L_c versus the polydispersity parameter λ at fixed average speed $\langle v \rangle = 1$ for $\phi = 0.1$ and $\phi = 0.2$. The points show the simulation results and the lines are the theoretical predictions for the CSD length scale L_c as given in eqn (11). For the simulations with $\lambda > 2.5$, the numerical results no longer agree with the theory, as discussed in Section IV. Other parameters: $\alpha = 2.5 \times 10^{-4}$, $v_{\max} = 4$, and $N = 10^4$.

for the clusters and gas regions, $F_c(l)$ and $F_g(l)$, respectively, in the one-component case. (For the purpose of calculating the distribution $F_g(l)$, the gas can be regarded as consisting of empty regions since the concentration of isolated particles is proportional to α; see below.) Such distributions were obtained by considering that the clustering process approaches a thermodynamic-like equilibrium where $F_c(l)$ maximizes an entropy for the number of possible cluster configurations. The resulting distributions follow an exponential decay with the region size l:

$$F_c(l) = A_c e^{-l/L_c} \quad \text{and} \quad F_g(l) = A_g e^{-l/L_g}. \quad (2)$$

$A_{c,g}$ and $L_{c,g}$ are the parameters of the distribution to be found, with $L_{c,g}$ corresponding to the stationary average cluster or gas region size. To find these parameters, we first invoke particle conservation, which implies that gas and cluster regions together cover the whole system, i.e.,

$$\sum_l l F_c(l) + \sum_l l F_g(l) = N. \quad (3)$$

Then, since we consider periodic boundary conditions, the number of gas regions must be equal to the number of cluster regions, which leads to

$$\sum_l F_c(l) = \sum_l F_g(l). \quad (4)$$

Moreover, since the the concentration of isolated particles is $\phi_g = \alpha/v$ as shown in Ref. [2], we assume that for high persistence, i.e., high v/α, one has $\phi_g \approx 0$. On the other hand, the particle concentration inside the cluster is $\phi_c = 1$ by definition.

Taking one typical cluster and one typical gas region, the total amount of particles reads

$$L_c\phi_c + L_g\phi_g = (L_c + L_g)\phi, \tag{5}$$

where, due to particle conservation, on the right hand side the size of the cluster and gas region taken together appears multiplied by the system's overall particle concentration ϕ.

We recall the assumption that cluster-cluster interactions are weak; in fact, these interactions are assumed to occur only through uncorrelated emissions of particles into the gas, and absorptions from there. This implies low tumbling rates ($\alpha/v \ll \phi$), consistently with the previous steps. In this regime, there are essentially no particles in each gas region ($\phi_g \approx 0$), meaning that these regions are almost empty. To close our system of equilibrium equations, we assume detailed balance (since the clustering process can be mapped onto a thermodynamic-like equilibrium process) and calculate the balance between production and evaporation of dimers (i.e., clusters of size $l = 2$). Their formation is controlled by emissions of particles at each boundary of a gas region. In order to compute the production rate of dimers, denoted by W_2^+, suppose that a particle at the right boundary of a gas region of typical size L_g is emitted from a cluster with velocity v. Such an emission occurs at rate $\alpha/2$. Because tumbling rates are small, we assume that the emitted particle will keep its velocity direction and reach the next cluster after a travel time L_g/v. For a new dimer to be created, a particle from the left boundary of the gas region must be emitted, also with probability $\alpha/2$ at each time step, before the first particle's arrival. As for the first particle, we assume that the second emitted particle will move ballistically, i.e., without tumbling, until it encounters the first particle. That is, if the second particle is emitted before the arrival of the first particle, then the dimer will form. For a particle to be emitted at a time τ, it must *not* be emitted before that, i.e., until a time $\tau - 1$. This happens with probability $(1 - \alpha/2)^{\tau-1}(\alpha/2)$. Therefore, the probability for a new dimer to form is given by

$$P_+ = \left(\frac{\alpha}{2}\right)\left[\left(\frac{\alpha}{2}\right)\sum_{\tau=0}^{L_g/v}(1-\alpha/2)^{\tau-1}\right],$$
$$\approx \left(\frac{\alpha}{2}\right)\left[1 - (1-\alpha/2)^{L_g/v}\right], \tag{6}$$

where the $\alpha/2$ outside the square brackets accounts for the emission of the first particle, whilst the terms inside the square brackets ensures that the second particle leaves its cluster before the travel time L_g/v. Note that, in order to approximate the expression, we used our assumption that α is small. The global production rate of dimers is then given by $W_2^+ = 2P_+ N_g$, where $N_g = \sum_l F_g(l)$ is the total number of gas regions and the factor two accounts for the spatially inverted case, where the particle on the left is emitted first. Putting this together yields

$$W_2^+ = \alpha\left[1 - (1 - \alpha/2)^{L_g/v}\right]N_g$$
$$\approx \frac{\alpha^2}{2}\frac{L_g}{v}N_g. \tag{7}$$

On the other hand, the rate of evaporation of dimers, W_2^-, can be calculated as follows. First, note that the only dimers that need to be considered are the ones composed of two particles facing each other, as otherwise the dimer would either have evaporated or be quickly absorbed by other clusters (and so it would not be noticed as a dimer in the steady-state statistics). That is, a dimer is destroyed whenever one of its constituents tumbles to a different direction, which occurs with probability $P_- = \alpha/2$ for each of them, leading to a total probability of evaporation equal to $2P_-$. Then, the global rate of dimer evaporation is obtained by multiplying this total tumbling probability of the dimer by the total number of dimers $F_c(l = 2)$, i.e., $W_2^- = \alpha A_c e^{-2/L_c} \approx \alpha A_c$, where we used that $2/L_c \approx 0$. Therefore, the final equilibrium condition used to close our system of equations, $W_2^- = W_2^+$, amounts to

$$A_c \approx \frac{\alpha}{2}\frac{L_g}{v}N_g. \tag{8}$$

To solve the system of equations (3)-(8) analytically, we replace the summations by integrals from $l = 0$ to ∞. The solution is

$$A_c \approx \frac{N\alpha(1-\phi)}{2v} \text{ and } L_c \approx l_c(v, \alpha, \phi) \equiv \sqrt{\frac{2v\phi}{\alpha(1-\phi)}}, \tag{9}$$

where, due to the small α limit (i.e., $\phi_g \approx 0$), all instances of $v\phi - \alpha$ and $v - \alpha$ have been replaced by $v\phi$ and v, respectively.

B. Speed-diverse systems

To include speed diversity, let us consider a multi-component system with an arbitrary number of particle types Q, which are characterized by speeds v_i and are present at global concentrations ϕ_i with $i = 1, \ldots, Q$ and $\sum_i \phi_i = \phi$. In this case, the size distributions of clusters and gas regions are still given by exponentials. This fact can be verified either via simulations, as shown in Section III,

or by noticing that the original derivation of the CSD in Ref. [2] becomes specific to the case of identical particles only *after* the exponential shape is obtained. We thus have $F_c(l) = A_c e^{-l/L_c}$ with new constants A_c and L_c to determine in the speed diverse case. The only expression that requires alteration is the one that explicitly involves self-propulsion speed, that is, the rate of dimer creation. It can now be written as

$$W_2^+ \approx \sum_{ij} \alpha \left[1 - (1 - \alpha/2)^{L_g/v_j} \right] N_g^{ij}, \quad (10)$$

where the summation runs over particle types and N_g^{ij} is the *total* number of gas regions bounded by particles of types i and j on the left and right borders, respectively. (Since N_g^{ij} is symmetric in i and j, the term $(1 - \alpha/2)^{L_g/v_j}$ does not need to be symmetrized.) A cluster border will be occupied by a particle of type i with a probability that depends only on the global concentration ϕ_i and α, independently of v_i. As a result, the chance to find a gas region simultaneously bounded by types i and j obeys $N_g^{ij} \sim \phi_i \phi_j$, which is then normalized so that $\sum_{ij} N_g^{ij}$ gives the total number of gas regions. On the other hand, W_2^- remains unchanged since it is speed-independent. The resulting CSD parameters can be expressed, using the function l_c defined in (9), as

$$A_c \approx \frac{N\alpha(1 - \phi)}{2v_{\text{eff}}} \quad \text{and} \quad L_c \approx l_c(v_{\text{eff}}, \alpha, \phi),$$
$$\text{with} \quad v_{\text{eff}} = \left(\sum_i \frac{\phi_i}{\phi v_i} \right)^{-1} \quad (11)$$

a *monodisperse* effective speed which is found to be the component-weighted *harmonic* mean speed of the mixture. As such, v_{eff} encodes speed diversity entirely and is the effective self-propulsion speed in the monodisperse case that gives the same CSD as in the speed-diverse case.

We briefly remind the reader that if a vehicle travels a certain distance at speed v_1 and returns the *same* distance at speed v_2, then its average speed is the harmonic mean of v_1 and v_2, not the arithmetic mean. The total travel time is the same as if it had travelled the whole distance at that average speed. However, if the vehicle travels for a certain amount of *time* at speed v_1 and then the same amount of time at a speed v_2, then its average speed is the arithmetic mean of v_1 and v_2. Therefore, our result can be understood as follows. The average cluster size is ultimately set by the time that an arbitrary particle takes to cross a typical gas region. This time is an arithmetic average over the times taken by each particle type, where we consider that the chance of having a particle of type i travelling in the gas is just proportional to ϕ_i. Thus, the average speed at which such typical fixed distance is covered is the harmonic average of the speeds, as given in eqn (11), not the arithmetic one.

For the 50-50% binary mixture, our derivation leads to

$$L_c^{\text{bi}} \approx \sqrt{\frac{2v_0(1 - \delta^2)\phi}{\alpha(1 - \phi)}}, \quad (12)$$

and, for the fully polydisperse mixtures, the result is

$$L_c^{\text{poly}} \approx \sqrt{\frac{2v_0\phi}{\alpha(1 - \phi)}} e^{-\frac{\lambda^2}{4}} = \sqrt{\frac{2\langle v \rangle \phi}{\alpha(1 - \phi)}} e^{-\frac{\lambda^2}{2}}, \quad (13)$$

where we have recalled that in the fully polydisperse case $\langle v \rangle = v_0 \exp\{(\lambda^2/2)\}$. Expressions (12) and (13) are in excellent agreement with the simulation results as presented in Section III. In particular, the theory predicts that the ratio between the speed-diverse and monodisperse values of L_c does not depend on ϕ or α as shown in the inset of Fig. 3b.

The gas concentration, albeit small and taken to be zero in the analytical derivation of Section IV A, can be written using the monodisperse effective speed v_{eff} and extending the monodisperse expression [2] to obtain $\phi_g = \alpha/v_{\text{eff}}$ or, equivalently, $\phi_g = \alpha_{\text{eff}}/\langle v \rangle$, where $\alpha_{\text{eff}} \equiv \alpha\langle v \rangle/v_{\text{eff}}$ is the effective tumbling rate in the monodisperse case that gives the same CSD as in the speed-diverse case. The resulting expression is highly accurate, as shown in Fig. 3c. Our theory in (11) becomes less accurate at high speed diversity as a large number of very slow particles arise, in which case the assumption $\alpha/v \ll \phi$ is no longer valid.

C. Swapping

Particle swapping is taken into account by adding appropriate terms to the monodisperse effective tumbling rate α_{eff}. This is because particle swapping enables an additional mechanism for cluster-border evaporation, other than tumbling. For simplicity, we consider only the 50-50% binary mixture as other cases are analogous. A satisfactory approximation for head-to-head crossing effects is to consider $\alpha_{\text{eff}} \to \alpha_{\text{eff}} + \kappa_{\text{HC}} p_s$, where κ_{HC} is a parameter proportional to the fraction of pairs susceptible to head-to-head crossing. In principle, it depends on p_s, too. But, to first order in p_s, it can be considered a constant. Because head-to-head crossing rates are proportional to the relative speed between the particle types, there should be a

corresponding dependence on δ, but in the 50-50% binary mixture this cancels out because the average of $v_i + v_j$ across particle types is independent of δ. By fitting data without overtaking for several values of δ and p_s, we find $\kappa_{HC} \approx 0.56$.

Proceeding similarly for the case with both head-to-head crossing and overtaking mechanisms, we use

$$\alpha_{eff} \to \alpha_{eff} + \kappa_{HC} p_s + \kappa_T p_s \delta, \qquad (14)$$

where κ_T is taken as a constant and the new term is linear in δ since overtaking depends on the relative speed only between the particles of faster-behind-slower pairs. Fitting data from simulations with overtaking, we obtain $\kappa_T \approx 1.97$. As shown in Fig. 5, this approach provides good results.

V. CONCLUSIONS AND DISCUSSION

In this work we showed how motility-induced self-clustering in confined active matter can be reduced by diversity of self-propulsion speeds. Also, cluster sizes are further reduced by confinement relaxation. We used a minimal quasi-1D discrete lattice model of run-and-tumble particles with a distribution of self-propulsion speeds. Neighbouring particles were allowed to perform head-to-head crossing and overtaking at rates proportional to their relative speeds, depending on whether they are face-to-face oriented or there is a faster particle behind a slower one. This mimics a narrow channel whose width is large enough to allow for some position swapping events. A binary mixture and a fully polydisperse system were studied. Without swapping, the average cluster size L_c *decreases* with diversity. This is equivalent to a system of identical particles whose speed is the *harmonic* mean speed of the mixture. With swapping, particles can escape from clusters more quickly. Consequently, L_c decreases with swapping rates and depends less strongly on diversity. At sufficiently high swapping rates, clustering then becomes controlled by head-to-head crossing and overtaking events and thus speed diversity becomes irrelevant. We derived an accurate dynamical equilibrium theory for the CSDs and gas concentrations that is applicable to all models studied here.

In order to calculate L_c for the mixture, at first glance one could be tempted to insert the arithmetic mean self-propulsion speed of the mixture into one-component theories or even to take the arithmetic mean of the one-component L_c's. Here we find that both ideas provide significantly wrong results (more below). Unlike the case of tumbling-rate diversity, we highlight that the above results

for speed diversity imply that the average cluster size follows a non-arithmetic generalized power mean of the one-component L_c's. In fact, our average cluster size result in eqn (11) can be rewritten as

$$L_c \approx \left(\sum_i \frac{\phi_i}{\phi} L_i^p \right)^{1/p} \quad \text{with} \quad L_i = \sqrt{\frac{2 v_i \phi}{\alpha(1 - \phi)}} \tag{15}$$

and the exponent $p = -2$. This is different from the case of tumbling-rate diversity previously studied where we had $p = 1$, that is, the arithmetic mean [30]. In fact, although for speed diversity an arithmetic mean of the one-component L_c's would qualitatively capture a decrease in the mixture's L_c, it is quantitatively wrong. Also, notice that the arithmetic mean result for tumbling-rate diversity can in principle also be obtained by the analytical method described in Section IV, but in that case we no longer have $N_g^{ij} \sim \phi_i \phi_j$ as the number of gas regions bounded by a certain particle type depends non-trivially on that particle's tumbling rate (still, we have inserted numerical results for N_g^{ij} and obtained accurate values for L_c; data not shown).

For the parameters investigated, our data do not show strong "fractionation", i.e., clusters with compositions that differ from the global composition. For small clusters, however, we do observe a systematic deviation: small clusters are typically richer in slower particles (data not shown). This is because faster particles travel too fast between clusters to allow for the emission of another particle that could meet it to form a dimer (and then other small cluster sizes). Instead, the first emitted particle is quickly reabsorbed by another cluster before a second particle is emitted. Investigating exactly which parameters would generate strong fractionation in 1D is beyond the scope of our work. This question would be more relevant in 2D, where one might speculate, for instance, that the faster particles would dominate the interior of big clusters since they would be more likely to occupy newly available vacancies.

To estimate whether our qualitative clustering behaviour is also valid for other models of active particles that are not on-lattice or run-and-tumble, one can look at the answer to this question for monodisperse systems. For off-lattice models, simulations show that monodisperse systems of run-and-tumble bacteria, active Brownian particles, and active Ornstein-Uhlenbeck particles, have an average cluster size given by $L_c \sim \sqrt{\rho u / \omega}$, where ρ is the off-lattice model global concentration, u is the active speed, and ω is an inverse persistence time parameter [11]. This is precisely the off-lattice version of the length scale expression

(9). Consequently, our qualitative results for mixtures are indeed expected to be universal across the most used models in 1D scalar active matter.

The clustering reduction effects presented here can indeed be relevant in biological experimental situations. Consider the unicellular microorganisms known as *choanoflagellates*. Although individually they are better described as smooth, active Brownian swimmers, their clustering behaviour is expected to be equivalent to that of run-and-tumble particles because of the equivalence between these models with respect to collective phenomena [70]. These cells naturally differentiate into slower and faster particles with a speed-diversity standard deviation that can be obtained from literature experimental data as in Ref. [33]. One can then calculate the error in estimating their L_c via one-component models that use the arithmetic mean speed instead of the harmonic mean speed, leading to $\Delta L_c \equiv [l_c(v_{\text{eff}}, \alpha, \phi) - l_c(\langle v \rangle, \alpha, \phi)]/l_c(v_{\text{eff}}, \alpha, \phi) \approx 30\%$.

In future work, it would be relevant to include features such as propulsion mechanisms that are so strong that the particles collectively move and merge clusters. In this kind of scenario, one expects that, for less soft active particles, swapping could become less frequent in narrow channels; at the same time, pushing effects due to direct contact would be enhanced. Also, one could consider lattice models with rules to mimic hydrodynamic interactions, which can substantially change cluster escape times [27]. Another interesting future avenue is the case of direction-dependent speeds arising due to external forces.

AUTHOR CONTRIBUTIONS

PdC contributed with conceptualization, simulations, theoretical analyses and the writing of the original draft. SD and FMR also helped with the conceptualization and developed most of the analytical derivations. All authors contributed equally to interpreting the simulation data, developing the main ideas of the theory and revising and editing the manuscript.

CONFLICTS OF INTEREST

There are no conflicts to declare.

ACKNOWLEDGEMENTS

[1] M. E. Cates and J. Tailleur, *Annu. Rev. Condens. Matter Phys.*, 2015, **6**, 219–244.

[2] R. Soto and R. Golestanian, *Physical Review E*, 2014, **89**, 012706.

[3] N. Sepúlveda and R. Soto, *Physical Review E*, 2016, **94**, 022603.

[4] A. Slowman, M. Evans and R. Blythe, *Physical review letters*, 2016, **116**, 218101.

[5] F. Ginot, I. Theurkauff, F. Detcheverry, C. Ybert and C. Cottin-Bizonne, *Nature Communications*, 2018, **9**, 1–9.

[6] G. S. Redner, C. G. Wagner, A. Baskaran and M. F. Hagan, *Physical Review Letters*, 2016, **117**, 148002.

[7] J. I. Quelas, M. J. Althabegoiti, C. Jimenez-Sanchez, A. A. Melgarejo, V. I. Marconi, E. J. Mongiardini, S. A. Trejo, F. Mengucci, J.-J. Ortega-Calvo and A. R. Lodeiro, *Scientific Reports*, 2016, **6**, 23841.

[8] L. Ranjard and A. Richaume, *Research in microbiology*, 2001, **152**, 707–716.

[9] J. Männik, R. Driessen, P. Galajda, J. E. Keymer and C. Dekker, *Proceedings of the National Academy of Sciences*, 2009, **106**, 14861–14866.

[10] L. E. Fuentes-Ramirez and J. Caballero-Mellado, in *PGPR: Biocontrol and biofertilization*, Springer, 2005, pp. 143–172.

[11] P. Dolai, A. Das, A. Kundu, C. Dasgupta, A. Dhar and K. V. Kumar, *Soft Matter*, 2020, **16**, 7077–7087.

[12] R. Dandekar, S. Chakraborty and R. Rajesh, *arXiv preprint arXiv:2006.05980*, 2020.

[13] C. Vanhille Campos, F. Alarcón Oseguera, I. Pagonabarraga, R. Brito and C. Valeriani, *arXiv preprint arXiv:1912.01282*, 2019.

[14] L. Caprini and U. M. B. Marconi, *Physical Review Research*, 2020, **2**, 033518.

[15] Y. Li, P. K. Ghosh, F. Marchesoni and B. Li, *Phys. Rev. E*, 2014, **90**, 062301.

[16] X. Ao, P. K. Ghosh, Y. Li, G. Schmid, P. Hänggi and F. Marchesoni, *The European Physical Journal Special Topics*, 2014, **223**, 3227–3242.

[17] P. K. Ghosh, P. Hänggi, F. Marchesoni and F. Nori, *Physical Review E*, 2014, **89**, 062115.

[18] X. Ao, P. K. Ghosh, Y. Li, G. Schmid, P. Hänggi and F. Marchesoni, *EPL (Europhysics Letters)*, 2015, **109**, 10003.

[19] L. Barberis and F. Peruani, *The Journal of Chemical Physics*, 2019, **150**, 144905.

[20] P. Illien, C. de Blois, Y. Liu, M. N. van der Linden and O. Dauchot, *Physical Review E*, 2020, **101**, 040602.

[21] F. Peruani, J. Starruß, V. Jakovljevic, L. Søgaard-Andersen, A. Deutsch and M. Bär, *Physical Review Letters*, 2012, **108**, 098102.

[22] H.-P. Zhang, A. Be'er, E.-L. Florin and H. L. Swinney, *Proceedings of the National Academy of Sciences*, 2010, **107**, 13626–13630.

[23] A. Be'er, B. Ilkanaiv, R. Gross, D. B. Kearns, S. Heidenreich, M. Bär and G. Ariel, *Communications Physics*, 2020, **3**, 1–8.

[24] J. E. Keymer, P. Galajda, G. Lambert, D. Liao and R. H. Austin, *Proceedings of the National Academy of Sciences*, 2008, **105**, 20269–20273.

[25] I. Buttinoni, J. Bialké, F. Kümmel, H. Löwen, C. Bechinger and T. Speck, *Physical Review Letters*, 2013, **110**, 238301.

[26] D. Levis and L. Berthier, *Physical Review E*, 2014, **89**, 062301.

[27] F. Alarcón, C. Valeriani and I. Pagonabarraga, *Soft matter*, 2017, **13**, 814–826.

[28] A. Villa-Torrealba, C. Chávez-Raby, P. de Castro and R. Soto, *Phys. Rev. E*, 2020, **101**, 062607.

[29] H. C. Berg, *E. coli in Motion*, Springer Science & Business Media, 2008.

[30] P. de Castro, S. Diles, R. Soto and P. Sollich, *Soft Matter*, 2021, **17**, 2050–2061.

[31] C. Bechinger, R. Di Leonardo, H. Löwen, C. Reichhardt, G. Volpe and G. Volpe, *Reviews of Modern Physics*, 2016, **88**, 045006.

[32] E. P. Ipiña, S. Otte, R. Pontier-Bres, D. Czerucka and F. Peruani, *Nature Physics*, 2019, **15**, 610–615.

[33] J. Sparacino, G. L. Miño, A. J. Banchio and V. Marconi, *Journal of Physics D: Applied Physics*, 2020, **53**, 505403.

[34] I. Berdakin, A. V. Silhanek, H. N. M. Cortéz, V. I. Marconi and C. A. Condat, *Central European Journal of Physics*, 2013, **11**, 1653–1661.

[35] P. de Castro and P. Sollich, *Phys. Chem. Chem. Phys.*, 2017, **19**, 22509–22527.

[36] P. B. Warren, *Physical Chemistry Chemical Physics*, 1999, **1**, 2197–2202.

[37] P. de Castro and P. Sollich, *The Journal of Chemical Physics*, 2018, **149**, 204902.

[38] P. de Castro and P. Sollich, *Soft Matter*, 2019, **15**, 9287–9299.

[39] P. S. de Castro Melo, *Phase separation of polydisperse fluids*, King's College London, 2019.

[40] J. Stenhammar, R. Wittkowski, D. Marenduzzo and M. E. Cates, *Physical Review Letters*, 2015, **114**, 018301.

[41] T. Kolb and D. Klotsa, *Soft Matter*, 2020, **16**, 1967–1978.

[42] C. Hoell, H. Löwen and A. M. Menzel, *The Journal of Chemical Physics*, 2019, **151**, 064902.

[43] R. Wittkowski, J. Stenhammar and M. E. Cates, *New Journal of Physics*, 2017, **19**, 105003.

[44] S. C. Takatori and J. F. Brady, *Soft Matter*, 2015, **11**, 7920–7931.

[45] A. Grosberg and J.-F. Joanny, *Physical Review E*, 2015, **92**, 032118.

[46] A. Curatolo, N. Zhou, Y. Zhao, C. Liu, A. Daerr, J. Tailleur and J. Huang, *Nature Physics*, 2020, 1–6.

[47] Y. Wang, Z. Shen, Y. Xia, G. Feng and W. Tian, *Chinese Physics B*, 2020, **29**, 053103.

[48] B. van der Meer, V. Prymidis, M. Dijkstra and L. Filion, *The Journal of Chemical Physics*, 2020, **152**, 144901.

[49] P. Dolai, A. Simha and S. Mishra, *Soft Matter*, 2018, **14**, 6137–6145.

[50] S.-T. Lin, P. K. Maiti *et al.*, *arXiv preprint arXiv:2109.00415*, 2021.

[51] L. Angelani, *Journal of Physics A: Mathematical and Theoretical*, 2017, **50**, 325601.

[52] A. Costanzo, R. Di Leonardo, G. Ruocco and L. Angelani, *Journal of Physics: Condensed Matter*, 2012, **24**, 065101.

[53] A. Costanzo, J. Elgeti, T. Auth, G. Gompper and M. Ripoll, *EPL (Europhysics Letters)*, 2014, **107**, 36003.

[54] J.-C. Wu, K. Lv, W.-W. Zhao and B.-Q. Ai, *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 2018, **28**, 123102.

[55] Z. Wang, K.-C. Chu, H.-K. Tsao and Y.-J. Sheng, *Physical Chemistry Chemical Physics*, 2021, **23**, 16234–16241.

[56] L. Caprini, C. Maggi and U. Marini Bettolo Marconi, *The Journal of Chemical Physics*, 2021, **154**, 244901.

[57] A. Daddi-Moussa-Ider, M. Lisicki, A. J. Mathijssen, C. Hoell, S. Goh, J. Bławzdziewicz, A. M. Menzel and H. Löwen, *Journal of Physics: Condensed Matter*, 2018, **30**, 254004.

[58] K. Bisht and R. Marathe, *Physical Review E*, 2020, **101**, 042409.

[59] J. Männik, F. Sekhavati, J. Keymer and C. Dekker, Proc of MicroTAS, 2010, pp. 1346–1348.

[60] H. Wioland, E. Lushi and R. E. Goldstein, *New Journal of Physics*, 2016, **18**, 075002.

[61] E. Locatelli, F. Baldovin, E. Orlandini and M. Pierno, *Physical Review E*, 2015, **91**, 022109.

[62] S. Ketzetzi, M. Rinaldin, P. Dröge, J. de Graaf and D. J. Kraft, *arXiv preprint arXiv:2103.07335*, 2021.

[63] J. Aguilar, D. Monaenkova, V. Linevich, W. Savoie, B. Dutta, H.-S. Kuan, M. Betterton, M. Goodisman and D. Goldman, *Science*, 2018, **361**, 672–677.

[64] N. Gravish, G. Gold, A. Zangwill, M. A. Goodisman and D. I. Goldman, *Soft matter*, 2015, **11**, 6552–6561.

[65] N. Figueroa-Morales, R. Soto, G. Junot, T. Darnige, C. Douarche, V. A. Martinez, A. Lindner and É. Clément, *Physical Review X*, 2020, **10**, 021004.

[66] U. M. B. Marconi, N. Gnan, M. Paoluzzi, C. Maggi and R. Di Leonardo, *Scientific reports*, 2016, **6**, 1–9.

[67] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen and O. Shochet, *Physical Review Letters*, 1995, **75**, 1226.

[68] L. Caprini, U. M. B. Marconi and A. Puglisi, *Physical review letters*, 2020, **124**, 078001.

[69] D. Debnath, P. K. Ghosh, V. R. Misko, Y. Li, F. Marchesoni and F. Nori, *Nanoscale*, 2020, **12**, 9717–9726.

[70] M. E. Cates and J. Tailleur, *EPL (Europhysics Letters)*, 2013, **101**, 20010.