



Cite this: *Soft Matter*, 2021,
17, 5613

Motility-induced inter-particle correlations and dynamics: a microscopic approach for active Brownian particles

J. K. G. Dhont, ^{ab} G. W. Park ^a and W. J. Briels ^{ac}

Amongst the theoretical approaches towards the dynamics and phase behaviour of suspensions of active Brownian particles (ABPs), no attempt has been made to specify the motility-induced inter-particle correlations as quantified by the pair-correlation function. Here, we derive expressions for the pair-correlation function for ABPs with very short-ranged direct interactions for small and large swimming velocities and low concentrations. The pair-correlation function is the solution of a differential equation that is obtained from the Fokker–Planck equation for the probability density function of the positions and orientations of the ABPs. For large swimming Peclet numbers, λ , the pair-correlation function is highly asymmetric. The pair-correlation function attains a large value, $\sim \lambda$, within a small region of spatial extent, $\sim 1/\lambda$, near contact of the ABPs when the ABPs approach each other. The pair-correlation function is small within a large region of spatial extent, $\sim \lambda^{1/3}$, when the ABPs move apart, with a contact value that is essentially zero. From the explicit expressions for the pair-correlation function, Fick's diffusion equation is generalized to include motility. It is shown that mass transport, in case of large swimming velocities, is dominated by a preferred swimming direction that is induced by concentration gradients. The expression for the pair-correlation function derived in this paper could serve as a starting point to obtain approximate results for high concentrations, which could then be employed in a first-principles analysis of the dynamics and phase behaviour of ABPs at higher concentrations.

Received 18th March 2021,
Accepted 18th April 2021

DOI: 10.1039/d1sm00426c

rsc.li/soft-matter-journal

1 Introduction

Active matter consisting of suspensions of self-propelling particles ranges from synthetic colloidal particles to cells and bacteria. Several different types of synthetic colloidal particles have been developed over the past few years with a variety of propulsion mechanisms. Self-propulsion affects the dynamics and stability of phases, induces new phases with no counterparts in systems of passive particles, and affects phase separation kinetics. For overviews on the various types of propulsion mechanisms and motility-induced phase separation (MIPS), see, for example, ref. 1 (an introduction in German) and ref. 2–8.

The inter-particle correlations that underly the behaviour of active systems originate from direct and hydrodynamic interactions between the particles. The central quantity that

specifies the degree of correlations between particles is the so-called pair-correlation function (see, for example, ref. 9 and 10). Since propulsion is achieved by forces exerted by the particles onto the surrounding fluid, there are additional contributions to their hydrodynamic interactions as compared to passive particles. These additional hydrodynamic interactions depend on the propulsion details and lead to highly anisotropic attractive/repulsive interactions depending on the relative distance between the particles (see, for example, ref. 8 and 11–16). Simulations in 3D on spherical active particles show that the differences in hydrodynamic interactions between pullers and pushers have a pronounced effect on the pair-correlation function at close contact (with a center-to-center distance of $2.01a$, where a is the radius of the spheres), but a much smaller effect for somewhat larger distances (equal to $2.50a$).¹⁷ The effect of hydrodynamic interactions on inter-particle correlations has been studied for model flagellum swimmers by means of simulations in ref. 18, where increased correlations between the rear parts of swimmers and decreased correlations between the rear and the front parts are found. Strong inter-particle correlations due to long-ranged hydrodynamic interactions in 3D have been observed in large-scale simulations in ref. 19,

^a Institute of Biological Information Processing, IBI-4, Biomacromolecular Systems and Processes, Forschungszentrum Jülich GmbH, D-52428 Jülich, Germany.
E-mail: j.k.g.dhont@fz-juelich.de

^b Heinrich Heine Universität, 40225 Düsseldorf, Germany

^c MESA+ Institute for Nanotechnology, University of Twente, P. O. Box 217, 7500 AE Enschede, The Netherlands. E-mail: w.j.briels@utwente.nl



where direct interactions (for example, due to excluded-volume interactions) have been neglected. In general, self-propelled particles with a solid core, interacting hydrodynamically in addition to hard-core interactions, are referred to as squirmers. For comparison, active Brownian particles (ABPs) are introduced that do not exhibit hydrodynamic interactions. In the present paper, we study inter-particle correlations between ABPs, thus neglecting hydrodynamic interactions. It remains a largely open question to which extent inter-particle correlations of spherical ABPs and squirmers with excluded-volume interactions differ in 3D bulk.

A microscopic statistical mechanics approach towards the behaviour of active systems can be based on the so-called Smoluchowski equation, which is a Fokker–Planck-type equation of motion for the probability density function of the positions and orientations of the active particles. This approach can be used, for example, to generalize the van't Hoff law for the osmotic pressure²⁰ and expressions for Casimir forces between two plates²¹ to active systems, and explain the alignment of active colloids in a gravitational field.²² These generalizations aim at dilute active systems, and therefore depart from the Smoluchowski equation neglecting inter-particle correlations. For concentrated systems, however, such inter-particle correlations are essential.

An approach that addresses concentrated systems on a microscopic level is dynamic density functional theory (DDFT)²³ (for an overview on DDFT approaches, see ref. 24). This approach is based on the equation of motion for the one-particle distribution function as obtained from the Smoluchowski equation. Approximations are made in ref. 23 to eliminate three-particle correlation functions which arise from the coupling between the hydrodynamic and direct interactions, in favour of two-particle correlation functions. The two-particle correlation functions are then largely ignored, which is stated to be a good approximation for weak and long-ranged direct interactions. In the present paper, we are interested in the pair-correlation function for very steep and short-ranged interactions between ABPs, for which the DDFT approach in ref. 23 is inappropriate.

The Smoluchowski equation has been employed, within certain approximations, to analyse the stability of MIPS in 2D.^{25–29} Here, a linear decrease of the average velocity, $v \sim v_0(1 - \zeta \rho)$ (where v_0 is the free-swimming velocity), of swimmers with the increasing concentration, ρ , is found, which is in accordance with semi-empirical approaches^{30,32–34} and simulations in 2D.³⁵ Without the specification of the pair-correlation function, however, the coefficient ζ remains an unknown parameter. In ref. 36 and 37, the analysis of MIPS for non-chiral and chiral ABPs in 2D is similarly based on the Smoluchowski equation, where inter-particle correlations are accounted for by numerical input obtained from simulations.³⁸ An independent calculation of the pair-correlation function is not pursued in these studies. It is the purpose of the present study to derive explicit expressions for the 3D pair-correlation function, for small concentrations where binary interactions are dominant. For an overview of several theoretical approaches, see ref. 27 and 39.

In ref. 40, an expression for the pair-correlation function has been derived on the two-particle level from a model where orientational motion is mimicked by introducing a Gaussian type of noise imposed on the propulsion velocity with a finite relaxation time. This allows for a description on the basis of position coordinates alone, thus eliminating orientational variables. Within this approach, it is possible to extract an effective potential, the Boltzmann exponent of which gives the low-density pair-correlation function. As will turn out, within the microscopic approach in the present work, where orientational motion is explicitly accounted for, it is not possible to represent activity by means of an effective pair-interaction potential.

The pair-correlation function (averaged with respect to the orientation of one of the particles) for a dilute 2D system of Brownian Janus particles with induced-charge electrophoresis activity is experimentally obtained in ref. 41, and similarly for externally driven binary mixtures in ref. 42. Good agreement is found in ref. 41 with numerical solutions of a differential equation for the pair-correlation function for weak inter-particle interactions, which is similar but not identical to the Smoluchowski equation in 2D (the difference is discussed in Section 3). It is stated in ref. 41 that “the depletion wake consists of two depletion wings”, which we will comment on at the end of Section 5.

An extensive overview of the current status and challenges concerning active systems can be found in ref. 43, and in ref. 44 with an emphasis on simulation methods.

Surely, the above-cited references represent only a biased fraction of the vast literature on active systems. There is, however, essentially no literature that addresses the analytical determination of the pair-correlation function. The aim of the present paper is to derive closed analytic expressions for the pair-correlation function for dilute 3D systems of spherical ABPs, where binary interactions dominate. It remains a future challenge to extend the theory given in the present paper to higher concentrations and to include hydrodynamic interactions.

This paper is organized as follows. In Section 2, we recapitulate the Smoluchowski equation for ABPs, the derivation of the equation of motion for the orientation-dependent local density from it, and the definition and relevance of the pair-correlation function for collective dynamics and MIPS. In Section 3, the fundamental differential equation for the pair-correlation function for small concentrations is derived, which is solved for small swimming Peclet numbers in Section 4, and for large Peclet numbers in Section 5. Section 6 discusses the collective diffusion to leading order in concentration for large Peclet numbers. The summary, conclusion, and outlook are given in Section 7.

2 Fokker–Planck approach and relevance of the pair-correlation function

A microscopic particle-based starting point to predict motility-induced collective motion and phase separation (MIPS) of active



Brownian particles (ABP) with a spherically shaped core is the N -particle Smoluchowski equation. This is a Fokker-Planck equation in the over-damped limit for the (non-equilibrium) probability density function (pdf), P_N , of the position coordinates, $\{\mathbf{r}_j\}$, and the orientations, $\{\hat{\mathbf{u}}_j\}$, of a total number of N ABPs within a given volume. The standard Smoluchowski equation is a generalized diffusion equation for the time-rate of change of P_N , which expresses probability conservation in terms of the divergence of a probability flux. There are several contributions to the probability flux. There is a contribution arising from translational Brownian motion, which gives rise to changes of position coordinates. The position coordinates also change due to direct inter-particle interactions, while rotational Brownian motion results in changes of the orientations. Including the flux due to active swimming, the Smoluchowski equation reads as follows:

$$\begin{aligned} \frac{\partial}{\partial t} P_N = & D_0 \sum_{j=1}^N \nabla_j \cdot [\nabla_j P_N + \beta P_N \nabla_j \Psi] \\ & + D_r \sum_{j=1}^N \hat{\mathcal{R}}_j^2 P_N - \sum_{j=1}^N \nabla_j \cdot [P_N v_0 \hat{\mathbf{u}}_j], \end{aligned} \quad (1)$$

where the various contributions are displayed in the same order as discussed above. Here, D_0 and D_r are the translational and rotational diffusion coefficients for a single sphere in an unbounded fluid, respectively, $\beta = 1/k_B T$ is the inverse thermal energy (where k_B is the Boltzmann constant and T is the temperature), ∇_j is the gradient operator with respect to the position coordinate, r_j , of the j th particle, and

$$\hat{\mathcal{R}}_j(\cdot) = \hat{\mathbf{u}}_j \times \nabla_{u_j}(\cdot), \quad (2)$$

is the rotation operator, with ∇_{u_j} the gradient operator with respect to the Cartesian coordinates of the unit vector, $\hat{\mathbf{u}}_j$, that specifies the orientation of the j th particle. The orientation defines the direction in which an ABP would actively move when submerged in an unbounded fluid without the presence of other ABPs. Furthermore, Ψ is the total potential interaction energy of the assembly of ABPs, which depends only on the position coordinates, not on orientations, and v_0 is the bare swimming velocity. The average swimming velocity in a concentrated dispersion of ABPs is different from v_0 as a result of interactions between the ABPs, the calculation of which requires the solution of the Smoluchowski equation (see Section 6).

The Smoluchowski equation assumes overdamped dynamics, where inertial effects are neglected, which is a valid approximation on time scales of the order of a few tens of nanoseconds for sizes of the ABPs not larger than a few microns. An analysis of inertial effects for larger ABPs can be found in ref. 45. What has been neglected in eqn (1) are hydrodynamic interactions between the swimmers. Including hydrodynamic interactions remains a future challenge, and would enable to make a distinction between, for example, pullers and pushers.

The formulation of the Smoluchowski equation given above relates to 3D. A 2D formulation leads to a different form of the rotational operators, but is most probably amenable to the same type of analysis given in the present 3D case. Most of

the experiments (like in ref. 41 and 42) have been conducted in 2D, where hydrodynamic interactions with the confining geometry will probably play a major role. Including such hydrodynamic interactions requires a major extension of the present approach. In addition, one should make a distinction whether the swimming direction is confined within the 2D plane of motion, or whether rotation in all directions is still possible. In the latter case, the rotational operators in the Smoluchowski equation remain unchanged, but the swimming contribution involves projections of the orientations onto the 2D plane.

The dynamics, phase separation kinetics, phase behaviour, spatial inhomogeneities, and local orientational order parameters can be computed from the orientation-dependent local density $\rho(\mathbf{r}, \hat{\mathbf{u}}, t) \equiv N P_1(\mathbf{r}, \hat{\mathbf{u}}, t)$, where P_1 is the one-particle pdf for the position and orientation of an ABP, which is defined as the integral of the N -particle pdf, P_N , with respect to all coordinates except for those of a single particle:

$$\begin{aligned} P_1(\mathbf{r}, \hat{\mathbf{u}}, t) = & \int d\mathbf{r}_2 \cdots \int d\mathbf{r}_N \oint d\hat{\mathbf{u}}_2 \cdots \oint d\hat{\mathbf{u}}_N P_N(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N, \hat{\mathbf{u}}, \hat{\mathbf{u}}_2, \dots, \hat{\mathbf{u}}_N, t), \end{aligned} \quad (3)$$

where the integrals $\oint d\hat{\mathbf{u}}_j$ range over the surface of a unit sphere, that is, over all orientations of $\hat{\mathbf{u}}_j$. The orientation-dependent local density is the number concentration of ABPs at a position \mathbf{r} with an orientation $\hat{\mathbf{u}}$. The equation of motion for P_1 is obtained by the integration of the N -particle Smoluchowski equation with respect to all position coordinates and orientations, with the exception of a single position and orientation. This leads to the following equation:

$$\begin{aligned} \frac{\partial}{\partial t} \rho(\mathbf{r}, \hat{\mathbf{u}}, t) = & [D_0 \nabla^2 + D_r \hat{\mathcal{R}}^2] \rho(\mathbf{r}, \hat{\mathbf{u}}, t) - \nabla \cdot [\rho(\mathbf{r}, \hat{\mathbf{u}}, t) v_0 \hat{\mathbf{u}}] \\ & + \beta D_0 \nabla \cdot [\rho(\mathbf{r}, \hat{\mathbf{u}}, t)] \\ & \times \int d\mathbf{r}' \oint d\hat{\mathbf{u}}' \rho(\mathbf{r}', \hat{\mathbf{u}}', t) g(\mathbf{r}, \mathbf{r}', \hat{\mathbf{u}}, \hat{\mathbf{u}}', t) \nabla V(|\mathbf{r} - \mathbf{r}'|), \end{aligned} \quad (4)$$

where the interaction energy of the assembly of ABPs is assumed to be pair-wise additive, that is,

$$\Psi = \sum_{n > m}^N V(|\mathbf{r}_n - \mathbf{r}_m|), \quad (5)$$

where V is the pair-interaction potential. Furthermore, the pair-correlation function that appears in the integral in eqn (4) is defined as follows:

$$P_2(\mathbf{r}, \mathbf{r}', \hat{\mathbf{u}}, \hat{\mathbf{u}}', t) = P_1(\mathbf{r}, \hat{\mathbf{u}}, t) P_1(\mathbf{r}', \hat{\mathbf{u}}', t) g(\mathbf{r}, \mathbf{r}', \hat{\mathbf{u}}, \hat{\mathbf{u}}', t), \quad (6)$$

where the two-particle pdf, P_2 , is equal to the integral of P_N with respect to all particle coordinates except for two of them (similar to the definition of P_1). For non-interacting ABPs, P_2 factorizes into a product of two P_1 s, so that $g = 1$. The pair-correlation function thus describes the degree of correlation



between two particles embedded in a (concentrated) system of remaining ABPs.

For an isolated ABP, for which the last term in eqn (4) involving the pair-correlation function is absent, the following well-known expression for the mean-squared displacement W can be derived⁴⁶:

$$W(t) = 6 \left[D_0 + \frac{v_0^2}{6D_r} \right] t + \frac{v_0^2}{2D_r^2} [\exp\{-2D_r t\} - 1]. \quad (7)$$

This expression has been verified experimentally in 2D in ref. 47, in 2D for small times in ref. 48, and by simulations in ref. 35 and 49. For times much larger than $1/D_r$, it is found that $W(t) = 6D_s t$, where the long-time self-diffusion coefficient is equal to $D_s = D_0 + v_0^2/6D_r$. As will be seen in Section 6, the collective diffusion coefficient at infinite dilution is equal to the self-diffusion coefficient, just like for passive systems.

An analysis of MIPS based on the Smoluchowski equation (eqn (4)) requires an explicit expression for the pair-correlation function. The aim of the present analysis is to derive an explicit expression for the pair-correlation function, depending on the bare swimming velocity, v_0 . This will be done on the two-particle level, where binary interactions dominate, thus neglecting the indirect interactions between two particles *via* the surrounding particles. The extension of the results obtained in this work to higher concentrations requires a separate study, where, for example, a similar Ornstein–Zernike approach may be followed as for systems in equilibrium.⁹

3 A differential equation for the pair-correlation function

Before addressing the equation of motion for the pair-correlation function, note that the pair-correlation function in eqn (4) for the orientation-dependent local density is only needed for short distances between two ABPs, since the pair-correlation function in the integral in eqn (4) occurs as a product with the pair-interaction potential. We therefore need a closure for $g(\mathbf{r}, \mathbf{r}', \hat{\mathbf{u}}, \hat{\mathbf{u}}', t)$ for distances $|\mathbf{r} - \mathbf{r}'| < R_v$, where R_v is the range of the pair-interaction potential, which is of the order of the size of the ABPs. This has two important consequences. First, the interest here is in spatial variations of the orientation-dependent local density on length scales much larger than the size of single particles. Unstable modes involve the growth of density variations with a relatively large spatial extent, much larger than the range of the pair-interaction potential. We can therefore evaluate the pair-correlation function as needed in eqn (4) for a homogeneous system, at the local density at \mathbf{r} to within a neighbourhood of extent R_v from \mathbf{r}' . Second, since structural relaxation over small distances (of the order R_v) is much faster than the relaxation or growth of inhomogeneities varying over large distances (of the order of the spatial extent of unstable density variations), the pair-correlation function may be assumed to instantaneously adapt to the local density. That is, on the time scale on which density variations relax or grow, the pair-correlation function attains its stationary form.

These considerations are similar to those in ref. 50, where the Cahn–Hilliard theory for spinodal gas–liquid phase separation is derived from the Smoluchowski equation. Also, in that case, the pair-correlation function instantaneously relaxes to its equilibrium value corresponding to the local density, which may be regarded as the statistical analogon of local thermodynamic equilibrium approximations.

Note that phase separation kinetics from a meta-stable state (like nucleation and growth) cannot be described on the basis of the above-discussed local homogeneity and fast relaxation of the pair-correlation function. Such a phase separation mechanism involves the formation of a nucleus, which may have a spatial extent that is not much larger than the range R_v of the pair-interaction potential.

The equation of motion for the two-particle pdf, P_2 , can be obtained by integrating the N -particle Smoluchowski equation (eqn (1)) with respect to all particle coordinates except for two, similar to the derivation of eqn (4) for the one-particle pdf. This leads to an equation of motion for the pair-correlation function from its definition in eqn (6). On the two-particle level, neglecting higher-order correlations related to the three-particle pdf leads to the following stationary equation of motion:

$$0 = 2D_0 \nabla \cdot \left\{ \nabla + \beta \nabla V(R) - \frac{v_0}{2D_0} (\hat{\mathbf{u}}_1 - \hat{\mathbf{u}}_2) \right\} g(\mathbf{R}, \hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2) + D_r \{ \hat{\mathcal{R}}_1^2 + \hat{\mathcal{R}}_2^2 \} g(\mathbf{R}, \hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2), \quad (8)$$

where ∇ , $\hat{\mathcal{R}}_1$, and $\hat{\mathcal{R}}_2$ are understood to act on the distance $\mathbf{R} = \mathbf{r}_1 - \mathbf{r}_2$ between two ABPs, and the orientations $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ of the two ABPs, respectively. Note that the pair-correlation function is a function of the relative distance $\mathbf{r}_1 - \mathbf{r}_2$ due to the approximate homogeneity on length scales larger than R_v , and is independent of time due to the fast relaxation of correlations over short distances, as discussed above.

We note that with the so-called linearized Dean equation applied to a dilute system in ref. 41 to analyze the experimental data for the pair-correlation function in 2D, the contribution $2D_0\beta\nabla \cdot [g\nabla V] = (2/\gamma)\nabla \cdot [g\nabla V]$ in eqn (8) (where γ is the Stokes–Einstein friction coefficient) is replaced by $(2/\gamma)\nabla^2 V$, thus neglecting pair-correlations to this contribution. Such a neglect is not allowed within the Smoluchowski-equation approach, neither for short-ranged nor for long-ranged pair-interaction potentials. In particular, including hard-core-excluded volume interactions, the product $g\nabla V$ is proportional to the delta distribution at the contact of the two ABPs (see eqn (27) in Section 6), while ∇V by itself is undefined.

Without activity, it follows from eqn (8) that the pair-correlation function is given by the Boltzmann exponential $g(R) = \exp\{-\beta V(R)\}$. This is the well-known leading order term in an expansion of the pair-correlation function with respect to the overall concentration for passive systems. For such low concentrations, binary inter-particle interactions are dominant. Terms that account for higher-order collisions involving more than two particles have been omitted in eqn (8). For excluded-volume interactions, it is expected that this limits the validity of our analysis to volume fractions less than about 0.10. It should



be noted that activity cannot be included in eqn (8) in terms of the gradient of an interaction potential. If this would have been possible, the pair-correlation function would simply be given by the afore-mentioned Boltzmann exponential of that potential.

Specializing to hard-core interactions (in which case, $V(R) = 0$ for $R > 2a$, where a is the radius of a particle, and $V(R) = \infty$ for $R < 2a$), with a possibly super-imposed very short-ranged attractive “sticky-sphere” interaction, eqn (8) reduces for $R > 2a$ to

$$0 = 2D_0 \nabla \cdot \left\{ \nabla - \frac{v_0}{2D_0} (\hat{\mathbf{u}}_1 - \hat{\mathbf{u}}_2) \right\} g(\mathbf{R}, \hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2) + D_r \{ \hat{\mathcal{R}}_1^2 + \hat{\mathcal{R}}_2^2 \} g(\mathbf{R}, \hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2), \quad \text{for } R > 2a. \quad (9)$$

In solving this equation, the hard-core interactions are accounted for by the no-flux boundary condition:

$$0 = \hat{\mathbf{R}} \cdot \left\{ \nabla - \frac{v_0}{2D_0} (\hat{\mathbf{u}}_1 - \hat{\mathbf{u}}_2) \right\} g(\mathbf{R}, \hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2), \quad \text{for } R = 2a, \quad (10)$$

where $\hat{\mathbf{R}} = \mathbf{R}/R$ is the unit vector normal to the spherical surface with radius $2a$. This boundary condition physically means that the cores of two particles cannot interpenetrate.

The differential equation (eqn 9) and the boundary condition (eqn 10) are most conveniently rewritten in terms of dimensionless variables, introducing the dimensionless distance and gradient operator:

$$\begin{aligned} \mathbf{R} &\rightarrow \mathbf{R}/2a, \\ \nabla &\rightarrow 2a \nabla. \end{aligned} \quad (11)$$

Hereafter, we use the same symbols for the dimensionless distance and gradient operator as before for the distance and gradient operator with dimensions. Note that $R = 1$ at the contact of the two particles. The free diffusion coefficients are equal to $D_0 = k_B T / 6\pi\eta_0 a$ and $D_r = k_B T / 8\pi\eta_0 a^3$, where η_0 is the shear viscosity of the solvent, so that the dimensionless Peclet number,

$$\lambda = \frac{av_0}{D_0}, \quad (12)$$

can be introduced (some times also denoted by Pe) to arrive at

$$\begin{aligned} 0 &= \nabla \cdot \{ \nabla - \lambda \mathbf{U} \} g + \frac{3}{2} \{ \hat{\mathcal{R}}_1^2 + \hat{\mathcal{R}}_2^2 \} g, \quad \text{for } R > 1, \\ 0 &= \hat{\mathbf{R}} \cdot \{ \nabla - \lambda \mathbf{U} \} g, \quad \text{for } R = 1, \\ g &\rightarrow 1, \quad \text{for } R \rightarrow \infty, \end{aligned} \quad (13)$$

where

$$\mathbf{U} = \hat{\mathbf{u}}_1 - \hat{\mathbf{u}}_2. \quad (14)$$

The dimensionless Peclet number, λ , characterizes the relative importance of the swimming velocity of an isolated ABP as compared to translational and rotational diffusion. Note that it is not necessary to make a distinction between a translational and rotational Peclet number: for the spherical particles under consideration, the ratio D_0/D_r for the translational and rotational diffusion coefficients is equal to $4a^2/3$, which

results in identical Peclet numbers (apart from a factor $3/2$) in the dimensionless form of the stationary equation of motion (eqn 13) for the translational and rotational contributions. The last boundary condition in eqn (13) at infinity results from the fact that particles become uncorrelated at large separations.

We have not been able to solve this set of equations analytically for arbitrary values of the Peclet number. In the sections below, we shall therefore determine the asymptotic solutions for small and large swimming velocities.

As will turn out, these asymptotic expressions are approximately cylindrically symmetric around the direction of \mathbf{U} , so that the pair-correlation function for these limiting Peclet numbers can be expressed in terms of the cylindrical coordinates $Z = R \cos \Theta$ and $\rho = R \sin \Theta$, where Θ is the angle between \mathbf{R} and \mathbf{U} (as depicted in Fig. 1), that is,

$$\cos \Theta = \hat{\mathbf{R}} \cdot \hat{\mathbf{U}}, \quad (15)$$

where $\hat{\mathbf{R}}$ and $\hat{\mathbf{U}}$ are the unit vectors in the directions of \mathbf{R} and \mathbf{U} , respectively. In addition, the asymptotic expressions for the pair-correlation function depend on the magnitude of \mathbf{U} , which is equal to

$$U = |\hat{\mathbf{u}}_1 - \hat{\mathbf{u}}_2| = \sqrt{2 - 2\alpha}, \quad (16)$$

where

$$\alpha = \hat{\mathbf{u}}_1 \cdot \hat{\mathbf{u}}_2 \quad (17)$$

is the cosine of the angle between $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ (also indicated within sphere 2 in Fig. 1). For intermediate Peclet numbers, there is an additional dependence on $\hat{\mathbf{R}} \cdot (\hat{\mathbf{u}}_1 + \hat{\mathbf{u}}_2)$. Due to the approximate independence of the asymptotic solutions for small and large Peclet numbers on this variable, the pair-correlation function might be a weak function of this variable for intermediate Peclet numbers as well.

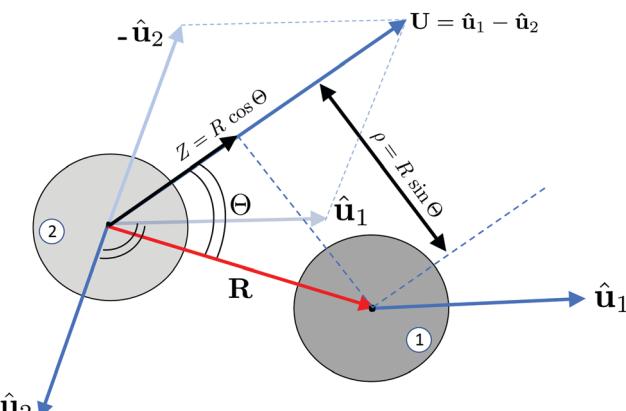


Fig. 1 Definition of the variables R , Z , ρ and U in relation to the previously defined orientations, $\hat{\mathbf{u}}_{1,2}$, and the center-to-center distance, \mathbf{R} . For simplicity, all vectors are drawn within the same 2D plane. The asymptotic expressions derived below are the functions of these four scalar variables.



4 The pair-correlation function for small Peclet numbers

For small Peclet numbers, the pair-correlation can be expanded as follows:

$$g = 1 + \lambda g_1 + \lambda^2 g_2 + \dots \quad (18)$$

After substitution into eqn (13), it is found that the leading order contribution satisfies the following set of equations:

$$\begin{aligned} 0 &= \nabla^2 g_1 + \frac{3}{2} \{ \hat{\mathcal{R}}_1^2 + \hat{\mathcal{R}}_2^2 \} g_1, \quad \text{for } R > 1, \\ 0 &= \hat{\mathbf{R}} \cdot \{ \nabla g_1 - \mathbf{U} \}, \quad \text{for } R = 1, \\ g_1 &\rightarrow 0, \quad \text{for } R \rightarrow \infty. \end{aligned} \quad (19)$$

The solution of this set of equations is determined in Appendix A by means of a spherical harmonics series representation, leading to

$$\begin{aligned} g(R, U, \Theta) &= 1 - \frac{1}{5 + 2\sqrt{3}} \left[\frac{\sqrt{3}}{R} + \frac{1}{R^2} \right] \\ &\quad \times \exp \left\{ -\sqrt{3}(R - 1) \right\} \lambda U \cos(\Theta), \end{aligned} \quad (20)$$

where the various coordinates are defined at the end of the previous section (see also Fig. 1). The contact value g^c of the pair-correlation function (for which $R = 1$) is thus equal to

$$g^c(U, \Theta) = 1 - \frac{1 + \sqrt{3}}{5 + 2\sqrt{3}} \lambda U \cos(\Theta). \quad (21)$$

A little consideration (see also Fig. 1) shows that the particles move toward each other when $\cos \Theta < 0$, and move away from each other when $\cos \Theta > 0$. Eqn (20) thus shows that the pair-correlation function is larger than unity when the particles move toward each other, and is equally less than unity in case they move away from each other.

Note that the dimensionless relative velocity of two particles is equal to λU . In case $U = 0$, for which $\hat{\mathbf{u}}_1 = \hat{\mathbf{u}}_2$, the two particles move side-to-side along with each other with the same velocity. The motility-induced distortion of the pair-correlation is zero in that case, as the particles do not interact with each other. Hence, no matter how large the Peclet number may be, for $U = 0$, there is no effect of motility on the pair-correlation function. This explains why the leading order correction for small velocities in eqn (20) is proportional to λU instead of just λ .

Numerical results for $g - 1$ are given in Fig. 2 for $\lambda U = 1$. The spherical white area with radius unity is the excluded volume by particle 1 for particle 2, within the coordinate frame that moves along with particle 1, the center of which is chosen at the origin. As discussed above, the two particles approach each other when $Z < 0$, leading to an increase of the pair-correlation function, and move apart when $Z > 0$, leading to a decrease.

The angular average of the contact value of the pair-correlation function is zero. Beyond the linear approximation discussed here, however, this angular average will be non-zero.

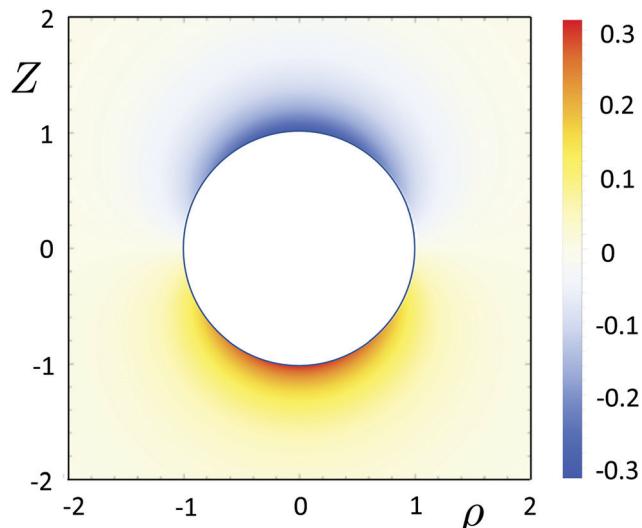


Fig. 2 Deviation of the pair-correlation function from unity for $\lambda U = 1$ according to eqn (20) for small Peclet numbers. The coordinates $Z = R \cos \Theta$ and $\rho = R \sin \Theta$ are the cylindrical coordinates around the direction of \mathbf{U} , which are defined in Fig. 1.

In particular, for the large Peclet numbers considered in the subsequent section, the angular averaged contact value varies like λ , and thus attains very large values.

5 The pair-correlation function for large Peclet numbers

For large velocities, the behavior of the pair-correlation is quite different in case the particles approach each other or move apart. We therefore define the so-called front-sector as the space where the two particles approach each other (for which $\cos \Theta < 0$), and the wake-sector as the space where the particles move apart (for which $\cos \Theta > 0$). Contrary to the behavior of the pair-correlation function for small velocities, where the increase in the front-sector is equal to the decrease in the wake sector, for large velocities, there is a much more pronounced asymmetry. When the two particles approach each other, there is a probability accumulation within a narrow region where the pair-correlation function attains large values, whereas when the two particles move apart, there is a probability shadow within an extended region where the pair-correlation function is essentially zero. The analysis in both sectors will be discussed separately in the following subsections.

5.1 The front-sector: $\cos \Theta < 0$

Multiplying eqn (13) by the small parameter $1/\lambda$, it is obvious that this small parameter multiplies the highest order spatial Laplacian derivative. A perturbation expansion approach where the second-order Laplacian derivative would be simply neglected (as it is multiplied by the small parameter $1/\lambda$) leads to a first-order differential equation. There are, however, two boundary conditions, which cannot be satisfied simultaneously by adjusting the single integration constant of the solution of



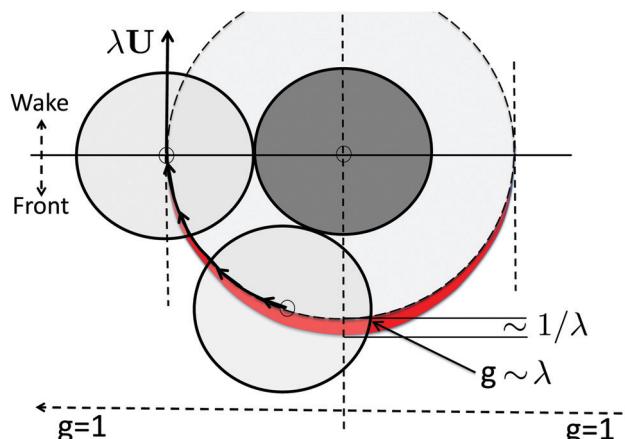


Fig. 3 Schematic representation of the boundary-layer behaviour of the pair-correlation function in the front-sector, where two particles approach each other: the light-grey sphere overtakes the dark-grey sphere. As soon as the two spheres are side-to-side (where $\cos \Theta = 0$), the light-grey sphere enters the wake-sector. Within the boundary layer (indicated in red), with a spatial extent $\sim 1/\lambda$, the pair-correlation function attains large values, $\sim \lambda$. The pair-correlation function is unity outside the boundary layer (but within the front-sector).

that first-order differential equation. This implies that the second-order derivative must be very large near the contact, so that the Lapacian derivative cannot be neglected, even though it is multiplied by the small parameter $1/\lambda$. This is the standard situation for singularly perturbed differential equations^{51,52} and leads, in the present case, to a narrow boundary layer near contact of the two particles where the pair-correlation function exhibits large spatial gradients, as schematically depicted in Fig. 3 in red. In this figure, the dark-grey sphere is overtaken by the light-grey sphere. Outside the boundary layer, the pair-correlation function is unity (which is the value that the pair-correlation function attains without interactions).

In a singular perturbation approach, the boundary layer is spatially stretched by introducing the so-called boundary-layer variable, $S = \lambda^\nu(R - 1)$, where the exponent ν is determined in such a way that the second-order derivative with respect to S is of the same order as the remaining dominant terms in the differential equation, commonly referred to as “dominant balance”, which in the present case is the contribution from the motility-induced flux. The boundary-layer analysis of eqn (13) is given in Appendix B, leading to the following expression for the pair-correlation function, valid up to the leading order of the expansion:

$$g_{\text{front}}(R, U, \Theta) = 1 - \frac{1}{2}\lambda U \cos(\Theta) \exp\{(R - 1)[\lambda U \cos(\Theta) - 2]\}, \quad (22)$$

where the index “front” indicates that this expression is only valid in the front-sector, where $\cos \Theta < 0$. As before, Θ is the angle between \mathbf{R} and $\mathbf{U} = \hat{\mathbf{u}}_1 - \hat{\mathbf{u}}_2$. Note that the contact value of the pair-correlation function, where $R = 1$, is equal to

$$g_{\text{front}}^c(U, \Theta) = 1 - \frac{1}{2}\lambda U \cos \Theta. \quad (23)$$

For $\cos \Theta = 0$, eqn (22) predicts that $g = 1$, as it should be, since the particles do not interact in that configuration (the two particles are side-by-side, like the left upper light-grey sphere and the dark-grey sphere in Fig. 3). Furthermore, the distance $R - 1$ from the contact over which the pair-correlation function relaxes to unity is $\sim 1/\lambda U$. This quantifies the behavior discussed above, which is schematically depicted in Fig. 3.

As shown in Appendix B, the rotational contributions to eqn (13) can be neglected. The physical interpretation of this result is that, due to the small extent of the boundary layer, the time during which the particles interact (where the light-grey sphere rolls over the dark-grey sphere, as depicted in Fig. 3) is sufficiently short that the rotational diffusion is not effective. Note that $g = 1$ in case $U = 0$, as it should be, since in that case the two spheres move along with each other with the same velocity.

Numerical results are given in Fig. 4 for $\lambda U = 10$ in the Z -versus- ρ plane as before, which shows the boundary-layer behavior of the pair-correlation function within the front-sector. The boundary-layer type of behaviour of the pair-correlation function in 3D at low concentrations is confirmed by simulations in ref. 53, where orientational averages show a very steep increase near contact. The probability accumulation in the front-sector in 2D has also been observed in simulations in ref. 28. Such a steep increase of the pair-correlation function and large contact value correspond to a sticky-sphere type of interaction for passive systems. This has been used in ref. 54 to compare the equation of state for active particles to the equation of state of a system of passive adhesive particles.

5.2 The wake-sector: $\cos \Theta > 0$

The result in eqn (22) diverges for large distances when $\lambda U \cos \Theta > 2$, so that the above result does not describe the behavior of the pair-correlation function in the wake-sector where $\cos \Theta > 0$. In the wake-sector, the behavior is very different from that in the front-sector. Once the particles enter the wake-sector (the light-grey spheres depicted in Fig. 5), there is a probability-depleted region indicated in blue in Fig. 5, which we will refer to as the probability shadow.

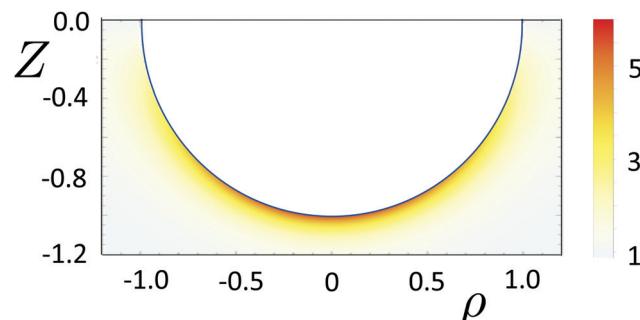


Fig. 4 The pair-correlation function in the front-sector for $\lambda U = 10$ according to eqn (22). As before, $Z = R \cos \Theta$ and $\rho = R \sin \Theta$ are the cylindrical coordinates relative to the direction of \mathbf{U} (see Fig. 1). The pair-correlation function takes large values, $\sim \lambda$, within a boundary layer of extent $\sim 1/\lambda$.



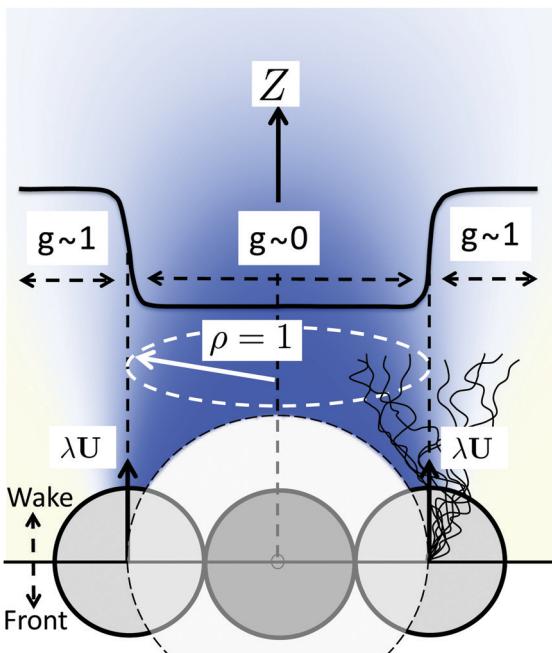


Fig. 5 Schematic representation of the probability shadow in the wake-sector, where the particles move away from each other. Within the blue region where $\rho < 1$, the pair-correlation function is essentially zero for inter-particle distances less than $\lambda^{1/3}$, while for $\rho > 1$, it is unity. For large distances in the direction of \mathbf{U} , and $\rho < 1$, the pair-correlation function relaxes to unity due to (rotational) diffusion into the probability shadow, as indicated by the wiggly lines for the light-grey particle on the right which overtakes the dark-grey sphere in the middle.

For infinite Peclet numbers, there is a cylindrical region of infinite extent in the direction of \mathbf{U} where the pair-correlation function is zero.⁵⁵ For finite Peclet numbers, there is an approximate cylindrically symmetric region around the direction of \mathbf{U} , where the pair-correlation function is essentially zero for short distances, as time has been too short for the overtaking particle to move into the probability shadow. Therefore, the pair-correlation function attains an almost step-function-like behavior for short distances as a function of the cylindrical coordinate $\rho = R \sin \Theta$ at $\rho = 1$, as indicated by the solid line in Fig. 5 (see Fig. 1 for the definition of coordinates). For $\rho < 1$, the pair-correlation function is essentially zero, while for $\rho > 1$, its value is approximately unity. The pair-correlation function relaxes to unity at large distances in the direction of \mathbf{U} due to translational and rotational diffusion into the probability shadow (as indicated by the thin solid lines for the light-grey sphere on the right in Fig. 5). In contrast to the large spatial derivatives in the direction perpendicular to \mathbf{U} at $\rho = 1$, spatial derivatives in the direction of \mathbf{U} are very small. Furthermore, the spatial derivatives along the direction of \mathbf{U} (and for $\rho < 1$) decrease with the increasing swimming velocity, while the extent of the probability shadow increases with the increasing velocity. This behavior is to be contrasted to the front-sector, where the opposite occurs: here, the spatial derivatives in the direction of \mathbf{U} are very large, and increase with the increasing swimming velocity, while the extent of the boundary layer decreases with the increasing velocity.

An approximate asymptotic solution of eqn (13) for large Peclet numbers within the wake-sector is constructed in Appendix B, the result of which is

$$g_{\text{wake}}(R, U, \Theta) = 1 - \int_0^\infty dC J_1(C) J_0\left(CR\sqrt{1 - \cos^2(\Theta)}\right) \times \exp\left\{-C^2 \frac{4 + U^2}{8U^2} \frac{R^3}{\lambda U} \cos^3(\Theta)\right\}. \quad (24)$$

where $J_{0,1}$ are the Bessel functions of the first kind (of order 0 and 1). The index “wake” refers to its validity within the wake-sector. As shown in Appendix B by numerical integration, the contact value of the above expression for the pair-correlation function decays to zero faster than $1/\lambda$. Therefore, the contact value is zero,

$$g_{\text{wake}}^c(U, \Theta) = 0, \quad (25)$$

to within leading order in $1/\lambda$, in contrast to the front-sector side where the contact value is very large. A constant contact value for the pair-correlation function, uniformly spread over the surface of the sphere, is found in ref. 40 on the basis of a theoretical approach that does not explicitly account for orientational diffusion, as discussed in the introduction. According to the above Smoluchowski-type of approach, orientational diffusion is, however, essential for small Peclet numbers as well as for high Peclet numbers in the wake-sector.

The extent of the probability shadow can be calculated from eqn (24) as follows. In the Z -direction, we have $\cos \Theta = 1$, so that the Bessel function J_0 in eqn (24) is equal to unity. The remaining integral can be calculated analytically as follows:⁵⁶

$$g_{\text{wake}}(R, U, \Theta = 0) = 1 - \int_0^\infty dC J_1(C) \exp \left\{ -C^2 \frac{4 + U^2}{8U^2} \frac{R^3}{\lambda U} \right\} \\ = \exp \left\{ -\frac{2U^2}{4 + U^2} \frac{\lambda U}{R^3} \right\}. \quad (26)$$

Note that the contact value for $\Theta = 0$ is exponentially small with the increasing λ . Furthermore, this result shows that the extent l of the probability shadow scales like $l \sim \lambda^{1/3}$ for non-zero relative velocities, U . As shown in Appendix B, this scaling is entirely due to rotational diffusion. Translational diffusion can be neglected in the wake-sector, contrary to the front-sector, where it is dominant. The scaling relation for the extent of the probability shadow can be understood as follows.⁵⁷ Let $\delta\theta$ denote a small angle over which the spheres depicted in Fig. 5 rotate relative to each other in a direction where the overtaking light-grey sphere moves into the blue, probability-depleted region. The time δt it takes for such a small rotation is $\delta t \sim (\delta\theta)^2/D_r$. Let l be the extent of the probability shadow in the Z -direction. Since the width of the probability shadow is $\sim a$, we have $\delta\theta \sim a/l$. Furthermore, $l \sim v_0\delta t$, so that eliminating $\delta\theta$ and δt in favour of l leads to $(l/a)^3 \sim v_0/(aD_r) \sim av_0/D_0 = \lambda$. This confirms the scaling found in eqn (24) and (26), and shows that the scaling is solely due to rotational diffusion.

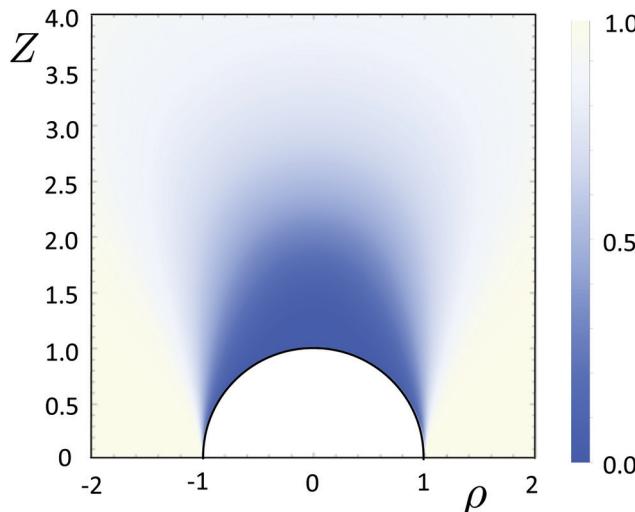


Fig. 6 The pair-correlation function in the wake-sector for $\lambda = 5$ and $U = 2$ according to eqn (24). As before, $Z = R \cos \Theta$ and $\rho = R \sin \Theta$ are the cylindrical coordinates relative to the direction of U (see Fig. 1). The pair-correlation function is essentially a step-function as a function of ρ at $\rho = 1$ when the spheres are not too far apart, while it converges to unity in the direction along U within the probability shadow over large distances of order $\sim \lambda^{1/3}$.

Numerical results are given in Fig. 6 for $\lambda = 5$ and $U = 2$. This density plot quantifies the expected step-like function behavior for $\rho = 1$ in case the spheres are not too far apart, and the very slow convergence to unity in the direction along U within the probability shadow.

In ref. 41, it is found that the depletion wake in 2D consists of two wings, where the two wings are surrounded by regions where the pair-correlation function is enhanced due to motility (see, for example, Fig. 1 in ref. 41). There, the pair-correlation function is plotted in a coordination frame that is fixed to the orientation, $\hat{\mathbf{u}}_1$, of a single colloid and averaged with respect to the orientation, $\hat{\mathbf{u}}_2$, of the second colloid. We have plotted the pair-correlation function in a coordination frame that is fixed to the relative velocity, $\sim (\hat{\mathbf{u}}_1 - \hat{\mathbf{u}}_2)$. The latter allows us to make a distinction between the approaching swimmers and the swimmers moving away from each other, and thus distinguishes between what we referred to as the wake-sector and the front-sector. Averaging over the orientation, $\hat{\mathbf{u}}_2$, mixes parts of the wake-sector and the front-sector, which probably explains the two-wing structure reported in ref. 41.

The 2D simulations in ref. 28 do not show a clear long-ranged probability shadow. These simulations have been performed at very high packing fractions, and thus indicate that higher-order inter-particle correlations affect the spatial extent of the probability shadow in the wake-sector.

6 Collective diffusion

With the expressions for the pair-correlation function derived in the previous sections, the equation of motion (eqn 4) can be evaluated in closed form. Since the pair-correlation function is evaluated to the leading order in concentration, that equation

of motion is also valid up to leading order in concentration. In this section, we derive expressions for the Fickian collective diffusion coefficient and the motility-induced polarization due to concentration gradients for hard-sphere interactions.

Since the pair-correlation function is zero when the spheres overlap, we have:

$$g(\mathbf{r}, \mathbf{r}', \hat{\mathbf{u}}, \hat{\mathbf{u}}') \nabla V(|\mathbf{r} - \mathbf{r}'|) = -\beta^{-1} g^c(\mathbf{r}, \mathbf{r}', \hat{\mathbf{u}}, \hat{\mathbf{u}}') \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|} \delta(|\mathbf{r} - \mathbf{r}'| - 2a), \quad (27)$$

where, as before, g^c is the contact value of the pair-correlation function, a is the radius of the particles, and $\delta(\cdot)$ is the delta distribution. Substitution into eqn (4) gives

$$\begin{aligned} \frac{\partial}{\partial t} \rho(\mathbf{r}, \hat{\mathbf{u}}, t) = & [D_0 \nabla^2 + D_r \hat{\mathcal{R}}^2] \rho(\mathbf{r}, \hat{\mathbf{u}}, t) - \nabla \cdot [\rho(\mathbf{r}, \hat{\mathbf{u}}, t) v_0 \hat{\mathbf{u}}] \\ & - \frac{D_0}{2a} \nabla \cdot \left[\rho(\mathbf{r}, \hat{\mathbf{u}}, t) \oint d\hat{\mathbf{u}}' \oint dS g^c(\mathbf{R}, \hat{\mathbf{u}}, \hat{\mathbf{u}}') \mathbf{R} \rho(\mathbf{r} - \mathbf{R}, \hat{\mathbf{u}}', t) \right]. \end{aligned} \quad (28)$$

Here, $\mathbf{R} = \mathbf{r} - \mathbf{r}'$ is the non-scaled distance between the two ABPs, and $\oint dS$ is the integral with respect to \mathbf{R} ranging over a spherical surface with radius $2a$. For small spatial gradients in the orientation-dependent local density, such that the density varies linearly over distances of the order of the size of the particles, the leading order gradient expansion,

$$\rho(\mathbf{r} - \mathbf{R}, \hat{\mathbf{u}}', t) = \rho(\mathbf{r}, \hat{\mathbf{u}}', t) - \mathbf{R} \cdot \nabla \rho(\mathbf{r}, \hat{\mathbf{u}}', t), \quad (29)$$

can be used in eqn (28) to obtain

$$\begin{aligned} \frac{\partial}{\partial t} \rho(\mathbf{r}, \hat{\mathbf{u}}, t) = & [D_0 \nabla^2 + D_r \hat{\mathcal{R}}^2] \rho(\mathbf{r}, \hat{\mathbf{u}}, t) - \nabla \cdot [\rho(\mathbf{r}, \hat{\mathbf{u}}, t) v_0 \hat{\mathbf{u}}] \\ & + \frac{D_0}{2a} \nabla \cdot \left[\rho(\mathbf{r}, \hat{\mathbf{u}}, t) \oint d\hat{\mathbf{u}}' \mathbf{I}_1(\hat{\mathbf{u}}, \hat{\mathbf{u}}') \rho(\mathbf{r}, \hat{\mathbf{u}}', t) \right] \\ & + \frac{D_0}{2a} \nabla \cdot \left[\rho(\mathbf{r}, \hat{\mathbf{u}}, t) \nabla \cdot \oint d\hat{\mathbf{u}}' \mathbf{I}_2(\hat{\mathbf{u}}, \hat{\mathbf{u}}') \rho(\mathbf{r}, \hat{\mathbf{u}}', t) \right], \end{aligned} \quad (30)$$

where

$$\begin{aligned} \mathbf{I}_1(\hat{\mathbf{u}}, \hat{\mathbf{u}}') = & - \oint dS g^c(\mathbf{R}, \hat{\mathbf{u}}, \hat{\mathbf{u}}') \mathbf{R}, \\ \mathbf{I}_2(\hat{\mathbf{u}}, \hat{\mathbf{u}}') = & \oint dS g^c(\mathbf{R}, \hat{\mathbf{u}}, \hat{\mathbf{u}}') \mathbf{R} \mathbf{R}. \end{aligned} \quad (31)$$

Since for large swimming velocities the contact value of the pair-correlation function is zero on the wake-sector side, the integrals in eqn (31) range over the hemispherical surface corresponding to the front-sector side, as depicted in Fig. 7. The large contact value on the front-sector side is given in eqn (23), so that the integrations in eqn (31) lead to

$$\begin{aligned} \mathbf{I}_1(\hat{\mathbf{u}}, \hat{\mathbf{u}}') = & 8\pi a^3 \left[1 + \frac{1}{3} \lambda U \right] \frac{\mathbf{U}}{U}, \\ \mathbf{I}_2(\hat{\mathbf{u}}, \hat{\mathbf{u}}') = & \left[\frac{16}{3} + \lambda U \right] 2\pi a^4 \hat{\mathbf{U}} + 2\pi a^4 \lambda U \frac{\mathbf{U} \mathbf{U}}{U^2}, \end{aligned} \quad (32)$$

where, as before, $\mathbf{U} = \hat{\mathbf{u}} - \hat{\mathbf{u}}'$ and $U = |\hat{\mathbf{u}} - \hat{\mathbf{u}}'|$.



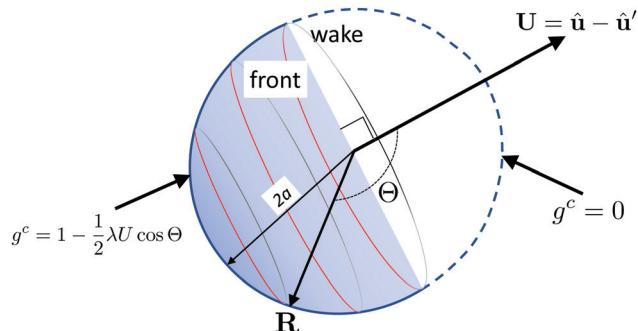


Fig. 7 For large Peclet numbers, the integrations in eqn (31) range over the semi-spherical surface with radius $2a$ indicated in blue, corresponding to the front-sector side where $\cos \theta < 0$.

The local fraction of the volume that is occupied by the cores of the particles with an orientation $\hat{\mathbf{u}}$, is by definition equal to

$$\phi(\mathbf{r}, \hat{\mathbf{u}}, t) = \frac{4\pi}{3} a^3 \rho(\mathbf{r}, \hat{\mathbf{u}}, t). \quad (33)$$

Fick's diffusion equation is the equation of motion for the volume fraction irrespective of their orientation:

$$\phi(\mathbf{r}, t) = \oint d\hat{\mathbf{u}} \phi(\mathbf{r}, \hat{\mathbf{u}}, t), \quad (34)$$

where the so-called polarization,

$$\mathbf{P}(\mathbf{r}, t) = \oint d\hat{\mathbf{u}} \hat{\mathbf{u}} \phi(\mathbf{r}, \hat{\mathbf{u}}, t), \quad (35)$$

plays an important role. The polarization is proportional to the particle volume flux in the direction of the average polar alignment. This is the most important contribution to the motility-induced mass transport. We will therefore use an expansion of $\phi(\mathbf{r}, \hat{\mathbf{u}}, t)$ (similar to what is discussed in, for example, ref. 21, 28, 39, 58 and 59) with respect to the orthonormal polyadic products of $\hat{\mathbf{u}}$ up to first order (see Appendix C for the discussion of such expansions):

$$\phi(\mathbf{r}, \hat{\mathbf{u}}, t) = \frac{1}{4\pi} \phi(\mathbf{r}, t) + \frac{3}{4\pi} \mathbf{P}(\mathbf{r}, t) \cdot \hat{\mathbf{u}}. \quad (36)$$

The next higher-order term in the expansion is proportional to

$$\oint d\hat{\mathbf{u}} \left[\hat{\mathbf{u}} \hat{\mathbf{u}} - \frac{1}{3} \hat{\mathbf{I}} \right] \phi(\mathbf{r}, \hat{\mathbf{u}}, t),$$

which measures the degree of motility-induced nematic order. Since the mass transport for large Peclet numbers is severely affected by the motility-induced polar alignment and much less by nematic ordering, this higher-order contribution to eqn (36) will not be considered.

For large swimming velocities, it is found from eqn (30) and (32), with some effort, that (see Appendix C for mathematical details)

$$\frac{\partial}{\partial t} \phi(\mathbf{r}, t) = \frac{4}{3} \lambda D_0 \nabla \cdot (\phi \nabla \phi) - v_0 \nabla \cdot \mathbf{P}(\mathbf{r}, t), \quad (37)$$

where only the dominant contributions for large Peclet numbers and leading order terms in spatial gradients are kept.

As will be seen below, \mathbf{P} is of the order $\nabla \phi$, so that eqn (37) is valid up to second order in spatial gradients, just like the original Fick's equation for passive systems. Keeping only the leading terms in the swimming velocities for the large Peclet numbers under consideration, the equation of motion for the polarization, \mathbf{P} , is similarly found, noting that $\lambda D_0/a = v_0$ (again, mathematical details are given in Appendix C):

$$\frac{\partial}{\partial t} \mathbf{P}(\mathbf{r}, t) = -2D_r \mathbf{P} - \frac{1}{3} v_0 \{1 - 2\phi(\mathbf{r}, t)\} \nabla \phi(\mathbf{r}, t), \quad (38)$$

where, as before, higher-order spatial gradients have been neglected. The solution of this equation is

$$\mathbf{P}(\mathbf{r}, t) = \int_{-\infty}^t dt' \mathbf{H}(\mathbf{r}, t') \exp\{2D_r(t' - t)\}, \quad (39)$$

$$\mathbf{H}(\mathbf{r}, t) = -(1/3)v_0\{1 - 2\phi(\mathbf{r}, t)\} \nabla \phi(\mathbf{r}, t).$$

In case spatial gradients are sufficiently small, so that these evolve on time scales much larger than $1/D_r$, this expression reduces for $t \gg 1/D_r$ to

$$\mathbf{P}(\mathbf{r}, t) = \frac{\mathbf{H}(\mathbf{r}, t)}{2D_r}. \quad (40)$$

This principle is commonly referred to as enslavement, where the non-conserved polarization, $P(\mathbf{r}, t)$, is enslaved by the conserved concentration, $\phi(\mathbf{r}, t)$.^{28,58,59} Physically, this amounts to a coarse graining to time scales where orientation is fully relaxed, while spatial gradients are sufficiently small to evolve on much larger time scales. It thus follows that,

$$\mathbf{P}(\mathbf{r}, t) = -\frac{2}{9} \lambda a \{1 - 2\phi(\mathbf{r}, t)\} \nabla \phi(\mathbf{r}, t). \quad (41)$$

Terms like $\sim \nabla^2 \mathbf{P}^2$ and $\sim \nabla \nabla \cdot \mathbf{P} \mathbf{P}$ that contribute to the right-hand side of eqn (37) (see Appendix C) are thus of fourth order in spatial gradients. Similarly, terms like $\sim \nabla^2 \mathbf{P}$, $\mathbf{P} \nabla \cdot \mathbf{P}$, $\nabla \cdot \mathbf{P} \mathbf{P}$, $[\nabla \phi] [\nabla \mathbf{P}]$, etc. are neglected in eqn (38), since these terms also contribute to higher-order gradients.

The condition for the steepness of the concentration gradients for which enslavement holds will be addressed below. Eqn (41) implies that there is a preferred orientation in the opposite direction to the concentration gradient. Thus, there is an enhanced motility-induced particle flux from high to low concentrations.

Substituting eqn (41) into the diffusion equation (eqn 37) finally leads to Fick's diffusion equation:

$$\frac{\partial}{\partial t} \phi(\mathbf{r}, t) = \nabla \cdot [D_c^{\text{eff}}(\mathbf{r}, t|\lambda) \nabla \phi(\mathbf{r}, t)], \quad (42)$$

where the effective collective diffusion coefficient is equal to

$$D_c^{\text{eff}}(\mathbf{r}, t|\lambda) = \frac{2}{9} \lambda^2 D_0 \{1 - 2\phi(\mathbf{r}, t)\}, \text{ for } \lambda \gg 1, \quad (43)$$

where terms proportional to orders of λ less than two have been neglected. Mass transport at high Peclet numbers is thus dominated by motility due to the polar alignment along the concentration gradients.

The enslavement argument used above holds when the relaxation time of the density is much larger than the rotational

relaxation time, $1/D_r$. From the above result for the effective diffusion coefficient, it thus follows that enslavement requires that

$$\Lambda^{-2} D_0 \lambda^2 \ll D_r \rightarrow \Lambda \gg \lambda a, \quad (44)$$

where Λ specifies the extent over which the concentration varies (which is the typical wavelength corresponding to the Fourier-mode decomposition of the volume fraction). This inequality is a condition for the validity of the above Fick's diffusion eqn (42) and (43) for large λ . Without motility, the wavelength Λ should be much larger than the radius a of the spheres. For Fick's equation to hold for active systems, the wavelength should be much larger than λa . The condition (eqn 44) physically means that the concentration should not change during the time needed for orientational relaxation, where an ABP migrates over a distance λa .

Including all terms of second order in the spatial gradients originating from the asymptotic solution of the pair-correlation function in the equations of motion for density and polarization leads to (see Appendix C for the list of all these terms)

$$D_c^{\text{eff}}(\mathbf{r}, t | \lambda) = D_0 \left[1 + 4\phi(\mathbf{r}, t) + \frac{20}{9} \lambda \phi(\mathbf{r}, t) \right. \\ \left. + \frac{2}{9} \lambda^2 \{1 - 2\phi(\mathbf{r}, t)\} \right]. \quad (45)$$

Note that the collective diffusion coefficient at infinite dilution is equal to the self-diffusion coefficient for long times, $t \gg 1/D_r$ (see the discussion below, eqn (7), in Section 2), just like for passive systems. An expression for the thermally averaged velocity, $\mathbf{v}(\mathbf{r}, t)$, of ABPs (including averaging with respect to the ABP orientations) can be found from the number conservation law $\partial\phi/\partial t = -\nabla \cdot [\mathbf{v}\phi]$, in combination with eqn (42) and (45):

$$\mathbf{v}(\mathbf{r}, t) = - \left[1 + 4\phi(\mathbf{r}, t) + \frac{20}{9} \lambda \phi(\mathbf{r}, t) \right. \\ \left. + \frac{2}{9} \lambda^2 \{1 - 2\phi(\mathbf{r}, t)\} \right] \nabla \phi(\mathbf{r}, t). \quad (46)$$

For large λ , where the last term dominates, the velocity $\mathbf{v} \sim (1 - 2\phi)\nabla\phi$ thus decreases (in the opposite direction of the spatial gradients) with increasing concentration due to binary ABP interactions. For high concentrations, a similar form, $\mathbf{v} \sim (1 - \zeta\phi)\nabla\phi$, is found in 2D simulations of disks, with $\zeta = 0.9$, essentially independent of v_0 and concentration.³⁵ Such a concentration-dependent velocity is introduced phenomenologically in, for example, ref. 30–34, and can be semi-empirically obtained from the Smoluchowski equation (in 2D), where ζ is related to an integral that involves the pair-correlation function, which remains unknown without specification of the pair-correlation function.^{25–29} At high concentrations, the decrease of the average translational velocity with the increasing concentration is at least in part due to multi-body interactions: when several ABPs approach each other head-on, the sliding motion along one of the neighbouring ABPs is blocked by the presence of the other ABPs. When such head-on interactions occur in large clusters, the dynamics of single ABPs is kinetically arrested. Such a diminishing translational motion due to multi-particle interactions is not accounted for in the above expression for the velocity,

since only binary collisions are considered. The expressions for the pair-correlation function derived in the present work may be used as a starting point to obtain the pair-correlation function for higher concentrations (for example, by means of an Ornstein-Zernike approach). This would establish a first-principles approach toward the dynamics and MIPS for ABPs with very short-ranged direct interactions.

For small Peclet numbers, it is found from the contact value in eqn (21) that the leading order contribution arising from motility varies like λ^2 . It may well be that the second-order contribution, g_2 , in the expansion of the pair-correlation function in eqn (18) contributes to the same order. We will therefore refrain here from the derivation of Fick's diffusion coefficient for small Peclet numbers. The diffusion equation as obtained from just the leading order term, g_1 , in eqn (18) is given at the end of Appendix C.

7 Summary and conclusions

Starting from the N -particle Smoluchowski equation, we derived analytic expressions for the low-density pair-correlation function for active Brownian particles (ABPs) in 3D for small and large swimming Peclet numbers, λ . The pair-interaction potential is assumed to be very short-ranged, consisting of an excluded-volume interaction and a possible sticky-like attractive interaction. In the analysis for large Peclet numbers, a distinction must be made between the front-sector, where two ABPs approach each other, and the wake-sector, where two ABPs move away from each other. The front-sector requires a singular perturbation analysis, leading to a boundary layer at ABP-contact with a spatial extent $\sim 1/\lambda$, and contact values of the order λ . Due to the small times during which the two ABPs interact while sliding over each other, rotational diffusion does not affect the pair-correlation function in the front-sector. The behaviour of the pair-correlation function in the wake-sector is quite opposite as compared to the front-sector: there is now a large region of extent $\sim \lambda^{1/3}$ where the pair-correlation function is essentially zero, and where rotational diffusion is solely responsible for the recovery of the pair-correlation function to unity at large distances. For passive colloids, such a pair-correlation function corresponds to a Janus particle with a very strong and short-ranged attraction at the front-sector side, and a strong but very long-ranged repulsion at the wake-side.

The same kind of behaviour of the pair-correlation function for large Peclet numbers is to be expected in 2D systems in case the orientations of the particles are also confined to the same 2D plane in which the particles move: a compact front-sector with large values of the pair-correlation function and an extended wake-sector with a near-zero pair-correlation function. In those 2D cases where rotation in all 3D directions is still possible, however, the physics may very well change, and the pair-correlation function may take a very different form.

The analytic expression for the pair-correlation function for large Peclet numbers is used to derive Fick's diffusion equation for the volume fraction of ABPs with excluded-volume interactions. We find that there is a preferred swimming direction



from high to low concentrations, which dominates the relaxation of concentration inhomogeneities over translational diffusion. Motility is thus expected to suppress thermodynamically driven spinodal decomposition, where diffusion occurs from regions of low to high concentrations due to strong attractive inter-particle interactions.

The present low-density calculation of the pair-correlation function obviously accounts only for binary collisions. It remains a future challenge to construct the pair-correlation function for concentrated systems of ABPs based on the low-density expressions as derived in the present paper. One possibility to achieve this extension is through an Ornstein-Zernike type of approach, where the low-density pair-correlation function identifies the direct-correlation function. This approach towards the dynamics and MIPS of ABPs would supplement the existing theories with a firm microscopic basis without the need for semi-empirical considerations and assumptions necessary to account for inter-particle correlations. The calculation of the pair-correlation function presented here also neglects hydrodynamic interactions. Even without hydrodynamic interactions, the analysis of pair-correlations for active systems poses a highly non-trivial problem. A further future challenge is to include hydrodynamic interactions, to be able to make a distinction between, for example, pullers and pushers.

The analysis of demixing kinetics and dynamics based on Fickian-like diffusion equations requires an additional extension, where higher-order spatial gradient contributions are included. These contributions stabilize the system against the demixing of short-wavelength concentration variations, that is, they prevent the unphysical growth of concentration variations with steep spatial gradients. Not including such higher-order gradient contributions still allows us to identify the phase-stability region, but excludes the analysis of demixing kinetics.

Conflicts of interest

There are no conflicts to declare.

Appendix A: The pair-correlation function for small Peclet numbers

The set of equations (eqn 19) is most conveniently solved by expanding the pair-correlation in terms of spherical harmonics:

$$g_1(\mathbf{R}, \hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2) = \sum_{l_1, m_1} \sum_{l_2, m_2} \sum_{p, q} F_{l_1, m_1, l_2, m_2}^{(p, q)}(R) Y_{l_1, m_1}(\hat{\mathbf{u}}_1) Y_{l_2, m_2}(\hat{\mathbf{u}}_2) Y_{p, q}^*(\hat{\mathbf{R}}). \quad (47)$$

The Laplace operator is rewritten as follows:

$$\tilde{\nabla}^2 = \frac{1}{R^2} \frac{\partial}{\partial R} \left(R^2 \frac{\partial}{\partial R} \right) + \frac{1}{R^2} \hat{\mathcal{R}}^2, \quad (48)$$

where $\hat{\mathcal{R}} = \hat{\mathbf{R}} \times \tilde{\nabla}_{\hat{\mathbf{R}}}$ is the rotation operator with respect to \mathbf{R} (where $\nabla_{\hat{\mathbf{R}}}$ is the gradient operator with respect to the Cartesian coordinates of the unit vector $\hat{\mathbf{R}}$). Using that,

$$\hat{\mathcal{R}}_1^2 Y_{l_1, m_1}(\hat{\mathbf{u}}_1) = -l_1(l_1 + 1) Y_{l_1, m_1}(\hat{\mathbf{u}}_1), \quad (49)$$

and similarly for $\hat{\mathcal{R}}^2 Y_{l_2, m_2}(\hat{\mathbf{u}}_2)$ and $\hat{\mathcal{R}}^2 Y_{p, q}(\hat{\mathbf{R}})$, and using orthonormality of the spherical harmonics, it is found that,

$$0 = \left[R^2 \frac{d^2}{dR^2} + 2R \frac{d}{dR} - \{k^2 R^2 + p(p + 1)\} \right] F_{l_1, m_1, l_2, m_2}^{(p, q)}(R), \quad (50)$$

where

$$k^2 = \frac{3}{2}[l_1(l_1 + 1) + l_2(l_2 + 1)]. \quad (51)$$

The solution of eqn (50) that tends to zero at infinity is as follows:

$$F_{l_1, m_1, l_2, m_2}^{(p, q)}(R) = C_{l_1, m_1, l_2, m_2}^{(p, q)} \frac{K_{p+\frac{1}{2}}(kR)}{\sqrt{kR}}, \quad (52)$$

where

$$K_{p+\frac{1}{2}}(x) = \sqrt{\frac{\pi}{2x}} \exp\{-x\} \sum_{s=0}^p \frac{(p+s)!}{s! (p-s)!} \frac{1}{(2x)^s}, \quad (53)$$

is a modified Bessel function. The integration constants $C_{l_1, m_1, l_2, m_2}^{(p, q)}$ must be determined such that the no-flux boundary condition is fulfilled. Using the addition theorem for spherical harmonics for the first-order Legendre polynomial (where Θ_1 is the angle between $\hat{\mathbf{R}}$ and $\hat{\mathbf{u}}_1$),

$$\hat{\mathbf{R}} \cdot \hat{\mathbf{u}}_1 = P_1(\cos \Theta_1) = \frac{4\pi}{3} \sum_{m=-1}^1 Y_{1, m}(\hat{\mathbf{u}}_1) Y_{1, m}^*(\hat{\mathbf{R}}), \quad (54)$$

and similarly for $\hat{\mathbf{R}} \cdot \hat{\mathbf{u}}_2$, and using that $Y_{0,0} = 1/\sqrt{4\pi}$, it is readily found from the orthonormality of the spherical harmonics that the no-flux boundary condition is fulfilled when

$$k \frac{d}{dx} \left(\frac{K_{3/2}(x)}{\sqrt{x}} \right) \Big|_{x=k} \times C_{1, m, 0, 0}^{(1, m)} = \frac{(4\pi)^{3/2}}{3}, \quad (55)$$

$$k \frac{d}{dx} \left(\frac{K_{3/2}(x)}{\sqrt{x}} \right) \Big|_{x=k} \times C_{0, 0, 1, m}^{(1, m)} = -\frac{(4\pi)^{3/2}}{3},$$

for $m = -1, 0, 1$, while all the remaining constants vanish, and $k = \sqrt{3}$. Hence,

$$g_1(\mathbf{R}, \hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2) = \frac{1}{\sqrt{3}} \left[\frac{d}{dx} \left(\frac{K_{3/2}(x)}{\sqrt{x}} \right) \right]_{x=\sqrt{3}}^{-1} \times \frac{K_{3/2}(\sqrt{3}R) \hat{\mathbf{R}} \cdot (\hat{\mathbf{u}}_1 - \hat{\mathbf{u}}_2)}{\sqrt{(\sqrt{3}R)}}. \quad (56)$$

Thus, from eqn (53), it is finally found that

$$g_1(\mathbf{R}, \hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2) = -C \left[\frac{\sqrt{3}}{R} + \frac{1}{R^2} \right] \times \exp\left\{-\sqrt{3}[R - 1]\right\} \hat{\mathbf{R}} \cdot (\hat{\mathbf{u}}_1 - \hat{\mathbf{u}}_2), \quad (57)$$

where

$$C = \frac{1}{3\sqrt{3}} \left[\frac{2}{3} + \frac{5}{3\sqrt{3}} \right]^{-1} \approx \frac{3}{25}. \quad (58)$$

This then leads to eqn (20) in the main text.



Appendix B: The pair-correlation function for large Peclet numbers

Due to rotational and translational invariance, the pair-correlation function depends on four variables: the distance, R , between the particles; (the cosine of) the angles between R and $\hat{\mathbf{u}}_{1,2}$, that is, $x_1 = \hat{\mathbf{R}} \cdot \hat{\mathbf{u}}_1$ and $x_2 = \hat{\mathbf{R}} \cdot \hat{\mathbf{u}}_2$; and (the cosine of) the angle between $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$, that is, $\alpha = \hat{\mathbf{u}}_1 \cdot \hat{\mathbf{u}}_2$. The differential equation (eqn (13)) in terms of these variables reads as follows:

$$0 = \text{Trans} + \text{Flux} + \text{Rot}, \quad (59)$$

where the contribution “Trans” from translational diffusion is equal to

$$\begin{aligned} \text{Trans} = \nabla^2 g &= \frac{1}{R^2} \frac{\partial}{\partial R} \left(R^2 \frac{\partial g}{\partial R} \right) \\ &+ \frac{1}{R^2} \frac{\partial}{\partial x_1} \left([1 - x_1^2] \frac{\partial g}{\partial x_1} \right) + \frac{1}{R^2} \frac{\partial}{\partial x_2} \left([1 - x_2^2] \frac{\partial g}{\partial x_2} \right) \\ &+ \frac{2}{R^2} [\alpha - x_1 x_2] \frac{\partial^2 g}{\partial x_1 \partial x_2}, \end{aligned} \quad (60)$$

the contribution “Flux” resulting from motility is equal to

$$\begin{aligned} \text{Flux} = -\lambda \mathbf{U} \cdot \nabla g &= -\lambda \left[\frac{1}{R} \frac{\partial g}{\partial x_1} (1 - \alpha - x_1^2 + x_1 x_2) \right. \\ &\left. + \frac{1}{R} \frac{\partial g}{\partial x_2} (\alpha - 1 + x_2^2 - x_1 x_2) + (x_1 - x_2) \frac{\partial g}{\partial R} \right], \end{aligned} \quad (61)$$

while the rotational contribution “Rot” is equal to

$$\begin{aligned} \text{Rot} = \frac{3}{2} \{ \hat{\mathcal{R}}_1^2 + \hat{\mathcal{R}}_2^2 \} g &= \frac{3}{2} \left\{ \frac{\partial}{\partial \alpha} \left([1 - \alpha^2] \frac{\partial g}{\partial \alpha} \right) \right. \\ &+ \frac{\partial}{\partial x_1} \left([1 - x_1^2] \frac{\partial g}{\partial x_1} \right) + 2(x_2 - \alpha x_1) \frac{\partial^2 g}{\partial x_1 \partial \alpha} \Big\}, \\ &+ \frac{3}{2} \left\{ \frac{\partial}{\partial \alpha} \left([1 - \alpha^2] \frac{\partial g}{\partial \alpha} \right) \right. \\ &+ \frac{\partial}{\partial x_2} \left([1 - x_2^2] \frac{\partial g}{\partial x_2} \right) + 2(x_1 - \alpha x_2) \frac{\partial^2 g}{\partial x_2 \partial \alpha} \Big\}. \end{aligned} \quad (62)$$

As will be shown below, the important related variables are as follows:

$$\begin{aligned} x &= x_1 - x_2 = \hat{\mathbf{R}} \cdot \mathbf{U} = U \cos \theta, \\ X &= x_1 + x_2, \end{aligned} \quad (63)$$

where

$$\mathbf{U} = \hat{\mathbf{u}}_1 - \hat{\mathbf{u}}_2, \quad (64)$$

and θ is the angle between \mathbf{R} and \mathbf{U} .

A little consideration shows that when $x < 0$, the two particles approach each other while for $x > 0$, they move apart. The region where $x < 0$ will be referred to as the front-sector, and where $x > 0$ as the wake-sector. The pair-correlation

function behaves very differently in both regions, as addressed in the main text on the basis of physical arguments. The mathematical analysis in the front-sector and wake-sector will be discussed separately in the following subsections.

7.1 The front-sector: $x < 0$

As will be shown below, the rotational contribution in eqn (62) is of lower order in the small parameter $1/\lambda$ as compared to the translational and flux contributions. Transforming (x_1, x_2) to (x, X) , and assuming that the pair-correlation function in this sector does not depend on X , the equation $0 = \text{Trans} + \text{Flux}$, with “Trans” and “Flux” given in eqn (60) and (61), becomes

$$\begin{aligned} 0 &= \frac{1}{R^2} \frac{\partial}{\partial R} \left(R^2 \frac{\partial g}{\partial R} \right) + \frac{1}{R^2} \frac{\partial}{\partial x} \left([U^2 - x^2] \frac{\partial g}{\partial x} \right) \\ &- \lambda \left[\frac{U^2 - x^2}{R} \frac{\partial g}{\partial x} + x \frac{\partial g}{\partial R} \right]. \end{aligned} \quad (65)$$

As will turn out, integration constants can be chosen such that the solution of this equation obeys both boundary conditions, which justifies the neglect of the X -dependence.

Multiplying the above differential equation by the small parameter $1/\lambda$ obviously shows that the highest-order derivative with respect to R is multiplied by this small parameter. This is the standard form of a singularly perturbed differential equation, the solution of which requires a boundary-layer analysis. Since the second-order derivative is needed to match two boundary conditions, this implies that the derivatives with respect to R are large near contact where $R \approx 1$. Thus, there is a small region close to contact where spatial derivatives are very large, which is referred to as the boundary layer. The fact that the variables x and α do not exhibit boundary-layer behavior will be verified below. Introducing the boundary-layer variable

$$S = \lambda^\nu (R - 1), \quad (66)$$

and insisting that the second-order derivative with respect to S in eqn (65) is of the same order in $1/\lambda$ as the flux contribution implies that

$$\nu = 1. \quad (67)$$

By introducing the variable S , the width $1/\lambda \ll 1$ of the boundary layer in terms of the original variable R is re-scaled to unity. This spatial stretching of the region close to contact renders the pair-correlation function a regular function of $1/\lambda$ when expressed in terms of S , and can therefore be expanded in a power series with respect to $1/\lambda$. Within the boundary layer, where $S = \lambda(R - 1) < 1$, eqn (65) thus reduces to

$$\frac{\partial^2 g}{\partial S^2} + \left(\frac{2}{\lambda} - x \right) \frac{\partial g}{\partial S} = 0. \quad (68)$$

As will turn out, the no-flux boundary condition can only be satisfied when the zeroth and the first-order contributions in $1/\lambda$ arising from derivatives with respect to S are kept. This is the reason why the first-order contribution $\sim 1/\lambda$ in the above differential equation is kept. The first-order contribution arising from $\partial g/\partial x$ is not needed to satisfy the boundary conditions,

which is therefore neglected. The solution of eqn (68) is

$$g = A + B \exp \left\{ -S \left(\frac{2}{\lambda} - x \right) \right\}, \quad (69)$$

where A and B are the integration constants. The no-flux boundary condition in eqn (13) in terms of the boundary variable S is

$$\frac{\partial g}{\partial S} - xg = 0, \text{ for } S = 0, \quad (70)$$

which leads to

$$g_{\text{inner}} = A \left[1 - \frac{1}{2} \lambda x \exp \left\{ -S \left(\frac{2}{\lambda} - x \right) \right\} \right]. \quad (71)$$

where the subscript “inner” is used to indicate that this is the solution within the boundary layer.

The outer solution, g_{outer} , where $S > 1$, is simply a constant, which is unity since for infinite distances between the two particles, the pair-correlation function tends to unity:

$$g_{\text{outer}} = 1. \quad (72)$$

Matching the inner and outer solutions implies that $A = 1$ in eqn (71) for the inner solution. It is thus finally found that, in terms of the original coordinates x and R ,

$$g_{\text{front}}(x, R, U) = 1 - \frac{1}{2} \lambda x \exp \{ (R - 1)[\lambda x - 2] \}, \quad (73)$$

where the index “front” is used to indicate that this solution is only valid within the front-sector. Recall that $x < 0$ within the front-sector.

Re-substituting this expression into eqn (59)–(62) shows that the above expression satisfies the differential equation to the leading order in $1/\lambda$, and by construction satisfies both boundary conditions. In doing so, it is important to notice that inside the boundary layer, $(R - 1) \sim 1/\lambda$, while outside the boundary layer, $g = 1$, where the exponential in eqn (73) is essentially zero. In particular, this justifies the neglect of the derivatives with respect to X and α in eqn (65), and thus shows that the rotational contribution can be neglected within the front-sector. The interpretation of the latter is that, for the large relative velocities under consideration, the time that the particles spent within the narrow boundary layer is sufficiently small that rotation does not occur. Furthermore, for $x = 0$ (and in particular, for $U = 0$), the pair-correlation function is identically equal to unity, as it should be.

7.2 The wake-sector: $x > 0$

The asymptotic solution in eqn (73) diverges for large distances when $x > 0$. This shows that the behavior in the wake-sector is quite different compared to that in the front-sector. On the basis of the discussion in the main text, spatial derivatives in the direction along the relative velocity $\sim U$ of the two particles are expected to be very small for large λU . We therefore introduce the scaled variable (not to be confused with the

S -coordinate introduced in the previous subsection),

$$S = f(\alpha) \frac{Rx}{\lambda U^2} = f(\alpha) \frac{R}{\lambda U} \cos \Theta, \quad (74)$$

where the function $f(\alpha)$ will be specified below. As a second variable, the distance ρ between the two particles perpendicular to \mathbf{U} ,

$$\rho = \frac{R}{U} \sqrt{U^2 - x^2} = R \sin \Theta, \quad (75)$$

is introduced. The reason for the introduction of this cylindrical variable is, as discussed in the main text, that a step-function-like behavior at $\rho = 1$ is expected for small separations between the particles, since diffusion has not yet been effective for the large relative velocities under consideration. As will be shown below, an approximate solution for the boundary-value problem within the wake-sector can be found, assuming that the pair-correlation function depends on R, x_1, x_2 , and α only through the combinations of S and ρ .

It is found by direct differentiation with respect to the Cartesian coordinates of \mathbf{R} , or from eqn (60), that

$$\text{Trans} = \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial g(\rho, S)}{\partial \rho} \right), \quad (76)$$

where terms of the order $1/\lambda$ corresponding to the differentiations with respect to S have been neglected. The differentiation with respect to S in the contribution to the flux in eqn (61) is multiplied with λ , and should therefore be kept. It is readily found that

$$\text{Flux} = -f(\alpha) \frac{\partial g(\rho, S)}{\partial S}. \quad (77)$$

Contrary to the front-sector, the rotational contribution cannot be neglected in the wake-sector. With some effort, the rotational contribution is found to be given by (the most efficient way to compute the differentiations is to first introduce the variable ρ^2 , instead of just ρ , and, after performing all differentiations, express the differentiations with respect to ρ^2 in terms of ρ)

$$\begin{aligned} \text{Rot} = & \left(\frac{Rx}{U} \right)^2 \frac{1}{U^2} \left[12 - \frac{U^2}{x^2} \left\{ 6 - \frac{3}{2} (x^2 + X^2) \right\} \right] \frac{1}{2\rho} \frac{\partial g(\rho, S)}{\partial \rho} \\ & + \left(\frac{Rx}{U} \right)^4 \frac{1}{U^2} \left[-6(1 + \alpha) + \frac{2U^2}{x^2} \left\{ 6 - \frac{3}{2} (x^2 + X^2) \right\} \right] \\ & \times \frac{1}{4\rho} \frac{\partial}{\partial \rho} \left(\frac{1}{\rho} \frac{\partial g(\rho, S)}{\partial \rho} \right), \end{aligned} \quad (78)$$

where, as before, $X = x_1 + x_2$. Here, the terms of the order R/λ resulting from differentiations with respect to S have been neglected. As will turn out, the pair-correlation function relaxes to unity for $R \sim \lambda^{1/3}$, so that the restriction of the analysis given below to distances where $R/\lambda \ll 1$ is not severe. In addition, it will turn out that the solution that we find converges to unity for large distances, as it should be, which thus correctly interpolates between $R \ll \lambda$ and $R \rightarrow \infty$. The variables R and x can be



expressed in terms of S , ρ , and U , using that

$$\begin{aligned} \frac{U^2}{x^2} &= 1 + \rho^2 \left(\frac{f(\alpha)}{\lambda US} \right)^2, \\ \frac{Rx}{U} &= \frac{\lambda US}{f(\alpha)}, \end{aligned} \quad (79)$$

leading to (here we leave X as it occurred before)

$$\begin{aligned} \text{Rot} &= -\frac{\rho}{2U^2} \frac{\partial g(\rho, S)}{\partial \rho} \left[6 - \frac{3}{2} X^2 \right] \\ &+ \frac{1}{2U^2} \left(\frac{\lambda US}{f(\alpha)} \right)^2 \left[\frac{1}{\rho} \frac{\partial g(\rho, S)}{\partial \rho} \{3 - 3\alpha + 3X^2\} \right. \\ &\left. + \frac{\partial^2 g(\rho, S)}{\partial \rho^2} \left\{ 6 - \frac{3}{2} X^2 \right\} \right] \\ &- \frac{3X^2}{4U^2} \left(\frac{\lambda US}{f(\alpha)} \right)^4 \left[-\frac{1}{\rho^3} \frac{\partial g(\rho, S)}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2 g(\rho, S)}{\partial \rho^2} \right]. \end{aligned} \quad (80)$$

Adding all contributions leads to a differential equation that we were unable to solve analytically. Instead, a solution will first be obtained in the region where $(\lambda US/f) = Rx/U = R \cos \theta \gg 1$ (note that this further limits the region where the analysis is valid, on top of the earlier condition $R \ll \lambda$). This is the region where the most significant relaxation of the pair-correlation function occurs. At first sight, only the last term in eqn (80) for the rotational contribution survives in this region. However, a little thought shows that

$$\begin{aligned} X^2 &= \frac{V^2 \rho^2}{\rho^2 + (\lambda US/f)^2} \approx \frac{V^2 \rho^2}{(\lambda US/f)^2}, \\ V^2 &= (2 + 2\alpha) \cos^2 \phi, \end{aligned} \quad (81)$$

where ϕ is the angle between $\tilde{\mathbf{U}} = \hat{\mathbf{u}}_1 + \hat{\mathbf{u}}_2$ (the magnitude of which is $\tilde{U} = \sqrt{2 + 2\alpha}$) and the projection of \mathbf{R} onto the plane perpendicular to \mathbf{U} . For $\lambda US/f \gg 1$, it follows that the last term in the rotational contribution in eqn (80) is of the order $(\lambda US/f)^2$ instead of the order $(\lambda US/\lambda)^4$. Expanding eqn (80) with respect to $(f/\lambda US) \ll 1$ thus leads to

$$\text{Rot} = \frac{3}{4U^2} \left(\frac{\lambda US}{f(\alpha)} \right)^2 \left[\frac{1}{\rho} \frac{\partial g(\rho, S)}{\partial \rho} \{U^2 + V^2\} + \frac{\partial^2 g(\rho, S)}{\partial \rho^2} \{4 - V^2\} \right]. \quad (82)$$

Since V varies between 0 and $2 + 2\alpha$ (depending on the angle ϕ), the combinations $U^2 + V^2$ and $4 - V^2$ both vary between $2 - 2\alpha$ and 4. For $\alpha = 1$, we have $U = 0$, for which the pair-correlation function must be trivially equal to 1 (since the relative velocity is 0 in that case). The largest relative velocity is attained for $\alpha = -1$, in which case $U^2 + V^2 = 4 - V^2 = 4$. We thus restrict the validity of our analysis further to values of α which are sufficiently negative, such that $U^2 + V^2$ and $4 - V^2$ are not very different. For such values of α , one can average V^2 with respect to the angle ϕ , for which both $U^2 + V^2$ and $4 - V^2$ take the same value equal to $3 - \alpha$, which also equals the average of the extreme values that both combinations

take. To quantify this further, note that

$$\begin{aligned} \frac{1}{\rho} \frac{\partial g(\rho, S)}{\partial \rho} \{U^2 + V^2\} + \frac{\partial^2 g(\rho, S)}{\partial \rho^2} \{4 - V^2\} &= \\ (3 - \alpha) \left[\frac{1}{\rho} \frac{\partial g(\rho, S)}{\partial \rho} + \frac{\partial^2 g(\rho, S)}{\partial \rho^2} \right] \\ - (1 + \alpha)(1 - 2 \cos^2 \phi) \left[\frac{1}{\rho} \frac{\partial g(\rho, S)}{\partial \rho} - \frac{\partial^2 g(\rho, S)}{\partial \rho^2} \right]. \end{aligned} \quad (83)$$

“Sufficiently negative α ” thus refers to the values of α where $3 - \alpha$ is at least an order of magnitude larger than $1 + \alpha$. For $\alpha < -7/11$, for example, $3 - \alpha$ is more than ten times larger than $(1 + \alpha)(1 - 2 \cos^2 \phi)$, for which the neglect of the last term on the right-hand side in eqn (83) against the first term is justified. The further analysis is thus restricted to angles between $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ larger than about 130° . Note that $3 - \alpha$ is never smaller than $(1 + \alpha)(1 - 2 \cos^2 \phi)$, for any α , so that the second term on the right-hand side in eqn (83) is never dominant over the first term.

Since the contribution “Trans” in eqn (76) is of the zeroth order in $\lambda US/f(\alpha)$, we have $\text{Flux} + \text{Rot} = 0$ for the large distances under consideration. It thus follows from eqn (77), (82) and (83), with the neglect of the last term in eqn (83), that

$$\begin{aligned} f(\alpha) \frac{\partial g(\rho, S)}{\partial S} &= \frac{3}{4} (3 - \alpha) \left(\frac{\lambda S}{f(\alpha)} \right)^2 \\ &\times \left[\frac{1}{\rho} \frac{\partial g(\rho, S)}{\partial \rho} + \frac{\partial^2 g(\rho, S)}{\partial \rho^2} \right]. \end{aligned} \quad (84)$$

Choosing the function $f(\alpha)$ as

$$f(\alpha) = \left[\frac{3}{4} (3 - \alpha) \right]^{1/3}, \quad (85)$$

reduces eqn (84) to

$$\frac{\partial g(\rho, S)}{\partial S} = \lambda^2 S^2 \left[\frac{1}{\rho} \frac{\partial g(\rho, S)}{\partial \rho} + \frac{\partial^2 g(\rho, S)}{\partial \rho^2} \right]. \quad (86)$$

As will be seen below, this choice of $f(\alpha)$ leads to a solution of the differential equation that satisfies the required boundary conditions. Separation of variables, writing $g(\rho, S) = F(S)H(\rho)$, gives

$$\begin{aligned} \frac{dF(S)}{dS} &= -C^2 \lambda^2 S^2 F(S), \\ \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial H(\rho)}{\partial \rho} \right) &= -C^2 H(\rho), \end{aligned} \quad (87)$$

where C^2 is a constant. The solutions of these equations are as follows:

$$F(S) \sim \exp \left\{ -\frac{C^2}{3} \lambda^2 S^3 \right\}, \quad (88)$$

$$H(\rho) \sim J_0(C\rho),$$

where J_0 is the zeroth order Bessel function of the first kind. Hence,

$$g = 1 - \int_0^\infty dC A(C) J_0(C\rho) \exp \left\{ -\frac{C^2}{3} \lambda^2 S^3 \right\}, \quad (89)$$



where unity is added in order to satisfy the boundary condition at infinity (the integral tends to zero for large S and ρ).

The final expression for the pair-correlation function has been subjected to the conditions that $R/\lambda \ll 1$, $Z = R \cos \Theta \gg 1$, and $\alpha < -7/11$. The above expression for the pair-correlation function satisfies the boundary condition $g \rightarrow 1$ for $R \rightarrow \infty$, and therefore correctly interpolates between distances where R/λ is small and where $R \rightarrow \infty$. Furthermore, as will be seen below, the function $A(C)$ can be chosen such that $g = 1$ for $Z = 0$, as it should be. Hence, the above expression for the pair-correlation function also correctly interpolates between large values of Z and $Z = 0$. Finally, for $U = 0$ (for which $\alpha = +1$ and $S \rightarrow \infty$), $g = 1$, as it should be, so that there is a correct interpolation with respect to the values of α not smaller than $-7/11$ as well. The interpolation arguments are very similar to the matching procedure of the inner and outer solutions in the boundary-layer analysis of the front-sector as discussed in the previous subsection. The outer region in the wake-sector corresponds to the region where the conditions $R/\lambda \ll 1$, $R \cos \Theta \gg 1$, and $\alpha < -7/11$ are not satisfied.

The function $A(C)$ is first chosen such that the expected step-function-like behavior at $\rho = 1$ for $S \ll 1$ is reproduced. This can be achieved using the following integral identity⁵⁶:

$$\int_0^\infty dC J_1(C) J_0(C\rho) = 0, \rho > 1, \\ = 1/2, \rho = 1, \\ = 1, \rho < 1, \quad (90)$$

where J_1 is the first-order Bessel function of the first kind. This identity implies that the step-function-like behavior is obtained for $A(C) = J_1(C)$, so that

$$g = 1 - \int_0^\infty dC J_1(C) J_0(C\rho) \exp\left\{-\frac{C^2}{3}\lambda^2 S^3\right\}. \quad (91)$$

As a last step, the no-flux boundary condition (eqn 13) should be verified, which, in terms of the variables R and x , reads as follows:

$$\frac{\partial g(x, R, U)}{\partial R} - \lambda x g(x, R, U) = 0, \text{ for } R = 1. \quad (92)$$

First consider the contact value, $g^c = g(x, R = 1, U)$, of the pair-correlation function, which is equal to

$$g^c(\cos \Theta, w) = 1 - \int_0^\infty dC J_1(C) J_0(C\sqrt{1 - \cos^2 \Theta}) \\ \times \exp\{-C^2 w \cos^3 \Theta\}. \quad (93)$$

where, as before, Θ is the angle between \mathbf{R} and \mathbf{U} , and where

$$w = \frac{14 + U^2}{\lambda - 8U^3}. \quad (94)$$

As will turn out, both terms in eqn (92) approach zero with decreasing values of w , *i.e.*, increasing values of λ .

The combination $(\cos \Theta/w)g^c \sim \lambda x g^c$ is plotted in Fig. 8a as a function of $w \sim 1/\lambda$ for various values of $\cos \Theta$, as indicated in the plots. As can be seen, $\lambda x g^c$ tends to zero for small w (corresponding to large λ). The particular value of w for which $\lambda x g^c$ tends to zero, however, decreases with decreasing values of $\cos \Theta$, and goes through a maximum. The point of contact where $\cos \Theta$ is zero, and therefore $\rho = \sqrt{1 - \cos^2 \Theta} = 1$, corresponds to a relative position for which $\mathbf{R} \perp \mathbf{U}$, that is, where the two particles just move from the front-sector to the wake-sector. For these positions at

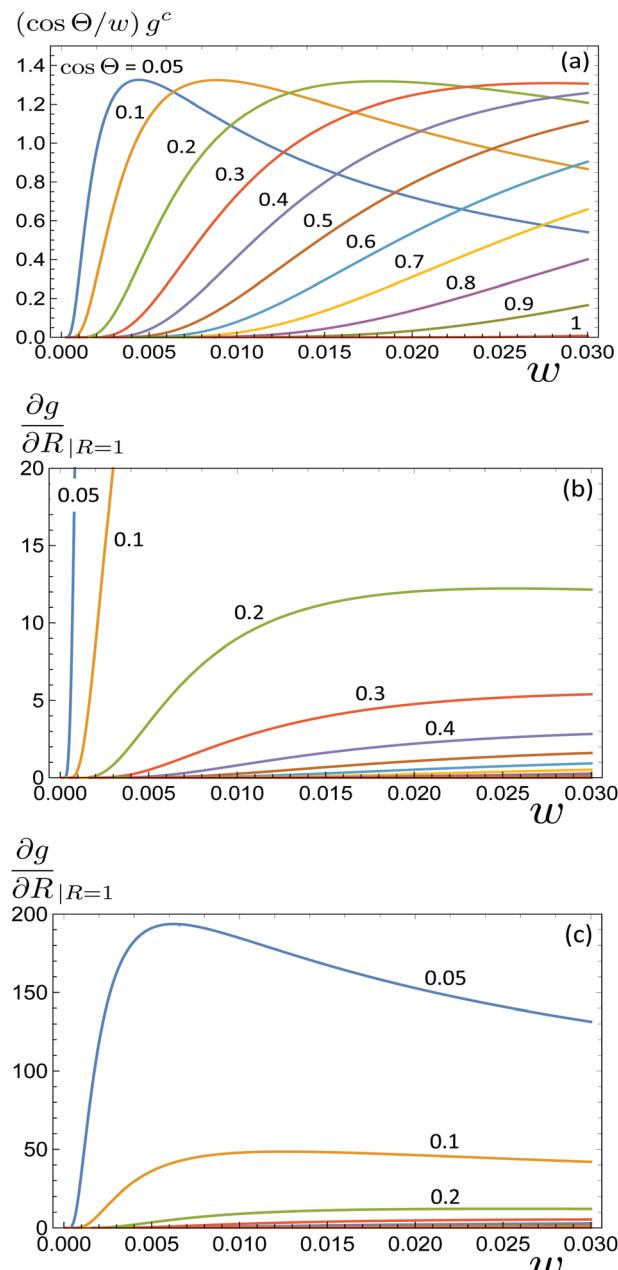


Fig. 8 (a) The combination $(\cos \Theta/w)g^c \sim \lambda x g^c$ (where Θ is the angle between \mathbf{R} and \mathbf{U} , $w \sim 1/\lambda$ given in eqn (94), and g^c is the contact value of the pair-correlation function) as a function of w , for various values of $\cos \Theta$, as indicated in the figure. (b and c) The radial derivative of the pair-correlation function at contact as a function of w for two different ranges of the derivative.



contact, diffusion becomes almost immediately effective, leading to a deviation from the step-function at $\rho = 1$. This is the physical origin of the relative large velocities required for which the contact value becomes zero for small values of $\cos \Theta$. Note that these numerical results imply that the contact value of the pair-correlation function tends to zero as $\lambda^{-\gamma}$, with $\gamma > 1$. Within the leading order expansion with respect to $1/\lambda$ considered here, the contact value is therefore zero.

The spatial derivative in eqn (92) is plotted in Fig. 8b and c (with different scales of the vertical axes), again as a function of w for various values of $\cos \Theta$. The spatial derivatives are seen to tend to zero for large values of λ . For the same reason as for the contact value of the pair-correlation function, the spatial derivative for small values of $\cos \Theta$ requires larger values of λ to converge to zero. Again, this is due to the step-function-like behavior at the point where $\rho = \sqrt{1 - \cos^2 \Theta} \approx 1$ and $\mathbf{R} \perp \mathbf{U}$.

Both terms in the no-flux boundary condition (eqn 92) thus tend to zero for large λ .

In terms of the original (dimensionless) coordinates, eqn (91) finally leads to the expression (eqn 24) as given in the main text.

Appendix C: Mathematical details concerning collective diffusion

First consider the derivation of Fick's diffusion equation for large Peclet numbers. Substituting eqn (32) into the equation of motion (eqn 30) gives

$$\begin{aligned} \frac{\partial}{\partial t} \phi(\mathbf{r}, \hat{\mathbf{u}}, t) &= [D_0 \nabla^2 + D_r \hat{\mathcal{R}}^2] \phi(\mathbf{r}, \hat{\mathbf{u}}, t) - \nabla \cdot [\phi(\mathbf{r}, \hat{\mathbf{u}}, t) v_0 \hat{\mathbf{u}}] \\ &+ 3 \frac{D_0}{a} \nabla \cdot \left[\phi(\mathbf{r}, \hat{\mathbf{u}}, t) \oint d\hat{\mathbf{u}}' \phi(\mathbf{r}, \hat{\mathbf{u}}', t) \left(1 + \frac{1}{3} \lambda U \right) \frac{\mathbf{U}}{U} \right] \\ &+ \frac{3}{4} D_0 \nabla \cdot \left[\phi(\mathbf{r}, \hat{\mathbf{u}}, t) \nabla \oint d\hat{\mathbf{u}}' \phi(\mathbf{r}, \hat{\mathbf{u}}', t) \left(\frac{16}{3} + \lambda U \right) \right] \\ &+ \frac{3}{4} D_0 \nabla \cdot \left[\phi(\mathbf{r}, \hat{\mathbf{u}}, t) \nabla \cdot \oint d\hat{\mathbf{u}}' \phi(\mathbf{r}, \hat{\mathbf{u}}', t) \lambda \frac{\mathbf{U} \mathbf{U}}{U} \right], \end{aligned} \quad (95)$$

where the orientation-dependent volume fraction, $\phi(\mathbf{r}, \hat{\mathbf{u}}, t)$, has been introduced in Section 6.

Next, $\phi(\mathbf{r}, \hat{\mathbf{u}}, t)$ is expanded in terms of orthonormal polyadic products of $\hat{\mathbf{u}}$ (which is equivalent to an expansion with respect to spherical harmonics):

$$\phi(\mathbf{r}, \hat{\mathbf{u}}, t) = F_0(\mathbf{r}, t) Q_0 + \mathbf{F}_1(\mathbf{r}, t) \cdot \mathbf{Q}_1(\hat{\mathbf{u}}) + \mathbf{F}_2(\mathbf{r}, t) : \mathbf{Q}_2(\hat{\mathbf{u}}) + \dots \quad (96)$$

The first few orthonormal polyadic products read as follows:

$$Q_0 = \sqrt{\frac{1}{4\pi}},$$

$$Q_{1;i} = \sqrt{\frac{3}{4\pi}} \hat{u}_i,$$

$$Q_{2:ij} = \sqrt{\frac{15}{8\pi}} \left[\hat{u}_i \hat{u}_j - \frac{1}{3} \delta_{ij} \right].$$

For $n \geq 2$, the orthonormal products are zero when contracted with respect to any two of its indices, so that $\mathbf{F}_{n \geq 2}$ is to zero when contracted with respect to any two of its indices without loss of generality. The polyadic products are symmetric upon interchanging any two of its indices, and therefore $\mathbf{F}_{n \geq 2}$ is similarly symmetric. These polyadic products are orthonormal in the sense that (\odot) stands for the contraction with respect to all adjacent indices

$$\begin{aligned} \oint d\hat{\mathbf{u}} [\mathbf{F}_n(\mathbf{r}, t) \odot \mathbf{Q}_n(\hat{\mathbf{u}})] \mathbf{Q}_m(\hat{\mathbf{u}}) &= \mathbf{F}_n(\mathbf{r}, t), \text{ for } n = m, \\ &= \mathbf{0}, \text{ for } n \neq m, \end{aligned} \quad (97)$$

where $\mathbf{0}$ is the matrix with zero entries. The term $\mathbf{Q}_1(\hat{\mathbf{u}}) \sim \hat{\mathbf{u}}$ in eqn (96) will contribute significantly to Fick's diffusion equation in case there is a preferred polar orientation induced by a concentration gradient, especially for large velocities. The last term in eqn (96) merely characterizes the degree of nematic ordering, which contributes much less: for perfect nematic order, for example, the two particles move along with each other, so that there is no effect of activity on the pair-correlation function. From the above orthonormality relations, it is readily found that eqn (96) reduces to

$$\phi(\mathbf{r}, \hat{\mathbf{u}}, t) = \frac{1}{4\pi} \phi(\mathbf{r}, t) + \frac{3}{4\pi} \mathbf{P}(\mathbf{r}, t) \cdot \hat{\mathbf{u}} + \dots, \quad (98)$$

where

$$\begin{aligned} \phi(\mathbf{r}, t) &= \oint d\hat{\mathbf{u}} \phi(\mathbf{r}, \hat{\mathbf{u}}, t), \\ \mathbf{P}(\mathbf{r}, t) &= \oint d\hat{\mathbf{u}} \hat{\mathbf{u}} \phi(\mathbf{r}, \hat{\mathbf{u}}, t), \end{aligned} \quad (99)$$

which reproduces eqn (36) in the main text.

In order to evaluate the integrals in eqn (95), the following integral identities must be used:

$$\begin{aligned} \oint d\hat{\mathbf{u}}' \frac{1}{|\hat{\mathbf{u}} - \hat{\mathbf{u}}'|} &= 4\pi, \\ \oint d\hat{\mathbf{u}}' \frac{\hat{\mathbf{u}}'}{|\hat{\mathbf{u}} - \hat{\mathbf{u}}'|} &= \frac{4\pi}{3} \hat{\mathbf{u}}, \\ \oint d\hat{\mathbf{u}}' \frac{\hat{\mathbf{u}}' \hat{\mathbf{u}}'}{|\hat{\mathbf{u}} - \hat{\mathbf{u}}'|} &= \frac{16\pi}{15} \left[\hat{\mathbf{I}} + \frac{3}{4} \hat{\mathbf{u}} \hat{\mathbf{u}} \right], \\ \oint d\hat{\mathbf{u}}' \frac{\hat{u}_i' \hat{u}_j' \hat{u}_k'}{|\hat{\mathbf{u}} - \hat{\mathbf{u}}'|} &= \frac{4\pi}{7} \hat{u}_i \hat{u}_j \hat{u}_k + \frac{16\pi}{105} [\delta_{ij} \hat{u}_k + \delta_{ik} \hat{u}_j + \delta_{jk} \hat{u}_i], \\ \oint d\hat{\mathbf{u}}' |\hat{\mathbf{u}} - \hat{\mathbf{u}}'| &= \frac{16\pi}{3}, \\ \oint d\hat{\mathbf{u}}' \hat{\mathbf{u}}' |\hat{\mathbf{u}} - \hat{\mathbf{u}}'| &= -\frac{16\pi}{15} \hat{\mathbf{u}}. \end{aligned} \quad (100)$$

Together with the expansion (eqn 98), these identities lead to the following results relating to each of the contributions in



eqn (95):

$$\begin{aligned}
 J_1 &= \oint d\hat{\mathbf{u}}' \phi(\mathbf{r}, \hat{\mathbf{u}}', t) \frac{\mathbf{U}}{U} = \frac{2}{3}\phi(\mathbf{r}, t)\hat{\mathbf{u}} - \frac{2}{5}\mathbf{P}(\mathbf{r}, t) \cdot [2\hat{\mathbf{I}} - \hat{\mathbf{u}}\hat{\mathbf{u}}], \\
 J_2 &= \oint d\hat{\mathbf{u}}' \phi(\mathbf{r}, \hat{\mathbf{u}}', t) \mathbf{U} = \phi(\mathbf{r}, t)\hat{\mathbf{u}} - \mathbf{P}(\mathbf{r}, t), \\
 J_3 &= \oint d\hat{\mathbf{u}}' \phi(\mathbf{r}, \hat{\mathbf{u}}', t) = \phi(\mathbf{r}, t), \\
 J_4 &= \oint d\hat{\mathbf{u}}' \phi(\mathbf{r}, \hat{\mathbf{u}}', t) U = \frac{4}{3}\phi(\mathbf{r}, t) - \frac{4}{5}\mathbf{P}(\mathbf{r}, t) \cdot \hat{\mathbf{u}}, \\
 J_5 &= \oint d\hat{\mathbf{u}}' \phi(\mathbf{r}, \hat{\mathbf{u}}', t) \frac{\mathbf{U}\mathbf{U}}{U} = -\frac{24}{35}[\mathbf{P}(\mathbf{r}, t)\hat{\mathbf{u}} + \hat{\mathbf{u}}\mathbf{P}(\mathbf{r}, t)] \\
 &\quad + \frac{4}{15}\left[\phi(\mathbf{r}, t) + \frac{3}{7}\mathbf{P}(\mathbf{r}, t) \cdot \hat{\mathbf{u}}\right][\hat{\mathbf{I}} + 2\hat{\mathbf{u}}\hat{\mathbf{u}}]. \tag{101}
 \end{aligned}$$

Substituting these results into eqn (95) and integrating with respect to $\hat{\mathbf{u}}$ leads, for each of the separate terms contributing to the equation of motion for $\phi(\mathbf{r}, t)$, to (here it is understood, for brevity, that in the right-hand sides, $\phi \equiv \phi(\mathbf{r}, t)$ and $P \equiv P(\mathbf{r}, t)$)

$$\begin{aligned}
 I_{\phi,1} &= 3\frac{D_0}{a}\nabla \cdot \oint d\hat{\mathbf{u}} \left[\phi(\mathbf{r}, \hat{\mathbf{u}}, t) \oint d\hat{\mathbf{u}}' \phi(\mathbf{r}, \hat{\mathbf{u}}', t) \frac{\mathbf{U}}{U} \right] = 0, \\
 I_{\phi,2} &= \lambda\frac{D_0}{a}\nabla \cdot \oint d\hat{\mathbf{u}} \left[\phi(\mathbf{r}, \hat{\mathbf{u}}, t) \oint d\hat{\mathbf{u}}' \phi(\mathbf{r}, \hat{\mathbf{u}}', t) \mathbf{U} \right] = 0, \\
 I_{\phi,3} &= 4D_0\nabla \cdot \oint d\hat{\mathbf{u}} \left[\phi(\mathbf{r}, \hat{\mathbf{u}}, t) \nabla \oint d\hat{\mathbf{u}}' \phi(\mathbf{r}, \hat{\mathbf{u}}', t) \right] \\
 &= 4D_0\nabla \cdot [\phi\nabla\phi], \\
 I_{\phi,4} &= \frac{3}{4}\lambda D_0\nabla \cdot \oint d\hat{\mathbf{u}} \left[\phi(\mathbf{r}, \hat{\mathbf{u}}, t) \nabla \oint d\hat{\mathbf{u}}' \phi(\mathbf{r}, \hat{\mathbf{u}}', t) U \right] \tag{102} \\
 &= \lambda D_0\nabla \cdot [\phi\nabla\phi] - \frac{3}{10}\lambda D_0\nabla^2 P^2, \\
 I_{\phi,5} &= \frac{3}{4}\lambda D_0\nabla \cdot \oint d\hat{\mathbf{u}} \left[\phi(\mathbf{r}, \hat{\mathbf{u}}, t) \nabla \cdot \oint d\hat{\mathbf{u}}' \phi(\mathbf{r}, \hat{\mathbf{u}}', t) \frac{\mathbf{U}\mathbf{U}}{U} \right] \\
 &= \frac{1}{3}\lambda D_0\nabla \cdot [\phi\nabla\phi] \\
 &\quad + \frac{3}{50}\lambda D_0\nabla \cdot [\nabla V^2 - 8\nabla \cdot \mathbf{P}\mathbf{P}].
 \end{aligned}$$

Hence, from

$$\frac{\partial}{\partial t}\phi(\mathbf{r}, t) = D_0\nabla^2\phi(\mathbf{r}, t) + \sum_{n=1}^5 I_{\phi,n}, \tag{103}$$

keeping only the leading order terms for large swimming velocities leads to eqn (37).

The equation of motion for \mathbf{P} is obtained by multiplying both sides of eqn (95) with $\hat{\mathbf{u}}$ and then integrate with respect to $\hat{\mathbf{u}}$. The separate terms that contribute to the equation of motion

for \mathbf{P} are similarly found to be equal to

$$\begin{aligned}
 I_{P,1} &= \frac{D_0}{a}\nabla \left[\frac{2}{3}\phi^2 + \frac{6}{25}V^2 \right] - \frac{D_0}{a}\frac{48}{25}\nabla \cdot \mathbf{P}\mathbf{P}, \\
 I_{P,2} &= \lambda\frac{D_0}{a}\frac{1}{3}\nabla\phi^2 - \lambda\frac{D_0}{a}(\nabla \cdot \mathbf{P}\mathbf{P}), \\
 I_{P,3} &= 4D_0[(\nabla\phi) \cdot (\nabla\mathbf{P}) + \mathbf{P}\nabla^2\phi], \\
 I_{P,4} &= -\lambda D_0\frac{1}{5}\nabla \cdot [\phi\nabla\mathbf{P}] + \lambda D_0[(\nabla\phi) \cdot (\nabla\mathbf{P}) + \mathbf{P}\nabla^2\phi], \\
 I_{P,5} &= \lambda D_0\frac{1}{25}[-4(\nabla\mathbf{P}) \cdot (\nabla\phi) - 4\nabla(\phi\nabla \cdot \mathbf{P}) \\
 &\quad - 4\phi\nabla(\nabla \cdot \mathbf{P}) + 7\mathbf{P}\nabla^2\phi + 7(\nabla\phi) \cdot (\nabla\mathbf{P}) \\
 &\quad + 2\nabla(\mathbf{P} \cdot \nabla\phi) + 2\nabla \cdot (\mathbf{P}\nabla\phi) + \nabla \cdot (\phi\nabla\mathbf{P})]. \tag{104}
 \end{aligned}$$

Furthermore, using

$$\hat{\mathbf{R}}^2\mathbf{Q}_n(\hat{\mathbf{u}}) = -n(n+1)\mathbf{Q}_n(\hat{\mathbf{u}}), \tag{105}$$

for $n = 1$ for the evaluation of the second term in eqn (95), including the third contribution, and adding all these integrals (similar to eqn (103)), keeping only the leading contributions for large λ , leads to the equation of motion (eqn 38) for \mathbf{P} .

The derivation of the diffusion equation for small Peclet numbers requires relatively little effort, due to the much simpler expressions:

$$\begin{aligned}
 \mathbf{I}_1(\hat{\mathbf{u}}, \hat{\mathbf{u}}') &= \frac{1 + \sqrt{3}}{5 + 2\sqrt{3}}\frac{32\pi}{3}a^3\lambda\mathbf{U}, \\
 \mathbf{I}_2(\hat{\mathbf{u}}, \hat{\mathbf{u}}') &= \frac{64\pi}{3}a^4\hat{\mathbf{I}}, \text{ for } \lambda \ll 1, \tag{106}
 \end{aligned}$$

as obtained from the contact value of the pair-correlation function in eqn (21). From these expressions, the equation of motion,

$$\frac{\partial}{\partial t}\phi(\mathbf{r}, t) = D_0\nabla \cdot [\{1 + 8\phi(\mathbf{r}, t)\}\nabla\phi(\mathbf{r}, t)] - v_0\nabla \cdot \mathbf{P}(\mathbf{r}, t), \tag{107}$$

is obtained (without having to resort to the expansion in eqn (98)). It is similarly found that (with $C = (1 + \sqrt{3})/(5 + 2\sqrt{3})$)

$$\begin{aligned}
 \frac{\partial}{\partial t}\mathbf{P}(\mathbf{r}, t) &= -2D_r\mathbf{P} + D_0\nabla^2\mathbf{P}(\mathbf{r}, t) - \frac{1}{3}v_0\nabla\phi(\mathbf{r}, t) \\
 &\quad + C\frac{4D_0}{3a}\lambda\nabla\phi^2(\mathbf{r}, t) - 4C\frac{D_0}{a}\lambda\nabla \cdot (\mathbf{P}(\mathbf{r}, t)\mathbf{P}(\mathbf{r}, t)) \\
 &\quad + 8D_0\mathbf{P}(\mathbf{r}, t)\nabla^2\phi(\mathbf{r}, t) + 8D_0[\nabla\phi(\mathbf{r}, t)] \cdot [\nabla\mathbf{P}(\mathbf{r}, t)]. \tag{108}
 \end{aligned}$$

Keeping only the leading order terms in spatial gradients, noting that \mathbf{P} is at least first order in such gradients, similar to what has been done for the large Peclet numbers above, this equation of motion reduces to

$$\frac{\partial}{\partial t}\mathbf{P}(\mathbf{r}, t) = -2D_r\mathbf{P} - \frac{1}{3}v_0\nabla\phi(\mathbf{r}, t) + C\frac{4D_0}{3a}\lambda\nabla\phi^2(\mathbf{r}, t) \tag{109}$$



As before, \mathbf{P} is enslaved by ϕ , and hence,

$$\mathbf{P}(\mathbf{r}, t) = -\frac{2}{9}\lambda a\{1 - 8C\phi(\mathbf{r}, t)\}\nabla\phi(\mathbf{r}, t). \quad (110)$$

Substituting this result into eqn (107) leads to Fick's diffusion eqn (42), where

$$D_c^{\text{eff}}(\mathbf{r}, t|\lambda) = D_0\{1 + 8\phi(\mathbf{r}, t)\} + \frac{2}{9}D_0\lambda^2\left\{1 + 8\frac{1 + \sqrt{3}}{5 + 2\sqrt{3}}\phi(\mathbf{r}, t)\right\}, \text{ for } \lambda \ll 1. \quad (111)$$

The above expression reduces for $\lambda = 0$ to the well-known diffusion coefficient for passive systems, where hydrodynamic interactions change the prefactor +8 in eqn (111) to +1.45. Contributions from hydrodynamic interactions are therefore important for passive systems. The contributions to Fick's diffusion coefficient due to hydrodynamic interactions originating from activity might be less important, as these interactions are short-ranged.

Acknowledgements

The authors acknowledge the various very helpful discussions with Prof. Vincent Démery. In particular, he provided them with the ideas concerning the scaling relation discussed in Section 5.2.

Notes and references

- 1 H. Stark, *Phys. J.*, 2007, **6**, 31.
- 2 W. F. Paxton, S. Sundararajan, T. E. Mallouk and A. Sen, *Angew. Chem., Int. Ed.*, 2006, **45**, 5420.
- 3 E. Lauga, *Soft Matter*, 2011, **7**, 3060.
- 4 E. Lauga and T. R. Powers, *Rep. Prog. Phys.*, 2009, **72**, 096601.
- 5 J. Elgeti, R. G. Winkler and G. Gompper, *Rep. Prog. Phys.*, 2015, **78**, 056601.
- 6 B. Bet, G. Boosten, M. Dijkstra and R. van Roij, *J. Chem. Phys.*, 2017, **146**, 084904.
- 7 W. C. K. Poon, From clarkia to escherichia and janus: The physics of natural and synthetic active colloids, in *Proceedings of the International School of Physics Enrico Fermi Course CLXXXIV Physics of Complex Colloids*, ed. C. Bechinger, F. Sciortino and P. Ziherl, IOS Press, 2013, p. 317.
- 8 A. Zöttl and H. Stark, *J. Phys.: Condens. Matter*, 2016, **28**, 253001.
- 9 J.-P. Hansen and I. R. McDonald, *Theory of simple liquids*, Academic Press, London, 1986.
- 10 J. K. G. Dhont, in *An introduction to dynamics of colloids* *Studies in Interface Science*, ed. D. Möbius and R. Miller, Elsevier, Amsterdam, 1996.
- 11 J. Clopes, G. Gompper and R. G. Winkler, *Soft Matter*, 2020, **16**, 10676.
- 12 T. Ishikawa, *J. R. Soc., Interface*, 2009, **6**, 815.
- 13 I. O. Götze and G. Gompper, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2010, **82**, 041921.
- 14 J. J. Molina, Y. Nakayama and R. Yamamoto, *Soft Matter*, 2013, **9**, 4923.
- 15 A. W. Zantop and H. Stark, *Soft Matter*, 2020, **16**, 6400.
- 16 T. J. Pedley, *IMA J. Appl. Math.*, 2016, **81**, 488.
- 17 B. Delmotte, E. Climent and F. Plouraboué, in *Franch hydrodynamic interactions among large populations of swimming micro-organisms*, *Computer Methods in Biomechanics and Biomedical Engineering*, 2013, vol. 16 (supp. 1), pp. 6–8, ISSN: 1025-5842.
- 18 A. Furukawa, D. Marenduzzo and M. E. Cates, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2014, **90**, 022303.
- 19 J. Stenhammar, C. Nardini, R. W. Nash, D. Marenduzzo and A. Morozov, *Phys. Rev. Lett.*, 2017, **119**, 028005.
- 20 J. Rodenburg, M. Dijkstra and R. van Roij, *Soft Matter*, 2017, **13**, 8957.
- 21 C. M. Kjeldbjerg and J. F. Brady, *Soft Matter*, 2021, **17**, 523.
- 22 M. Enculescu and H. Stark, *Phys. Rev. Lett.*, 2011, **107**, 058301.
- 23 A. M. Menzel, A. Saha, C. Hoell and H. Löwen, *J. Chem. Phys.*, 2016, **144**, 024115.
- 24 H. H. Wensink, H. Löwen, M. Marechal, A. Härtel, R. Wittkowski, U. Zimmermann, A. Kaiser and A. M. Menzel, *Eur. Phys. J.: Spec. Top.*, 2013, **222**, 3023.
- 25 T. Speck, J. Bialké, A. M. Menzel and H. Löwen, *Phys. Rev. Lett.*, 2014, **112**, 218304.
- 26 T. Speck, A. M. Menzel, J. Bialké and H. Löwen, *J. Chem. Phys.*, 2015, **142**, 224109.
- 27 T. Speck, *Soft Matter*, 2020, **16**, 2652.
- 28 J. Bialké, H. Löwen and T. Speck, *Eur. Phys. Lett.*, 2013, **103**, 30008.
- 29 R. van Damme, J. Rodenburg, R. van Roij and M. Dijkstra, *J. Chem. Phys.*, 2019, **150**, 164501.
- 30 S. Hermann, P. Krininger, D. de Las Heras and M. Schmidt, *Phys. Rev. E*, 2019, **100**, 052604.
- 31 S. Hermann, D. de Las Heras and M. Schmidt, *Phys. Rev. Lett.*, 2019, **123**, 268002.
- 32 M. C. Marchetti, Y. Fily, S. Henkes, A. Patch and D. Yllanes, *Curr. Opin. Colloid Interface Sci.*, 2016, **21**, 34.
- 33 J. Stenhammar, A. Tiribocchi, R. J. Allen, D. Marenduzzo and M. E. Cates, *Phys. Rev. Lett.*, 2013, **111**, 145702.
- 34 Y. Fily, S. Henkes and M. C. Marchetti, *Soft Matter*, 2014, **10**, 2132.
- 35 Y. Fily and M. C. Marchetti, *Phys. Rev. Lett.*, 2012, **108**, 235702.
- 36 J. Bickmann and R. Wittkowski, *J. Phys.: Condens. Matter*, 2020, **32**, 214001.
- 37 J. Bickmann, S. Bröker, J. Jeggle and R. Wittkowski, 2020, arXiv:2010.05262v1.
- 38 J. Jeggle, J. Stenhammar and R. Wittkowski, *J. Chem. Phys.*, 2020, **152**, 194903.
- 39 M. C. Marchetti, J. F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao and R. A. Simha, *Rev. Mod. Phys.*, 2013, **85**, 1143.
- 40 U. Marini Bettolo Marconi and C. Maggi, *Soft Matter*, 2015, **11**, 8768.
- 41 A. Poncet, O. Bénichou, V. Deméry and D. Nishiguchi, *Phys. Rev. E*, 2021, **103**, 012605.



42 A. Poncet, O. Bénichou, V. Démery and G. Oshanin, *Phys. Rev. Lett.*, 2017, **118**, 118002.

43 G. Gompper, *et al.*, *J. Phys.: Condens. Matter*, 2020, **32**, 193001.

44 M. R. Shaebani, A. Wysocki, R. G. Winkler, G. Gompper and H. Rieger, *Nat. Rev. Phys.*, 2020, **2**, 181.

45 H. Löwen, *J. Chem. Phys.*, 2020, **152**, 040901.

46 B. ten Hagen, S. van Teeffelen and H. Löwen, *J. Phys.: Condens. Matter*, 2011, **23**, 194119.

47 J. R. Howse, R. A. L. Jones, A. J. Ryan, T. Gough, R. Vafabakhsh and R. Golestanian, *Phys. Rev. Lett.*, 2007, **99**, 048102.

48 I. Buttinoni, J. Bialké, F. Kümmel, H. Löwen, C. Bechinger and T. Speck, *Phys. Rev. Lett.*, 2013, **110**, 238301.

49 M. T. Downton and H. Stark, *J. Phys.: Condens. Matter*, 2009, **21**, 204101.

50 J. K. G. Dhont, *J. Chem. Phys.*, 1996, **105**, 5112.

51 A. H. Nayfeh, *Introduction to Perturbation Techniques*, Wiley & Sons, New York, 1981.

52 W. Eckhaus, *Asymptotic Analysis of Singular Perturbations, Studies in Mathematics and its Applications*, North-Holland, Amsterdam, 1979.

53 F. Alarcón and I. Pagonabarraga, *J. Mol. Liq.*, 2013, **185**, 56.

54 F. Ginot, I. Theurkauff, D. Levis, C. Ybert, L. Bocquet, L. Berthier and C. Cottin-Bizonne, *Phys. Rev. X*, 2015, **5**, 218103.

55 T. Arnoux de Pirey, G. Lozano and F. van Wijland, *Phys. Rev. Lett.*, 2019, **123**, 260602.

56 I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, London, 2007, 4th edn.

57 Personal communication with Prof. Vincent Démery.

58 M. E. Cates and J. Tailleur, *Eur. Phys. Lett.*, 2013, **101**, 20010.

59 A. Baskaran and M. C. Marchetti, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2008, **77**, 011920.

