Countoscope for self-propelled particles

Tristan Cerdin [1,2] Talia Calazans [3] C. Douarche [1] and Sophie Marbach [2,*]
[1] *Université Paris-Saclay, CNRS, FAST, 91405 Orsay, France*
[2] *CNRS, Sorbonne Université, Physicochimie des Electrolytes et Nanosystèmes Interfaciaux, F-75005 Paris, France*
[3] *Department of Physics and Astronomy, University of Pennsylvania, Philadelphia 19104, USA*

(Received 31 March 2026; accepted 29 September 2026; published 9 November 2026)

Particle number fluctuations $N(t)$, measured in virtual observation boxes of an image or a simulation, offer a way to quantify particle dynamics when particle tracking is impractical, such as in high-density systems. While traditionally limited to equilibrium diffusive systems, we extend this approach—named "Countoscope"—to out-of-equilibrium self-propelled particles: Active Brownian (ABPs), run and tumble (RTPs), and active Ornstein-Uhlenbeck particles (AOUPs). For AOUPs, we leverage their Gaussian statistics to derive a general formula applicable to any Gaussian system. For ABPs and RTPs, we derive the self-intermediate scattering function (ISF)—and thus the correlations of $N(t)$—using an exact perturbative expansion over the probability density fields, revealing key physical features of the ISF and of the number correlations. Our theoretical predictions for the mean-squared number difference $\langle \Delta N^2(t) \rangle = \langle (N(t) - N(0))^2 \rangle$ match stochastic simulations and exhibit three time-dependent scaling regimes: diffusive, ballistic, and long-time enhanced diffusive, reflecting the regimes of the mean-squared particle displacement. We further uncover limiting laws in each of these regimes that are useful to quantify self-propulsion properties.

DOI: 10.1103/9rnw-g4xw

I. INTRODUCTION

More than a century ago, Smoluchowski introduced the concept of measuring particle dynamics by probing the number $N(t)$ of particles within a microscope's field of view [1–4]. As particles undergo Brownian motion, they diffuse in and out of the observation region, inducing fluctuations in $N(t)$. Smoluchowski developed a theory—later verified experimentally [5,6]—to relate diffusion coefficients to the correlations of $N(t)$ at equilibrium. While a limited number of studies expanded on this principle through the 1980s, most remained theoretical, with even fewer addressing out-of-equilibrium systems [7–11]. A notable 1953 study by Rothschild examined the dynamics of sperm cells [8], a canonical example of self-propelling particles that consume energy to perform directed motion [12]. Assuming sperm cells swim straight without reorientation, Rothschild derived a theory to quantify their swim speed from fluctuating counts. Rothschild also introduced the idea of counting particles in many virtual observation boxes of an image to improve statistical accuracy. However, due to the practical challenges of manual counting and correlation analysis at the time, the potential of number fluctuations was progressively forgotten [13].

Smoluchowski's modern legacy lies instead in techniques such as dynamic light scattering (DLS) and fluorescence correlation spectroscopy (FCS), which rely on correlations of the *intensity* scattered by particles within an illuminated region [14,15]. These methods are typically used to measure diffusion coefficients in equilibrium suspensions and are less commonly applied to out-of-equilibrium systems. Since self-propelled particles exhibit both directed motion and reorientation over time, their long-time behavior is akin to enhanced diffusion. DLS and FCS can quantify this enhanced diffusion [16,17], and ongoing research continues to refine their ability to distinguish self-propulsion speed and reorientation time from the signals [18].

Advances in modern microscopy and automated analysis have removed the obstacles that once made Smoluchowski's method impractical. Recent work by some of us has revived his approach to investigate the diffusive properties of colloidal suspensions at equilibrium [19]. Importantly, we introduced the "Countoscope" concept, which involves probing $N(t)$ in observation boxes of varying sizes to resolve dynamics across different length scales. This approach has enabled us to identify a peculiar hydrodynamic enhancement of *collective* diffusion in dense colloidal suspensions [20]. However, the potential of this method for analyzing out-of-equilibrium systems—particularly self-propelled particles—remains unexplored.

The primary challenge in applying Smoluchowski's formalism to self-propelled particles is the absence of a theoretical framework linking self-propelled dynamical parameters to fluctuating counts. Since $N(t)$ can be expressed as the integral of particle density within a box, a model for density correlations is sufficient to infer the correlation functions of $N(t)$. The intermediate scattering function (ISF), defined as the spatial Fourier transform of the density correlation, is a standard tool for investigating such correlations. Recent efforts have focused on solving the exact ISF for noninteracting self-propelled systems [21–24], achieving remarkable agreement with experimental data from Janus colloids and bacteria [22,23]. While some of these approaches are model specific [21] or rely on sophisticated theoretical frameworks requiring

*Contact author: sophie.marbach@cnrs.fr

 ©2026 American Physical Society

numerical solutions to obtain the ISF [22,24], they highlight the rich potential of ISFs to investigate self-propelled particles.

To investigate number correlations with ISFs, it is useful to manipulate simple expressions of the ISFs to increase computational efficiency as number correlations require additional integration steps. Recently, several theories have been developed to solve the probability distribution functions of self-propelled particles through the Fokker-Planck equation (FPE) of the system [25–28]. These approaches, while perturbative, offer simplified solutions and, therefore, hold significant potential for deriving analytical laws for number fluctuations. Such a perturbative strategy on the FPE has recently been successfully applied to obtain a simplified ISF of anisotropic active Brownian particles [27]. Another advantage is that these theories allow different models of self-propelled particles to be treated within the same framework [26,28], making them highly versatile.

Here we derive theoretical expressions for the number fluctuations of three canonical models of self-propelled noninteracting particles: Active Brownian particles (ABP), run-and-tumble particles (RTP), and active Ornstein-Uhlenbeck particles (AOUP). We use a perturbative expansion of their Fokker-Planck equations, conducted over increasing modes of orientational order. This expansion can also be interpreted as a perturbation in the Péclet number governing orientational relaxation. Our approach to obtain the Fourier-time ISF goes beyond prior works since we obtain all terms of the expansion, producing exact results in the form of an infinite series of fractions—a so-called *continued* fraction. The real-time ISF is obtained via inverse Fourier transform, with explicit expressions available at select perturbation levels. Number correlations are then derived as spatial integrals of the ISF. Comparison with numerical simulations confirms that our expressions are exact, with accuracy maintained for sufficiently small Péclet numbers when truncated at finite perturbation orders.

Our theoretical expansion also clarifies key features of the ISF and number correlations: Oscillations in both correlation functions arise from reorientation dynamics. Importantly, number correlations are very sensitive to differences in reorientational behavior between the different models. We identify three dynamical regimes in the number correlations: translational diffusion at short times, self-propulsion at intermediate times, and enhanced diffusion at long times. These limiting regimes reflect behavior of the mean-squared displacement of particles and are each more pronounced in small or large observation boxes depending on the regime. Finally, we obtain simple theoretical laws for number correlations in each of these regimes and discuss them in the practical context of extracting dynamical parameters from experimental data.

II. SETUP FOR COUNTING SELF-PROPELLED PARTICLES

A. Counting method

We study the number $N(t)$ of self-propelled particles in square boxes of size $L \times L$. Note that although we provide a reasoning here for two dimensions, the tools we use may readily be adapted to three dimensions [27]. The number $N(t)$ fluctuates as particles randomly enter and exit the box, undergoing various dynamics, see an example in Fig. 1. To read out the particles' dynamic properties from the fluctuating counts, we will here mostly explore two correlation functions: (i) the time correlation function $C_N(t) = \langle N(t)N(0)\rangle - \langle N\rangle^2$ for the particle number and (ii) the mean-squared change in particle number

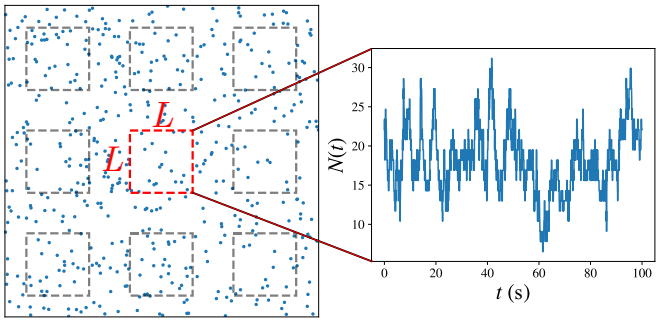


FIG. 1. Probing number fluctuations. "Countoscope" approach where we probe the number of particles in virtual observation boxes of an image or simulation.

$$\begin{aligned}\langle \Delta N^2(t)\rangle &= \langle (N(t) - N(0))^2\rangle \\ &= 2(\langle N^2\rangle - \langle N\rangle^2) - 2(\langle N(t)N(0)\rangle - \langle N\rangle^2) \\ &= 2(\langle N^2\rangle - \langle N\rangle^2) - 2C_N(t). \end{aligned} \tag{1}$$

In the above, $\langle\cdot\rangle$ corresponds to an average over realizations of the noise, initial conditions, and boxes of same size in an image. While we do not explicitly highlight this in our notations signatures, we note that $N(t)$ crucially depends on the size of the virtual box and probes different motion occurring at different length scales L.

$\langle \Delta N^2(t)\rangle$ can be put in parallel, for particle numbers, to the mean-squared displacement (MSD) $\langle \Delta \boldsymbol{r}^2(t)\rangle = \langle |\boldsymbol{r}(t) - \boldsymbol{r}(0)|^2\rangle$, where $\boldsymbol{r}(t)$ refers to a particle's position at time t. We will refer throughout this article to $\langle \Delta N^2(t)\rangle$ as the number mean-squared difference (NMSD). One main difference is that although for the MSD the averaging $\langle\cdot\rangle$ is done over individual particle trajectories, for the NMSD the averaging is instead taken over all boxes L "drawn" over the observation window. Note, that while they do not in Fig. 1, typically observation boxes overlap to allow for increased statistics, as demonstrated in a previous work [20]. Typically, here the number of boxes ranges between 10^1 and 10^3 for each box size, with fewer boxes for bigger box sizes.

In our investigation, we will consider for simplicity that particles do not interact. The presence or not of a particle in a box at fixed time t is therefore independent of that of others and, therefore, follows a Poisson process. Hence, we can write $\langle N^2\rangle - \langle N\rangle^2 = \langle N\rangle$ [29]. We can thus simplify Eq. (1) to $\langle \Delta N^2(t)\rangle = 2\langle N\rangle - 2C_N(t)$. It is thus sufficient to obtain a theoretical expression for either $C_N(t)$ or the NMSD to obtain predictions on both quantities. Here we will focus on $C_N(t) = \langle N(t)N(0)\rangle - \langle N\rangle^2$. We can write $N(t)$ in terms of individual particle contributions $N(t) = \sum_i n_i(t)$, where $n_i = 1$ when particle i is inside the box and 0 when outside

[30]. Then, for noninteracting particles,

$$\langle N(t)N(0)\rangle = \sum_i \langle n_i(t)n_i(0)\rangle + \sum_{i,j\neq i}\langle n_j(t)n_i(0)\rangle$$
$$= \langle N\rangle P_{\rm in}(t) + \langle N\rangle^2,$$

where $P_{\rm in}$ is the probability that a particle present in the box at $t=0$ is still in it at time t ($n_i(t)=n_i(0)=1$). $C_N(t)$ can thus be reexpressed as

$$C_N(t) = \langle N\rangle P_{\rm in}(t). \tag{2}$$

Note that this relationship was established quite some time ago by first principles by Smoluchowski [1,3]. The factor $P_{\rm in}(t)$ was then referred to as "probability after-effect" [4]. While its expression for strictly diffusive systems is known, it is much less intuitive for self-propelling particles. Overall, $\langle N(t)N(0)\rangle$ has an $\langle N\rangle^2$ contribution from different particles and an excess contribution from how long each of the $\langle N\rangle$ particles initially inside the box keeps contributing to correlations by staying in the box [30].

We can further express $P_{\rm in}(t)$ in terms of $a_0(\boldsymbol{r},\boldsymbol{r}',t)$, the probability of a particle to be at position $\boldsymbol{r}$ at time t with initial position $\boldsymbol{r}'$. We do so by integrating on $\boldsymbol{r}$ over the size of the box and averaging over all possible starting positions $\boldsymbol{r}'$ within the box. Rewriting a_0 as $a_0(\boldsymbol{r}-\boldsymbol{r}',t)$, since a_0 should only depend on the distance between start and end point due to translational invariance, we obtain

$$P_{\rm in}(t) = \frac{1}{L^2}\iint_{[-\frac{L}{2},\frac{L}{2}]^2} d\boldsymbol{r}d\boldsymbol{r}'\, a_0(\boldsymbol{r}-\boldsymbol{r}',t). \tag{3}$$

It will often be useful to operate in spatial Fourier space, and so we define $F(\boldsymbol{k},t)$, the Fourier transform of a_0 as

$$F(\boldsymbol{k},t) = \int d\boldsymbol{r} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} a_0(\boldsymbol{r},t) \tag{4}$$

for a given wave number $\boldsymbol{k}$. $F(\boldsymbol{k},t)$ is the self-intermediate scattering function of the particles, which we refer to as ISF for simplicity. When particles are noninteracting, and writing $\boldsymbol{r}'$ their initial position, $F(\boldsymbol{k},t)$ can be written [14] as

$$F(\boldsymbol{k},t) \equiv \langle e^{-i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')}\rangle = \int d(\boldsymbol{r}-\boldsymbol{r}') e^{-i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')} a_0(\boldsymbol{r}-\boldsymbol{r}',t), \tag{5}$$

where the average indicates an average over all particle trajectories. This last formulation shows that $F(\boldsymbol{k},t)$ is indeed the Fourier transform of a_0. We thus obtain

$$C_N(t) = \frac{\langle N\rangle}{L^2}\iint_{[-\frac{L}{2},\frac{L}{2}]^2} d\boldsymbol{r}d\boldsymbol{r}' \int \frac{d\boldsymbol{k}}{(2\pi)^2} e^{i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')} F(\boldsymbol{k},t). \tag{6}$$

Note that this expression can be simplified by carrying out the integrals on $\boldsymbol{k}$—see Eq. (A1). In this paper, we will establish expressions for either directly $P_{\rm in}(t)$ or $F(\boldsymbol{k},t)$ to obtain expressions for the correlation function $C_N(t)$, through Eq. (2) or Eq. (6).

B. Three models for self-propelled particles

We consider three broadspread models of self-propelled particles to benchmark the "Countoscope" approach in this nonequilibrium context. First, we consider RTP, which describe well a range of biological systems, including many bacteria and swimming microalgae [23,31–33]. In its simplest form, in two dimensions (2D), the position $\boldsymbol{r}(t)$ of a particle whose orientation is given by $\boldsymbol{u}(\theta) = (\cos\theta, \sin\theta)$ evolves as

$$\dot{\boldsymbol{r}}(t) = v\boldsymbol{u}(\theta_i) + \sqrt{2D_t}\boldsymbol{\eta}(t), \tag{7}$$

$$\theta \underset{\alpha}{\rightarrow} \theta' \in [0, 2\pi], \tag{8}$$

where v is the self-propulsion speed, D_t a translational diffusion coefficient, and $\boldsymbol{\eta}(t)$ a Gaussian white noise satisfying $\langle\boldsymbol{\eta}(t)\rangle = 0$ and $\langle\eta_i(t)\eta_j(t')\rangle = \delta_{ij}\delta(t-t')$, where $\langle\cdot\rangle$ represents an average over realizations of the noise. Equation (8) corresponds to exponentially distributed tumbling events at rate α. Note that some models include two rates to distinguish tumbling and running phases, but in the vast majority of cases, one can take the approximation of instantaneous tumbles [31]. Second, we consider the ABP model, which describes better systems where reorientation is smooth and diffusive. In that case the evolution equation for $\boldsymbol{r}(t)$ is the same as for the RTP Eq. (7), but the orientation evolves as

$$\dot{\theta}(t) = \sqrt{2D_r}\eta_r(t), \tag{9}$$

where $\eta_r(t)$ is another Gaussian white noise and D_r is the angular diffusion coefficient. ABP models capture well the motion of Janus particles [34–37]. For both RTP and ABP, fluctuations come from passive diffusivity and reorientation. It can, however, be useful to allow the self-propulsion speed to fluctuate for certain applications, which is captured by the AOUP model. In this case,

$$\dot{\boldsymbol{r}} = \boldsymbol{v}_{\rm OU}(t) + \sqrt{2D_t}\boldsymbol{\eta}(t), \tag{10}$$

$$\tau_r\dot{\boldsymbol{v}}_{\rm OU} = -\boldsymbol{v}_{\rm OU}(t) + \sqrt{2D_v}\boldsymbol{\eta}_v(t), \tag{11}$$

where τ_r is a timescale associated with velocity relaxation, D_v a velocity diffusion coefficient, and $\boldsymbol{\eta}_v(t)$ another Gaussian random noise. In this model the mean particle velocity is $v \equiv \langle\sqrt{|\boldsymbol{v}_{\rm OU}(t)|^2}\rangle = \sqrt{2D_v/\tau_r}$, allowing us to build correspondence between the different models.

For all three models, the MSD can be captured by a single equation

$$\langle\Delta\boldsymbol{r}^2(t)\rangle = 4\left(D_t + \frac{v^2}{2D_r}\right)t + 2\frac{v^2}{D_r^2}(e^{-D_r t} - 1), \tag{12}$$

where $D_r = \alpha$ for RTPs, and $D_r = 1/\tau_r$ and $v = \sqrt{2D_v/\tau_r}$ for AOUPs. This equation captures the three regimes of motion displayed by the particles: At very short times, motion is diffusive $\langle\Delta r^2(t)\rangle = 4D_t t$ until a crossover time $t_{\rm ball} = 4D_t/v^2$, after which motion is ballistic $\langle\Delta r^2(t)\rangle = v^2t^2$. Reorientation dynamics make motion diffusive again after another crossover time $t_{\rm diff} = 4D_{\rm eff}/v^2$ and $\langle\Delta r^2(t)\rangle = 4D_{\rm eff}t$ with an effective diffusion coefficient $D_{\rm eff} = D_t + v^2/2D_r$ that is enhanced by activity. We expect these three regimes to also appear in the NMSD due to the similarities in how they probe particle dynamics in time.

We quantify the impact of activity on the motion of active particles with a dimensionless Péclet number Pe,

$$\mathrm{Pe} = \frac{v\ell}{D_t} = \frac{v}{2\sqrt{D_r^\alpha D_t}}, \tag{13}$$

where $D_r^\alpha = D_r$ for ABP, α for RTP, and $1/\tau_r$ for AOUPS. For convenience, we use the equivalent notation $D_r^\alpha = D_r + \alpha$. The Pe number compares the timescale of active self-propulsion to that of diffusion over a length scale ℓ. We choose ℓ to characterize the distance that a particle diffuses before significant reorientation happens, so such that $4\ell^2/D_t = 1/D_r^\alpha$, giving $\ell = \sqrt{D_t/D_r^\alpha}/2$. The synthetic Janus particles typically described by the ABP model tend to have a low Pe $\simeq 3-6$ [34,36,38]. Bacteria such as *Escherichia coli* usually have higher self-propulsion speed with Pe $\simeq 15-20$ [24,31].

All theoretical calculations are checked against numerical simulations of the three particle models. For more details on numerical implementation see Appendix A.

C. Calculating ISFs for self-propelled particles

1. Non-Gaussian models

For both the noninteracting ABP and RTP models—which are non-Gaussian models—exact results of ISFs have been derived in recent years [21–23]. However, these solutions are either model specific or computationally expensive. Given that investigating number fluctuations requires to perform an integral on $\boldsymbol{k}$ on top of the ISF, as given in Eq. (6), we attempt to find a numerically efficient expression of the ISF. In addition, given the complexity of the previously obtained expressions, it is hard to infer behavior in limiting regimes of parameters.

We note that one common alternative approach to obtain ISFs involves stochastic density field theories (sDFT) or Dean-Kawasaki equations [19,39,40,40–44]. For this application, stochastic density field theory is impractical as there is no consensus yet on a stochastic density field theory for self-propelled particles [45–48]. Rather, we will use approaches based on the FPE—for which a consensus exists on the expression of the FPE for noninteracting suspensions of active particles. Thus we obtain results directly on the probability distribution $a_0(\boldsymbol{r}, t)$ rather than on the microscopic density $\sum_i \delta(\boldsymbol{r} - \boldsymbol{r}_i(t))$ studied by sDFT.

2. Gaussian models

The case of AOUPs is slightly different as it is a Gaussian model. Indeed, the self-propulsion speed $\boldsymbol{v}_{\mathrm{OU}}(t)$ in AOUPs evolves smoothly as a Gaussian colored noise. Thus, the probability distribution of the position of noninteractive AOUPs is Gaussian. It therefore satisfies the common property that all of the cumulants of its position variables above 2 are equal to 0. Looking back on Eq. (5) with initial condition $\boldsymbol{r}(0)$ we can show through a cumulant expansion [49] that

$$\begin{aligned} F(\boldsymbol{k}, t) &= \langle e^{-i\boldsymbol{k}\cdot(\boldsymbol{r}(t)-\boldsymbol{r}(0))}\rangle = e^{-i\boldsymbol{k}\cdot\langle \boldsymbol{r}-\boldsymbol{r}(0))\rangle - \frac{k^2}{4}\langle|\boldsymbol{r}(t)-\boldsymbol{r}(0)|^2\rangle + \ldots} \\ &= e^{-\frac{k^2}{4}\langle\Delta r^2(t)\rangle}, \end{aligned} \tag{14}$$

obtaining a well-known result for AOUPs using Eq. (12).

This result can further be plugged in Eq. (6). Then, one can split up integrals on x and y and, realizing they are the same, obtain

$$\begin{aligned} C_N(t) &= \langle N\rangle \left[f\left(\frac{\langle \Delta r^2(t)\rangle}{L^2}\right)\right]^2 \\ \text{with } f(\tau) &= \left[\sqrt{\frac{\tau}{\pi}}(e^{-1/\tau} - 1) + \mathrm{erf}\left(\sqrt{\frac{1}{\tau}}\right)\right], \end{aligned} \tag{15}$$

where erf is the error function. For more details on this calculation we refer the reader to Ref. [19].

It then follows that Countoscope formulas for *any* Gaussian model can easily be obtained so long as one knows the MSD of the model and simply by plugging the MSD into the function f. Remarkably the intuition for this formula was brought forwards recently in Ref. [50] but not derived. Note that in Ref. [50] it was applied to both Gaussian (fractional Brownian motion) and non-Gaussian random walks (obstructed diffusion) with only a mild deviation observed for non-Gaussian cases. Equation (15) will be useful as a reference point for the study of ABPs and RTPs to see which role non-Gaussianity plays in number fluctuations.

III. ANGULAR MOMENTS OF THE PROBABILITY DENSITY FOR NON-GAUSSIAN MODELS

A. Derivation of the probability densities

1. Formulation in terms of an infinite hierarchy

To obtain expressions for the ISFs of our models we will rely on a coarse-graining procedure going from the stochastic equations of motions to probability space. Using a FPE formalism we study the time evolution of $P(\boldsymbol{r}, \theta, t)$, the probability density that a particle is found at position $\boldsymbol{r}$ with orientation θ at time t. For either an ABP or an RTP,

$$\begin{aligned} \partial_t P(\boldsymbol{r}, \theta, t) = &-\nabla_r \cdot (v\boldsymbol{u}(\theta)P - D_t \nabla_r P) + D_r \partial_\theta^2 P \\ &- \alpha P + \frac{\alpha}{2\pi}\int_0^{2\pi} d\theta'\, P(\boldsymbol{r}, \theta', t), \end{aligned} \tag{16}$$

where for an ABP one sets $\alpha = 0$ in the above expression, and respectively for an RTP one sets $D_r = 0$. To find $a_0(\boldsymbol{r}, t)$, it is sufficient to find the marginal of $P(\boldsymbol{r}, \theta, t)$, i.e., $P(\boldsymbol{r}, \theta, t)$ integrated over all values of θ. To get rid of θ in this FPE, we closely follow the method of Dinelli *et al.* [26] for its simplicity when working with 2D active particles. Note, however, that we use a different closure for our purpose. The angular dependences are separated from the radial ones via a Fourier series

$$P(\boldsymbol{r}, \theta, t) = \frac{1}{2\pi}\sum_{n=-\infty}^{+\infty} e^{in\theta} a_n(\boldsymbol{r}, t), \tag{17}$$

introducing the different coefficients $a_n(\boldsymbol{r}, t)$ for $n \in \mathbf{Z}$. Introducing the scalar product $\langle f, g\rangle = \int_0^{2\pi} d\theta\, f^*(\theta) g(\theta)$, we note that each component may be obtained as an integral over the mode n, as

$$\langle e^{in\theta}, P\rangle = \int_0^{2\pi} d\theta\, e^{-in\theta}\frac{1}{2\pi}\sum_{k=-\infty}^{+\infty} e^{ik\theta} a_k(\boldsymbol{r}, t) = a_n.$$

We now aim to reexpress the FPE Eq. (16) as a function of the $a_n(\boldsymbol{r},t)$ so as to obtain $a_0(\boldsymbol{r},t)$. To do so, a few more relationships are worth deriving:

$$\begin{aligned}\langle e^{in\theta}, \partial_\theta^2 P\rangle &= \int_0^{2\pi} d\theta\; e^{-in\theta}\frac{1}{2\pi}\sum_{k=-\infty}^{+\infty} -k^2 e^{ik\theta}a_k(\boldsymbol{r},t)\\ &= -n^2 a_n\\ \langle e^{in\theta}, \boldsymbol{u}(\theta)P\rangle &= \int_0^{2\pi} d\theta\; e^{-in\theta}\frac{1}{2\pi}\sum_{k=-\infty}^{+\infty}\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}e^{ik\theta}a_k(\boldsymbol{r},t)\\ &= \frac{1}{2}[\hat{i}_+ a_{n+1} + \hat{i}_- a_{n-1}],\end{aligned}$$

where we also introduced the vectors $\hat{i}_\pm = \binom{1}{\pm i}$. In addition we note that

$$\begin{aligned}\int_0^{2\pi} d\theta\, P(\boldsymbol{r},\theta,t) &= \int_0^{2\pi} d\theta\,\frac{1}{2\pi}\sum_{n=-\infty}^{+\infty} e^{in\theta}a_n(\boldsymbol{r},t)\\ &= \sum_{n=-\infty}^{+\infty} a_n\int_0^{2\pi} d\theta\frac{e^{in\theta}}{2\pi}\\ &= \sum_{n=-\infty}^{+\infty} a_n\,\delta_{n,0} = a_0.\end{aligned}$$

We can then straightforwardly compute $\langle e^{in\theta}, \partial_t P\rangle$ and obtain a series of equations for the evolution of all the a_n coefficients,

$$\begin{aligned}\partial_t a_n = &-\frac{v}{2}\nabla_r\cdot[\hat{i}_+ a_{n+1} + \hat{i}_- a_{n-1}] + D_t\Delta a_n\\ &- D_r n^2 a_n - \alpha(1-\delta_{n,0})a_n.\end{aligned} \tag{18}$$

In the following it will be useful to compare different levels of truncation of these series and so we explicitly write out the first three equations of this series. Note that we also add a term as $\delta(t)\delta(\boldsymbol{r})$ in a_0 to account for initial conditions starting at $t=0$ and $\boldsymbol{r}=\boldsymbol{0}$:

$$\partial_t a_0(\boldsymbol{r},t) = -\frac{v}{2}\nabla_r\cdot[\hat{i}_+ a_1 + \hat{i}_- a_{-1}] + D_t\Delta a_0 + \delta(t)\delta(\boldsymbol{r}), \tag{19}$$

$$\begin{aligned}\partial_t a_{\pm1}(\boldsymbol{r},t) = &-\frac{v}{2}\nabla_r\cdot[\hat{i}_\pm a_{\pm2} + \hat{i}_\mp a_0] + D_t\Delta a_{\pm1}\\ &-(D_r+\alpha)a_{\pm1},\end{aligned} \tag{20}$$

$$\begin{aligned}\partial_t a_{\pm2}(\boldsymbol{r},t) = &-\frac{v}{2}\nabla_r\cdot[\hat{i}_\pm a_{\pm3} + \hat{i}_\mp a_{\pm1}] + D_t\Delta a_{\pm2}\\ &-(4D_r+\alpha)a_{\pm2}.\end{aligned} \tag{21}$$

The approach on the infinite series of a_n can be related to other techniques consisting in obtaining evolution equations for the fields describing the density ρ, the polar $\boldsymbol{p}(\boldsymbol{r},t)$, and nematic orders of the system $\underline{\underline{q}}(\boldsymbol{r},t)$ [25,27,51–54]. The advantage of working at the level of the a_n is that the n coefficients are directly related to one another through a recurrence relation while the link with higher orders when working with these fields is not obvious. Although the physical meaning of the a_n coefficients is less straightforward than the meaning of these fields, it is possible to connect these approaches, via the relations

$$\begin{aligned}\rho(\boldsymbol{r},t) &= \int P(\boldsymbol{r},\boldsymbol{u},t)\,d\boldsymbol{u} = a_0,\\ \boldsymbol{p}(\boldsymbol{r},t) &= \int \boldsymbol{u}P(\boldsymbol{r},\boldsymbol{u},t)\,d\boldsymbol{u} = \begin{pmatrix}a_1+a_{-1}\\ i(a_1-a_{-1})\end{pmatrix},\\ \underline{\underline{q}}(\boldsymbol{r},t) &= \int\left(\boldsymbol{u}\boldsymbol{u} - \frac{\mathbb{I}}{2}\right)P(\boldsymbol{r},\boldsymbol{u},t)\,d\boldsymbol{u}\\ &= \begin{pmatrix}\frac{1}{2}(a_2+a_{-2}) & i(a_2+a_{-2})\\ i(a_2+a_{-2}) & -\frac{1}{2}(a_2+a_{-2})\end{pmatrix},\end{aligned}$$

shown in Appendix B. These relations show in particular that $a_0(\boldsymbol{r},t)$ represents the probability density of finding particles in $\boldsymbol{r}$, $\rho(\boldsymbol{r},t)$.

We note that although their notation in the literature are very similar, the density $\rho(\boldsymbol{r},t)$ is a different object than the microscopic density which characterizes the ensemble of individual particle positions. One is a probability density that is a macroscopic, smooth object corresponding to a distribution of positions while the other is a sum of delta peaks corresponding to a microscopic function that needs to be handled with more care. We will stick with the a_0 notation instead of $\rho(\boldsymbol{r},t)$ as a reminder that it is a probability obtained through an expansion.

To make progress, it will be useful in the following to use the time and space Fourier transform,

$$\widetilde{a}_n(\boldsymbol{k},\omega) = \iint \mathrm{d}\boldsymbol{r}\,\mathrm{d}t\; a_n(\boldsymbol{r},t)e^{-i\omega t}e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}.$$

2. *Closure procedure and infinite continued fraction*

Equation (18) describes an infinite hierarchy of equations, with each a_n coefficient depending on $a_{n\pm1}$. A truncation will therefore be necessary to obtain expressions for a_0. A standard procedure [25–28,51,52] consists in cutting higher orders, $n>N$, corresponding to either the polar, the nematic, or higher angular dependences. Truncating at order N implies taking all components $\{a_{\pm n}\}_{n\geqslant N+1}=0$. With a truncation at order N, using Eq. (18), the Fourier components $\widetilde{a}_{\pm N}$ can then be expressed as

$$\begin{aligned}i\omega\widetilde{a}_{\pm N}(\boldsymbol{k},\omega) = &-\frac{v}{2}i\boldsymbol{k}\cdot\hat{i}_\mp\widetilde{a}_{\pm(N-1)} - D_t k^2\widetilde{a}_{\pm N}\\ &- N^2 D_r\widetilde{a}_{\pm N} - \alpha\widetilde{a}_{\pm N},\end{aligned}$$

This expression may be recast into

$$\widetilde{a}_{\pm N}(\boldsymbol{k},\omega) = \frac{-\frac{v}{2}i\boldsymbol{k}\cdot\hat{i}_\mp\widetilde{a}_{\pm(N-1)}}{b_N},$$

introducing the notation

$$b_n = i\omega + D_t k^2 + n^2 D_r + \alpha(1-\delta_{n,0}),\ \ \forall n\in\mathbf{N}. \tag{22}$$

Terms of order $N-1$ are then found iteratively

$$\begin{aligned}i\omega\widetilde{a}_{\pm(N-1)} = &-\frac{v}{2}i\boldsymbol{k}\cdot\hat{i}_\mp\widetilde{a}_{\pm(N-2)}\\ &-\left(\frac{\frac{v^2k^2}{4}}{b_N} + D_t k^2 + (N-1)^2 D_r + \alpha\right)\widetilde{a}_{\pm(N-1)},\end{aligned} \tag{23}$$

which can be recast into

$$\widetilde{a}_{\pm(N-1)} = \frac{-\frac{v}{2} i\boldsymbol{k}\cdot\hat{i}_{\mp}\widetilde{a}_{\pm(N-2)}}{b_{N-1}+\frac{c_N}{b_N}},$$

introducing the notation

$$c_n = \frac{v^2k^2}{4},\ \forall n\in\mathbf{N}. \tag{24}$$

Continuing this scheme, it is clear we can express all $\widetilde{a}_{n\leqslant N}$ as a function of c_n and b_n, with $c_n \equiv c$ always constant. We thus introduce the notation for continued fraction

$$\underset{n=1}{\overset{N}{\mathrm{K}}}\frac{c_n}{b_n} = \cfrac{c_1}{b_1+\cfrac{c_2}{b_2+\cfrac{c_3}{b_3+\ddots}}}$$

such that for $\widetilde{a}_{\pm1}$,

$$\widetilde{a}_{\pm1} = \frac{-\frac{v}{2} i\boldsymbol{k}\cdot\widetilde{a}_0}{b_1+\mathrm{K}_{n=2}^{N}\frac{c_n}{b_n}}.$$

Finally an expression for $\widetilde{a}_0$ correct to order N in the truncation is

$$\widetilde{a}_0^{-1} = b_0 + 2\underset{n=1}{\overset{N}{\mathrm{K}}}\frac{c_n}{b_n}. \tag{25}$$

This compact expression is one of the main results of our work. Notice that when we go to the limit of a passive tracer, with $v=0$, we recover $\widetilde{a}_0(\boldsymbol{k},\omega) = 1/(i\omega + D_t k^2)$ which is the standard expression for a purely diffusive particle [55].

In the case of RTPs, Eq. (25) can be further simplified. Since neither the denominators nor numerator depend on n in this case, as given by Eqs. (22) and (24), $b_n \equiv b$ and $c_n \equiv c$, the continued fraction converges to an analytic expression when taking the limit of $N\to\infty$, which is

$$\widetilde{a}_0^{-1} = b_0 + \underset{n=1}{\overset{\infty}{\mathrm{K}}}\frac{c}{b} = b_0 + \frac{4c}{b+\sqrt{b^2+4c}}.$$

Using the expressions for b and c as provided in Eqs. (22) and (24) yields

$$\widetilde{a}_0(\boldsymbol{k},\omega) = \frac{1}{\sqrt{(i\omega + D_t k^2+\alpha)^2 + v^2k^2} - \alpha}, \tag{26}$$

which is another main result of our work. This expression is analytically exact as all the terms are accounted for and the hierarchy can be closed without the need for a truncation (since we took $N\to\infty$). This is, to our knowledge the first exact result obtained for the probability distribution for noninteracting RTPs in 2D from a perturbative series on the orientational orders. Our result in Eq. (26) is similar to Eq. (20) of Ref. [21]. In comparison, Ref. [21] does not contain translational diffusion, which can actually be added in the entire derivation by replacing $i\omega \to i\omega + D_t k^2$. Reference [21] relies on a perturbative series on the number of tumbling events [56] rather than on the distribution of orientational orders. As such, it is not easily transferable to other choices of reorientation dynamics.

For ABPs, one has no choice but to choose an order at which to close the series. Taking a truncation at order N essentially corresponds to choosing the value for which $\frac{c}{b_{N+1}}$ will be negligible compared to b_N. Therefore, truncation will be more accurate when $c \ll b_n^2$. A competition between the values of the speed v in c and of the diffusion coefficients in b_n at order k^2 can therefore be identified, and $\frac{c}{b_n^2}$ will be smaller at smaller Pe. The truncation will thus also converge quicker for ABPs, due to the factor n^2 appearing in front of D_r in the definition of b_n in Eq. (22). This contribution comes from the diffusive reorientation in space for ABPs and is absent in RTPs for which reorientation dynamics are sudden. As such, the truncation consisting in neglecting higher orientational moments will be more accurate for systems where the orientation avoids sudden changes and evolves smoothly. For ABPs we can for instance stop at polar order, taking $N=1$ such that $\{a_n\}_{n\geqslant2}=0$, and obtain

$$\widetilde{a}_{0,1}(\boldsymbol{k},\omega) = \frac{i\omega + D_t k^2 + D_r + \alpha}{(i\omega + D_t k^2)(i\omega + D_t k^2 + D_r + \alpha) + \frac{v^2k^2}{2}}, \tag{27}$$

where we used the notation $\widetilde{a}_{0,N}$ to designate $\widetilde{a}_0$ approximated with a truncation at order N. Stopping at nematic order, taking $N=2$ and $\{a_n\}_{n\geqslant3}=0$, yields

$$\widetilde{a}_{0,2}(\boldsymbol{k},\omega) = \frac{2(i\omega + D_t k^2 + 4D_r + \alpha)(i\omega + D_t k^2 + D_r + \alpha) + \frac{v^2k^2}{2}}{(i\omega + D_t k^2)\left[2(i\omega + D_t k^2 + 4D_r + \alpha)(i\omega + D_t k^2 + D_r + \alpha) + \frac{v^2k^2}{2}\right] + v^2k^2(i\omega + D_t k^2 + 4D_r + \alpha)}. \tag{28}$$

B. Intermediate scattering functions

1. *From frequency domain to real time*

Using Eq. (4), we can relate $\tilde{a}_0(\boldsymbol{k},\omega)$ to the ISF doing an inverse Fourier transform in time $F(\boldsymbol{k},t) = \int \frac{\mathrm{d}\omega}{2\pi} e^{i\omega t}\tilde{a}_0(\boldsymbol{k},\omega)$. At order $N=1$ in the expansion, or polar order, we can simply take the inverse Fourier transform of Eq. (27). This integral can be solved using the residue theorem, with two poles canceling the denominator,

$$\omega_{\pm} = i(D_t k^2 + D_r) \pm \sqrt{\frac{k^2v^2}{2} - \frac{D_r^2}{4}},$$

which yields

$$F_1(\boldsymbol{k},t) = e^{-(D_t k^2 + \frac{D_r^\alpha}{2})t}\left[\cos\left(\frac{t}{2}\sqrt{2k^2v^2 - \left(D_r^\alpha\right)^2}\right) + \frac{D_r^\alpha}{2\sqrt{2k^2v^2 - \left(D_r^\alpha\right)^2}}\sin\left(\frac{t}{2}\sqrt{2k^2v^2 - \left(D_r^\alpha\right)^2}\right)\right], \tag{29}$$

where we abbreviated $D_r^\alpha = D_r + \alpha$. The fact that we can factor out D_r and α in the same way means that at this order of the truncation, D_r and α play the same role. This is also expected from looking at $a_{\pm1}$ terms in Eq. (19) where D_r and α play the same role. The polar order truncation thus smears out differences between reorientation modes.

The same strategy can be used to obtain an ISF at nematic order. The denominator is of order 3 in this case and we find three poles $(\omega_1, \omega_2, \omega_3)$ in the superior part of the complex plane (see Appendix C for details). Writing $h(\omega) = (i\omega + D_t k^2 + D_r + \alpha)(i\omega + D_t k^2 + 4D_r + \alpha) + \frac{v^2k^2}{4}$, we get

$$F_2(\boldsymbol{k},t) = -\sum_{i=1}^{3} \frac{h(\omega_i)e^{i\omega_i t}}{\prod_{j\neq i}(\omega_i - \omega_j)}. \tag{30}$$

For higher orders, it gets increasingly harder to obtain analytical results. Equations (27) and (28) highlight that the denominator of $\widetilde{a}_0(\boldsymbol{k},\omega)$ when the truncation is stopped at order N will be of order $N+1$ in ω. Therefore, analytic solutions for $F(\boldsymbol{k},t)$ can only be found up to order $N=3$ in the truncation as only polynomials of up to order 4 admit general solutions for their roots. For higher orders, one thus has to invert the Fourier transform numerically, with fast Fourier transform (FFT) algorithms. We do not solve at order 3 analytically in this paper for the sake of brevity, and all the figures with order $N \geqslant 3$ are obtained from perfoming that numerical integration.

In the case of RTP, where we have an exact expression for $\widetilde{a}_0$ given via Eq. (26), we also have to perform a numerical inverse Fourier transform. Numerically obtaining $F(\boldsymbol{k},t)$ from $\widetilde{a}_0$ with FFTs is a fast and robust process, compared to alternatives which rely on numerically solving a system of equations [24]. Inspired by Ref. [21], we can reexpress Eq. (26) as

$$\widetilde{a}_0 = \frac{P_0}{1-\alpha P_0} = \sum_{n=0}^{\infty} \alpha^n P_0^{n+1},$$

with $P_0 = \sqrt{(i\omega + D_t k^2 + \alpha)^2 + v^2k^2}^{-1}$. The inverse Fourier transform can then be carried out within the sum and we find a rather straightforward expression, in the case of RTP only, as

$$F(k,t) = e^{-\alpha t - D_t k^2 t} \sum_{n=0}^{\infty} \frac{\sqrt{\pi}}{2^{\frac{n}{2}}\Gamma\left(\frac{n+1}{2}\right)} \left(\frac{\alpha^2 t}{kv}\right)^{\frac{n}{2}} J_{\frac{n}{2}}(kvt)$$

where the $J_{n/2}$ are Bessel functions of the first kind. In practice, this direct expression in real time could be computationally advantageous. However, it requires to conduct the sum at high values of n where commonly used algorithms to calculate Bessel functions are inaccurate [57].

2. *Convergence of the obtained expressions*

We now compare our analytical predictions for $F(\boldsymbol{k},t)$ with measured ISFs from simulated data, both for ABP and RTP (Fig. 2). We especially compare the results at different orders N. Since our particles do not interact, we are looking at the ISF which satisfies $F(\boldsymbol{k},0) = 1$ and decays towards zero in time. For both RTP and ABP, we find this decay is accompanied with significant oscillations—a fact that has been identified though seldom explained in several works prior [24,34].

To understand the emergence of these oscillations, a key lies in the convergence of our predictions at increasing truncation order N. Indeed, the main improvement between simulation data and theory at higher orders is in the description of the oscillations. At order 1 the theory overestimates their amplitude and they decay slower than they should, as higher angular harmonics are neglected and a particle's rotation is not accurately accounted for. Order 3 already describes ABP motion quite accurately—at least for the set of investigated parameters—while there are still some minor differences for RTPs. This is an expected consequence as we know that the continued fraction converges slower in the case of RTPs. Eventually, the exact result for RTPs is fully converged and describes the motion perfectly at all length and timescales, with minor differences seen at long times and small k due to numerical noise. This convergence of how well we describe oscillations hints to the way that these oscillations are linked to reorientational dynamics of the considered particles.

Overall, oscillations do not appear for small k and are damped at long times when the dynamics can be described effectively as a diffusive process. For the AOUP model [Fig. 3(a)], the ISF is a simple decreasing exponential exhibiting no oscillations. The oscillations of the ABP and RTP models are thus, here, a marker of non-Gaussianity, as the ISF of every Gaussian model has to be a decreasing exponential of the MSD [Eq. (14)]. Similar oscillations can be observed in scattering measurements of particles moving with a mean drift such as sedimenting particles [58] and are also a feature of ISFs found from the propagator of free running particles [59]. The oscillations thus likely come from persistent motion in one direction, with fixed velocity, unlike for AOUPs for which the magnitude of the velocity is Gaussian distributed.

Further physical interpretation of oscillatory features is hardly possible. Comparing the ISFs of ABPs and RTPs [Fig. 3(b)] shows that the difference between models manifests as subtle differences in the oscillatory behavior. Notably oscillations decrease slightly slower and have a slightly larger amplitude in the case of ABPs, but it is not obvious to link the diffusive reorientation of ABPs to this observation. Similarly, other studies have shown that for some exponents of active Lévy particles, oscillations in the ISF are more important than

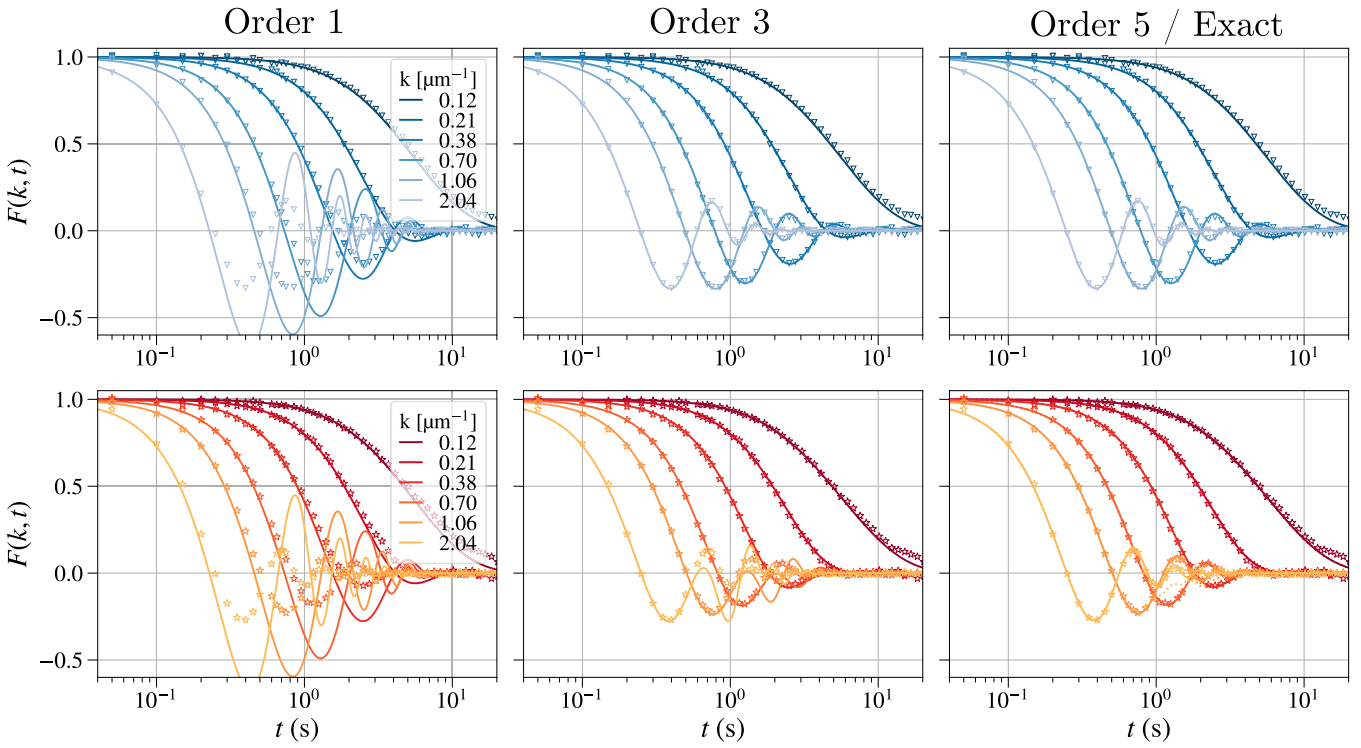


FIG. 2. Intermediate scattering functions for active Brownian particles (top) and run-and-tumble particles (bottom) obtained at different orders in the truncation. k values are taken to be the same for all three plots and go from small length scales (in light colors) to large ones (in dark colors). Dynamic parameters are $v = 5\ \mu\mathrm{m\,s^{-1}}$, $D_t = 0.1\ \mu\mathrm{m^2\,s^{-1}}$, and $D_r = 1\ \mathrm{s^{-1}}$. For order 5/Exact, in the top ABP graph the full line corresponds to an order $N = 5$ truncation, while in the bottom RTP plot, the full line is the exact result from the inverse FFT of Eq. (26) and the dotted line from an order $N = 5$ truncation.

for RTPs, but little physical explanation is found beyond these phenomenological observations [59]. ISFs thus relax differently according to the different reorientation mechanisms, but the direct link between oscillations and physical behavior remains unclear. This further justifies the investigation of other statistical observables that could be more directly linked to physical interpretation.

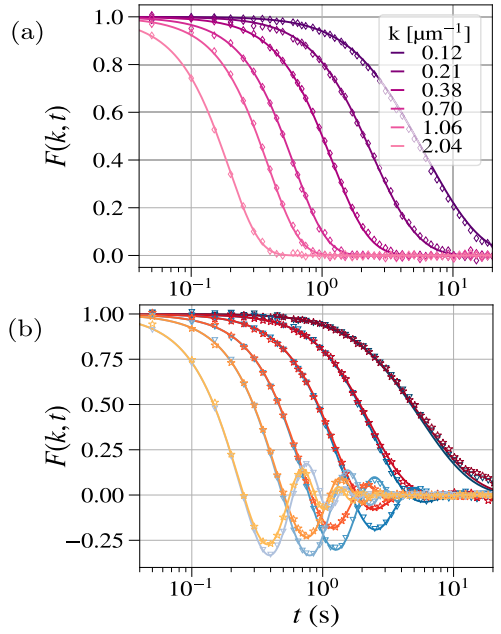


FIG. 3. Intermediate scattering function of AOUPs (a) and of ABPs and RTPS overlapped (b). The dynamic parameters used are equivalent to the ones in Fig. 2.

IV. COUNTOSCOPE FOR SELF-PROPELLED PARTICLES

A. Number fluctuations

To obtain the number correlation function $C_N(t)$ and the NMSD $\langle (N(t) - N(0))^2 \rangle$, we can use our predictions for $F(\boldsymbol{k}, t)$ and simply input them in Eq. (6), meaning taking a weighted integral over $\boldsymbol{k}$. Apart from the AOUPs, for which we have seen an explicit analytical solution exists to this integral [see Eq. (15)], for ABPs and RTPs, one has to perform a numerical integration. Using the numerical scheme detailed in Appendix A, this can be done quickly at little computational expense.

As a canonical illustration of the behavior of the number fluctuations, we present in Fig. 4(a), the NMSD curves for RTPs using the analytical expressions at order $N = 2$, for different box sizes (lines). We compare these with results from numerical simulations (stars) and find excellent agreement between the two. The NMSD exhibits several peculiar features. First, the number fluctuations increase in time, before reaching a plateau. Starting from an initial number configuration, the longer one observes particles inside a box the more likely one is to sample number configurations further away from the initial configuration; hence the number fluctuations increase. The plateau corresponds to the time when the memory of the

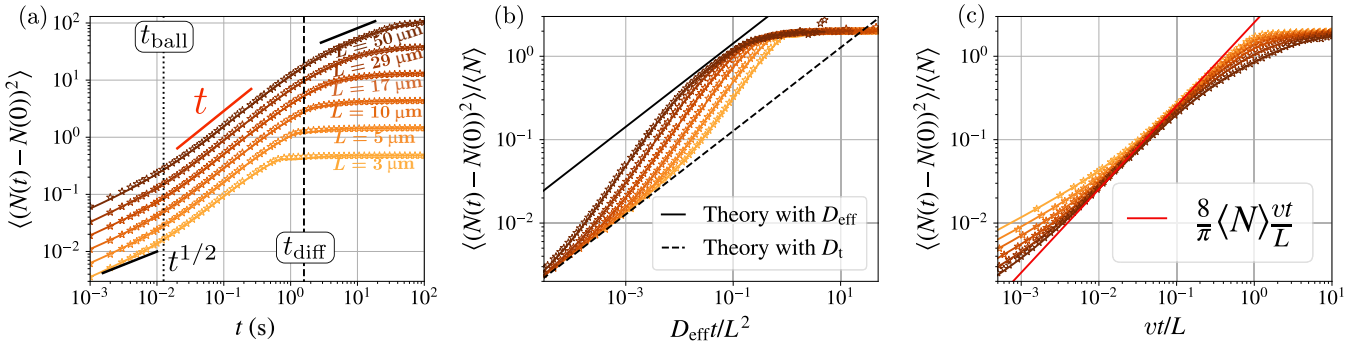


FIG. 4. NMSD behavior for run-and-tumble particles. (a) NMSD behavior with respect to lag time for increasing box sizes going from yellow to brown; (b) same as (a) but where time is rescaled by a typical diffusion timescale $L^2/D_{\rm eff}$ and the NMSD by $\langle N \rangle$ and (c) same as (b) but where time is rescaled by a typical ballistic time L/v. In all of the plots, stars correspond to numerical simulations and lines to the theory obtained from integrating $F_2(k,t)$ given by Eq. (30) inputted in Eq. (6). Physical parameters are the same as in Fig. 2 on simulations of 1500 noninteractive particles in a window of side 250 μm with periodic boundary conditions. Here we have $t_{\rm ball} = \pi \frac{D_t}{v^2}$ and $t_{\rm diff} = \pi \frac{D_{\rm eff}}{v^2}$ the crossover times between regimes for the NMSD.

initial configuration is lost, essentially when correlations in the number fluctuations have decayed sufficiently. The value of the plateau is given by the variance of the particle number $\lim_{t\to\infty} \langle \Delta N(t)^2 \rangle = 2(\langle N^2 \rangle - \langle N \rangle^2)$, as one can see from Eq. (1). In a noninteracting suspension, $\langle N^2 \rangle - \langle N \rangle^2 = \langle N \rangle$. The plateau behavior is confirmed as all NMSDs collapse at long times when they are rescaled by $\langle N \rangle$ [Figs. 4(b) and 4(c)]. The plateau is reached faster for smaller boxes where correlations are smaller, as initial particles are less likely to stay in the box or come back to the box.

The NMSD exhibits three different regimes in time, similarly to the MSD of RTP [38]. First, growing as $t^{1/2}$, then increasing faster, as $\propto t$ at intermediate times, and slower as $t^{1/2}$ at long times. Different rescalings can reveal this phenomena more accurately. First, we rescale time by $L^2/D_{\rm eff}$, the time it takes to diffuse across a box [Fig. 4(b)], and curves collapse at short and long times. Second, we rescale time by v/L, the time it takes for a particle to be advected across the box (c) and find excellent collapse of NMSDs at intermediate time. These regimes thus correspond to the two diffusive regimes at short and long time and the ballistic regime at intermediate time known for these self-propelled models. Notice that all regimes in NMSD space grow in time typically with a power that is the half of the power seen in the MSD, whereas in a MSD diffusive behavior scales linearly with t and ballistically as t^2, diffusive regimes in the NMSD increase as $t^{1/2}$ and the ballistic one as t. Overall, our theoretical predictions for the NMSD recover all the dynamical aspect of self-propelled particles, demonstrating their potential to probe and investigate dynamical behavior of self-propelled suspensions.

B. Limiting regimes

Further physical and analytical insights may be drawn by investigating the different limiting regimes of the NMSD.

1. Gaussian case

For Gaussian models this can easily be done, as the NMSD can be expressed directly in terms of the MSD, and Taylor expansions may readily be obtained at different timescales. At short timescales, Eq. (15) translates to

$$C_N(t) \simeq \langle N \rangle \left(1 - \frac{2}{\sqrt{\pi}} \sqrt{\frac{\langle \Delta r^2(t) \rangle}{L^2}} \right), \tag{31}$$

$$\langle \Delta N^2(t) \rangle \simeq \langle N \rangle \frac{4}{\sqrt{\pi}} \sqrt{\frac{\langle \Delta r^2(t) \rangle}{L^2}}. \tag{32}$$

One immediately sees that the exponents in time of the MSD are thus halved in the NMSD space. Plugging in the limiting expressions of the MSD (ballistic or diffusive) thus yields limiting expressions for the NMSD. For the case of AOUPs, diffusive regimes with diffusive coefficient D (where $D = D_t$ at short timescales or $D = D_{\rm eff}$ at longer timescales) can thus be expressed as

$$\langle \Delta N^2(t) \rangle \simeq \frac{8}{\sqrt{\pi}} \langle N \rangle \sqrt{\frac{Dt}{L^2}}, \tag{33}$$

and the ballistic regime as

$$\langle \Delta N^2(t) \rangle \simeq \frac{4}{\sqrt{\pi}} \langle N \rangle \frac{vt}{L}. \tag{34}$$

2. Diffusive limits in non-Gaussian cases

For a non-Gaussian model, limiting regimes can also be obtained but require more subtle investigations. For ABPs and RTPs, the first limiting regime can be obtained by neglecting the effect of activity, which is equivalent to truncating at order $N = 0$ in Eq. (25). This corresponds to a passive diffusive tracer with diffusion coefficient D_t. One then obtains a prediction for $C_N(t)$ given by Eq. (15) where $\langle \Delta r^2(t) \rangle = 4D_t t$; and hence we obtain at short times the same result as for AOUPs, given by Eq. (33) with $D = D_t$.

Long timescales and large length scales can be described through an alternative construction of a_0. Following the "fast field" truncation procedure used in Ref. [26], we take (i) $\partial_t a_{\pm 1} = 0$ as we suppose coefficients of order $n > 1$ relax "quickly," and (ii) over large length scales terms of order $\mathcal{O}(\nabla_r^2)$ can be neglected, as they sense small length scale variations—the so-called diffusion-drift approximation.

Looking back on Eq. (18) with these approximations we find

$$\partial_t a_0(\boldsymbol{r},t) = -\frac{v}{2}\nabla_r \cdot [\hat{i}_+ a_1 + \hat{i}_- a_{-1}] + D_t \Delta a_0$$
$$0 = -\frac{v}{2}\nabla_r \cdot [\hat{i}_\mp a_0] - (D_r + \alpha) a_{\pm 1} + \mathcal{O}(\nabla_r^2),$$

which we can easily solve for to find a partial differential equation for a_0 only,

$$\partial_t a_0(\boldsymbol{r},t) = D_{\text{eff}} \nabla^2 a_0,$$

where we recall that $D_{\text{eff}} = (\frac{v^2}{2(D_r+\alpha)} + D_t)$. This is the evolution equation of a diffusive process with effective diffusion coefficient D_{eff}. In this limit, the process is once again Gaussian and we recover Eq. (33) with $D = D_{\text{eff}}$. Diffusive regimes behave similarly for our Gaussian and non-Gaussian models. This is likely because the non-Gaussianity in RTP and ABP comes from the self-propulsion part of the dynamics which dominate only at intermediate timescales.

To check the relevance of these limiting diffusive laws, we present in Fig. 4(b), where time is rescaled by a diffusive timescale, the limit regime of Eq. (33) with $D = D_t$ and $D = D_{\text{eff}}$. The curves indeed all collapse at short timescales onto the limiting law with D_t and at long timescales onto that with D_{eff} confirming our analysis.

3. *Intermediate timescales*

Finding an expression for the ballistic regime of ABPs and RTPs requires more work. To do so, one option could be to neglect translational diffusion, taking $D_t = 0$ in Eq. (27), and inquire then about short-time behavior. However, this strategy is not sufficient as still the integrals in Eq. (6) cannot be solved. To circumvent this issue, we resort to the other expression of $C_N(t) = \langle N \rangle P_{\text{in}}(t)$ given in Eq. (2). Clearly, to obtain approximate predictions for $C_N(t)$, it is sufficient to search for approximate expressions for $P_{\text{in}}(t)$, the probability for a particle to still be in the box after some time t knowing it was in the box initially.

To address the intermediate timescales regime, we can consider that at these timescales, particles move with speed $v\boldsymbol{u}(\theta)$ and do not yet reorient ($D_r = \alpha = 0$). We also consider they do not diffuse ($D_t = 0$). Consider such a particle starting anywhere in a box at $t = 0$ – then, its probability distribution cloud can be represented as a blue square—as sketched in Fig. 5(a). At some later time t, this probability distribution is displaced by a vector $vt\boldsymbol{u}(\theta)$, as represented by the yellow square. The probability of the particle of still being in the box at time t corresponds to the overlap between the blue box and the yellow box. We can then write the probability $p_{\text{in}}(\theta,t)$ of the particle with orientation θ of still being in the box as

$$p_{\text{in}}(\theta,t) = \begin{cases} \frac{(L-v|\cos\theta|t)(L-v|\sin\theta|t)}{L^2} \\ 0 \text{ if } v\cos\theta t \geqslant L \text{ or } v\sin\theta t \geqslant L. \end{cases} \tag{35}$$

To unveil the limiting intermediate regime we restrict ourselves to short-enough times in the expression above such that $p_{\text{in}}(\theta,t)$ is nonzero and neglect terms of order $\mathcal{O}(t^2)$.

We can then average $p_{\text{in}}(\theta,t)$ over all the possible orientations θ to obtain at short times

$$P_{\text{in}}(t) = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\theta\, p_{\text{in}}(\theta,t) \simeq 1 - \frac{4}{\pi}\frac{v}{L}t. \tag{36}$$

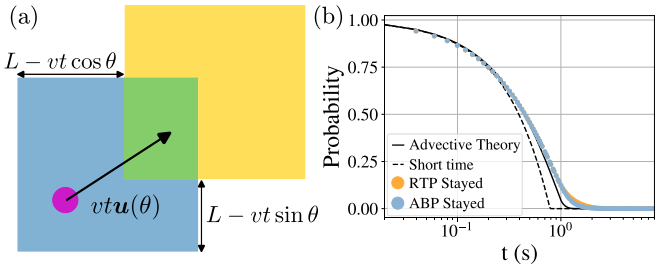


FIG. 5. Probability distribution from geometric arguments. Schematics (a) illustrating the probability distribution clouds to find an ballistic particle in a box after some time t (yellow) given it started in the box initially (blue). Probability distribution of ABPs or RTPs to remain in the box they start in (b) obtained from simulation against ones obtained from geometric arguments in the ballistic case. Simulation parameters are the same as in Fig. 2 and we consider boxes of size $L = 5\,\mu\text{m}$.

As an intermediate validation, we check that this equation accurately reproduces the probability that either an ABPs or an RTPs is still in a box at short-enough times in the intermediate regime, as obtained from simulations Fig. 5(b). The short-time approximation in Eq. (36) captures behavior accurately at short-enough times (black dashed line). Averaging over θ the "all time" expression of Eq. (35) numerically (black full line) pushes the agreement to longer timescales. At long-enough times, $P_{\text{in}}(t)$ is sensitive to reorientation dynamics and the probability that ABPs and RTPs are still in a box is, accordingly, no longer well captured by our approximations.

The probability $P_{\text{in}}(t)$ then gives $C_N(t)$ and the NMSD

$$\langle \Delta N(t)^2 \rangle \simeq \frac{8}{\pi}\langle N \rangle \frac{v}{L}t, \tag{37}$$

which is valid at short timescales. Equation (37) is presented in Fig. 4(c), and agrees perfectly with the full theory and simulation data at intermediate regimes. We thus obtain a very similar law as for Eq. (34) but with a prefactor that is higher compared to the AOUP case.

Why does this difference arise? Our probabilistic method for $P_{\text{in}}(t)$ can also be used to investigate AOUPs, but in that case, one has to realize that the probability distribution $p_{\text{in}}(\theta, |\boldsymbol{v}_{\text{OU}}|, t)$ of a particle with orientation θ to still be in the box at time t also depends on the norm of the velocity $|\boldsymbol{v}_{\text{OU}}|$, which is itself a distribution. Clearly the same reasoning as in Eq. (35) still applies, but then one needs to average over all possible values of θ and $|\boldsymbol{v}_{\text{OU}}|$. Each component of the speed vector $\boldsymbol{v}_{\text{OU}}$ of an AOUP is Gaussian distributed, thus the magnitude $|\boldsymbol{v}_{\text{OU}}|$ follows a Rayleigh distribution $P(|\boldsymbol{v}_{\text{OU}}|) = \frac{|\boldsymbol{v}_{\text{OU}}|\tau_r}{D_v} e^{-\frac{|\boldsymbol{v}_{\text{OU}}|^2\tau_r}{4D_v}}$. One then gets

$$P_{\text{in}}(t) = \frac{1}{2\pi}\int \mathrm{d}|\boldsymbol{v}_{\text{OU}}| P(|\boldsymbol{v}_{\text{OU}}|) \int_0^{2\pi} \mathrm{d}\theta\, p_{\text{in}}(\theta, |\boldsymbol{v}_{\text{OU}}|, t)$$
$$\simeq 1 - \frac{2}{\sqrt{\pi}}\frac{\sqrt{2D_v/\tau_r}}{L}t.$$

Identifying $\sqrt{2D_v/\tau_r} = v$ as the mean velocity of an AOUP, and inserting this expression for $P_{\text{in}}(t)$ in the NMSD, one recovers Eq. (34). The differences in the distribution of self-propulsion speeds thus translates to different prefactors in

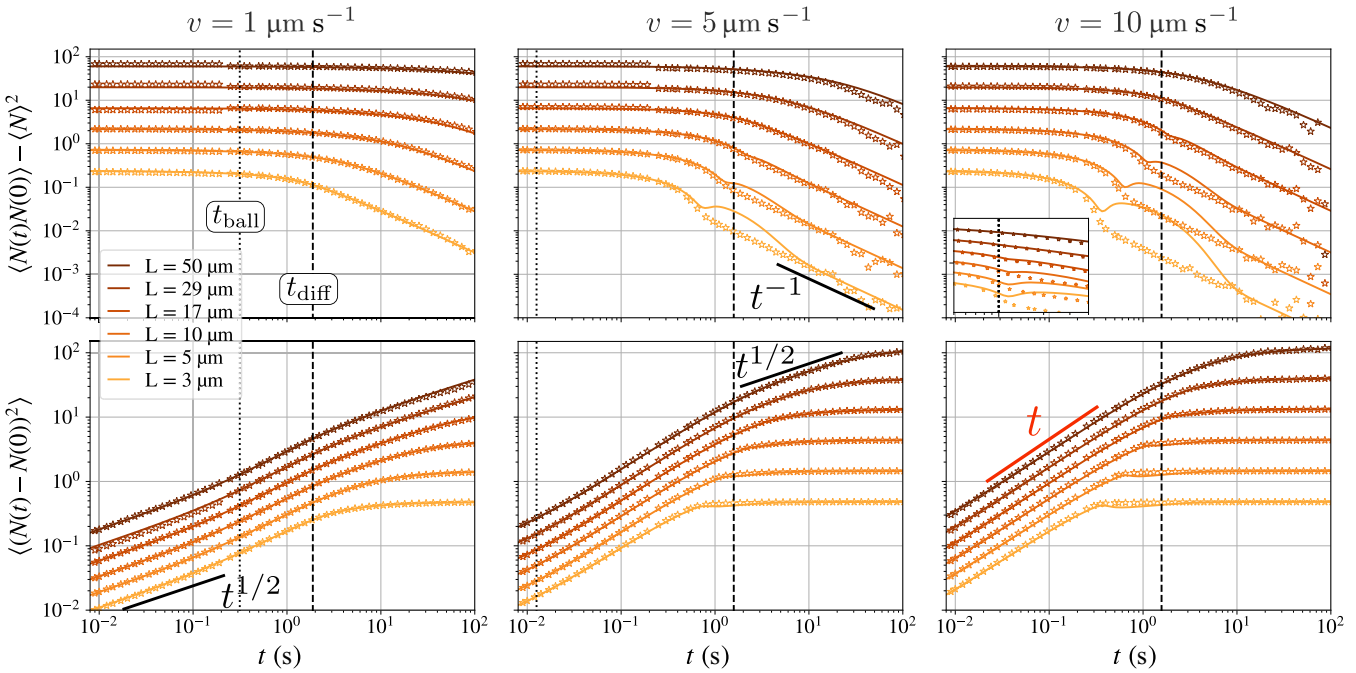


FIG. 6. Number fluctuations for run-and-tumble particles at different Péclet numbers. (Top row) Number correlations $C_N(t)$ and (bottom row) NMSD, with v increasing from left to right, $v = \{1, 5, 10\}$ μm s^{-1} corresponding to Pe $\simeq \{1.6, 8, 16\}$. All other dynamic parameters are kept constant at the same values as in Fig. 4. Stars correspond to simulated data and colored lines to the theory truncated at $N = 2$, using Eqs. (6) and (30). The inset in the rightmost column shows $C_N(t)$ with time rescaled by an ballistic timescale v/L and a vertical black dotted line at $t = v/L$.

the ballistic, intermediate time regime of the Countoscope. The difference in prefactors is small, of about 10%, since $4/\sqrt{\pi} \simeq 2.25$ and $8/\pi \simeq 2.55$; however, it is present. The non-Gaussianity of the ABP and RTP models is thus seen in the ballistic regime, and compared to a Gaussian model, differences in the NMSD originate from a peaked velocity distribution rather than from reorientation dynamics. We can show that differences in reorientation dynamics are more easily seen in the correlation function $C_N(t)$ [60].

All the limiting regimes that we have obtained provide simple analytical descriptions linking the behavior of each regime with the dynamic parameters that are relevant to quantify it (v, D_t, D_{eff}). These laws provide one way to quantify the motion of particles, using simple fits on number fluctuations calculated from experimental data.

C. Convergence at different orders for Countoscope metrics

It is known from study on the MSD of active particles that fitting the full expression should be preferred to fitting limiting regimes to obtain more accurate values of the dynamic parameters [61]. In this context, it is important to know at which order in the truncation one should stop to obtain accurate-enough results for the NMSD.

1. Convergence with respect to the Pe number

An important parameter affecting the accuracy of the truncation is the value of the Péclet number. This is expected from the construction of our $\widetilde{a}_0(\boldsymbol{k}, \omega)$ as highlighted in Sec. II A, given that the value of the continued fraction terms $\frac{c_n}{b_n}$ is larger at higher Pe numbers. Several authors have previously highlighted that typical truncation schemes lose precision at higher Pe [27,62]. Focusing on order $N = 2$ truncation, the highest order for which we obtained analytical results for $F(\boldsymbol{k}, t)$, we plot NMSD and $C_N(t)$ curves for RTPs at increasing Pe in Fig. 6. For this investigation we focus on RTPs, since we know RTP approximations converge more slowly than for ABP. By keeping all dynamic parameters the same, except for the self-propulsion speed that takes values $v = \{1, 5, 10\}$ μm s^{-1}, we probe Pe $\simeq \{1.6, 8, 16\}$, covering the range of most experimental systems.

For the lowest values of Pe, the particles' behavior is close to that of passive particles with the ballistic regime barely appearing in the NMSD. Our analytical expressions are accurate in this situation for both $C_N(t)$ and the NMSD. At intermediate Pe value, deviations between the theory and simulation data appear, especially in the number correlations $C_N(t)$ for the smallest boxes. This phenomenon is more pronounced at higher Pe value and visible in bigger boxes. Generally, the disagreement starts at timescales of order $t = L/v$, as seen in the inset of Fig. 6. Beyond this timescale, particles have typically traversed the box ballistically, and contributions to $P_{\text{in}}(t)$ are dominated by reorientation dynamics which determine whether a particle returns to a box it has left. Naturally, this effect is more pronounced in small boxes that particles are more likely to cross via self-propulsion without reorienting while they cross through it, and, therefore, the truncated theory fails to correctly predict behavior on smaller boxes. This behavior can be described through a critical lengthscale, $L_c \sim t_{\text{diff}} v = \frac{\pi D_{\text{eff}}}{v}$, which characterizes the typical maximum value of box sizes for which these effects should be seen. We can calculate that $L_c = \{1.9, 7.9, 15.7\}$ μm for our values of

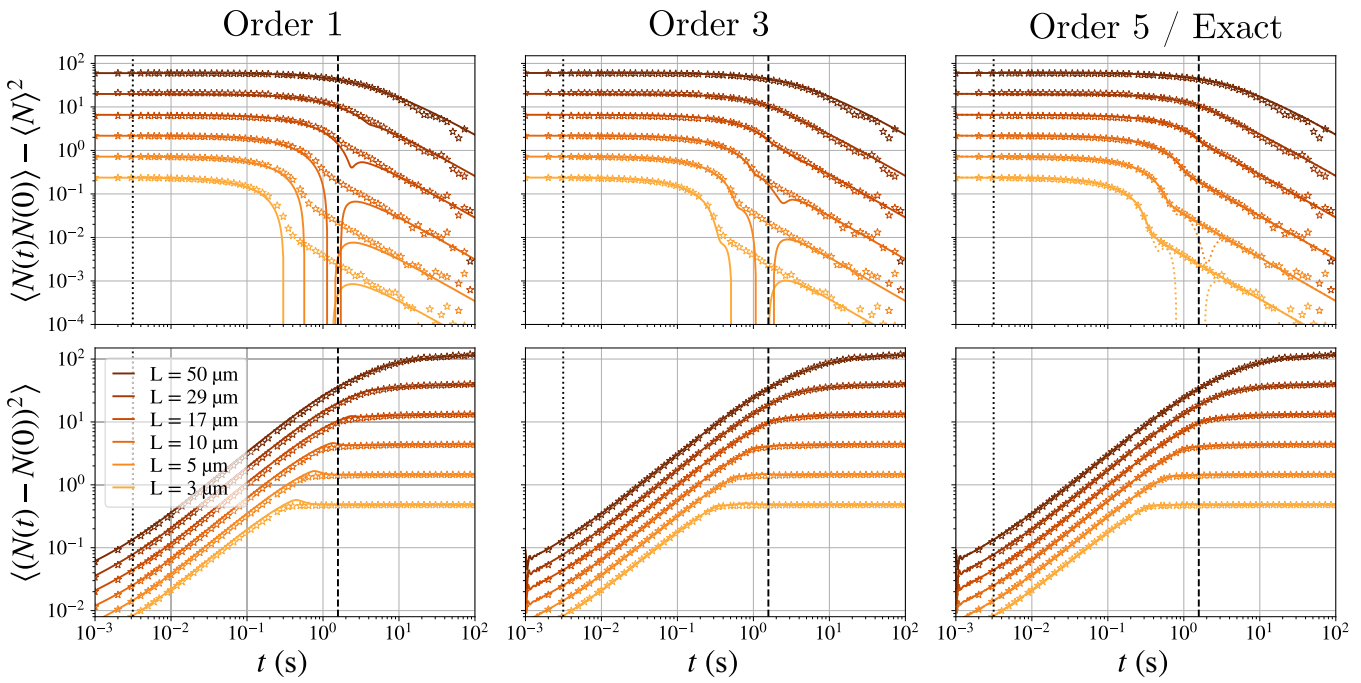


FIG. 7. Number correlations and NMSD for run-and-tumble particles at different orders in the truncation. Orders are increasing from left to right, with the rightmost column showing the exact expression in full lines and the order 5 result in dotted lines.

v. We find indeed that so long as $L \geqslant L_c$, the theory is in excellent agreement with the simulations.

Interestingly the departure from the theory in the correlation functions $C_N(t)$ are not as apparent in the NMSD (bottom row of Fig. 6). In fact, deviations occur primarily at long times in the correlation function, and their amplitude is small compared to the value of $C_N(t=0) = \langle N^2\rangle - \langle N\rangle^2$. In the NMSD space, this corresponds to small differences between the NMSD at late times and its plateau, and deviations between the theory and simulations are effectively hidden. The most noticeable differences occur only for the smallest box sizes (yellow) around the time where the NMSD reaches its plateau, at the highest Pe values. There the theory exhibits an unphysical albeit very little hump. Therefore, the visible inaccuracy in $C_N(t)$ has limited impact on the NMSD. The truncation at order $N = 2$ presents only noticeable disagreements in the NMSD for the highest Pe values and smallest box sizes, which are close to that which can be observed in the motion of low density bacteria.

2. *Convergence with respect to the order N of the truncation scheme*

We now restrict ourselves to investigations at high Pe number where disagreement between approximate theory and simulations is the most apparent, e.g., at a high self-propulsion speed $v = 10\ \mu\mathrm{m\,s}^{-1}$. We then look at how this disagreement can be reduced at increasingly higher truncation orders N in Fig. 7 (see also Fig. 10).

At order $N = 1$ (left column of Fig. 7) the theoretical approximation is, as expected, less accurate than for $N = 2$ that was investigated in Fig. 6 (right column). Differences between the theory and simulations are more striking and can even be seen clearly in the NMSD, at the time where the NMSD reaches its plateau, for several box sizes. In addition, at order 1 the prediction for the correlation function is that it becomes negative, which translates into a huge drops in log-log scale on smaller boxes. Because of an odd-even periodicity, these drops are typically only observed for odd truncation orders N, and not for even N.

At order $N = 3$ the accuracy is already noticeably better. Values of the correlation function still appear to become negative for some boxes but over a reduced time range and box size range. Accordingly, the NMSD presents a seemingly perfect agreement between theory and simulation for all observed box sizes. Order $N = 5$ is even more accurate, with some minor differences remaining in $C_N(t)$ and none for the NMSD. The exact formula in Eq. (26) shows excellent agreement at all time and length scales. Similar conclusions at all orders are also true for ABPs (see Fig. 9).

This investigation allows us to draw some intermediate conclusions on model choosing. At the NMSD level, there are barely any noticeable differences between orders $N = 3, 5$ and the exact expansion. It is thus natural to use $N = 2$ or $N = 3$ to capture the behavior of the NMSD in a practical application. For the correlation function however, higher-order approximations are necessary. While an order 3–5 is sufficient for ABPs (see Fig. 9), more orders are required to accurately reproduce RTPs, and hence one should use the exact formulation when considering specifically C_N.

3. *Convergence of the limiting regimes*

We can quantify this increasing accuracy by investigating the limiting regime at intermediate timescales. The exact behavior at intermediate timescales is given by Eq. (37), and

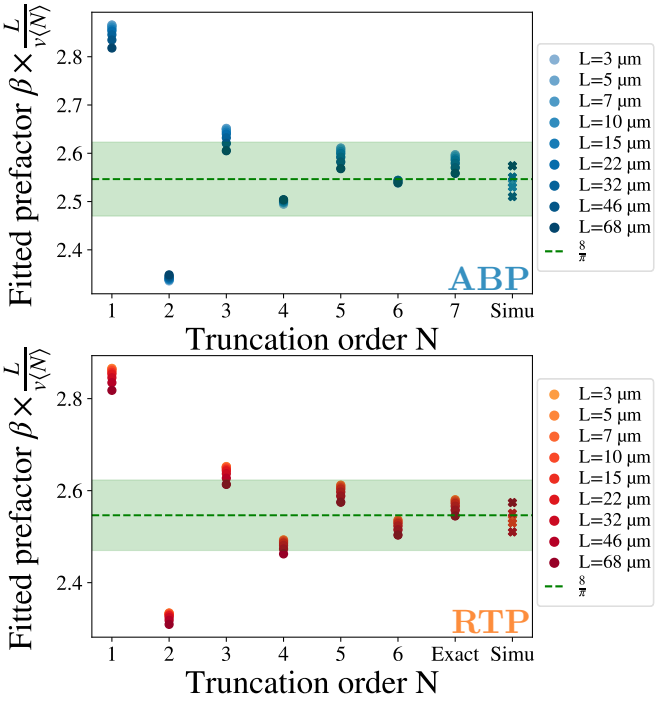


FIG. 8. Value of the prefactor to the ballistic regime for ABPs (top) and RTPs (bottom) obtained from fitting the theory at different orders compared to the theoretical value of $\frac{8}{\pi}$, with the green area representing a 3% confidence range.

our aim is to quantify how fast with increasing N we can capture this exact prefactor. To do so, we investigate a system at high Pe number. We then fit a linear law at short times, $\langle \Delta N^2(t)\rangle = \beta t$, where β is a fitting parameter. We then compare β with our theoretical expression in Eq. (37), and plot the results with increasing truncation order N in Fig. 8. Given we plot $\beta L/v\langle N\rangle$, we expect this prefactor converges to $8/\pi$.

The results obtained clearly approach the $8/\pi$ theoretical prediction with increasing truncation order N. Using the exact expression for the ISF of RTPs Eq. (26), we get as close to the correct value as our fit precision allows. Overall, the prefactor β oscillates with increasing order N, as odd orders overestimate the coefficient β while even orders underestimate it (see Fig. 10 for even values of N). This is intimately linked to the behavior observed in the curves of $C_N(t)$ where an upward or downward bump appeared compared to simulations for even and odd orders, respectively.

This can be explained from the structure of our continued fractions. Indeed we can rewrite through an equivalence transformation

$$\widetilde{a}_0^{-1} = b_0 + \underset{n=1}{\overset{N}{\mathrm{K}}} \frac{c}{b_n} = b_0 + \underset{n=1}{\overset{N}{\mathrm{K}}} \frac{1}{d_n b_n}$$

with

$$d_{2n} = 1, \text{ and } d_{2n+1} = \frac{1}{c}, \ \forall n \in \mathbb{N}.$$

This explains our oscillating behavior, as stopping at an even order will give us a smaller continued fraction, thus a bigger $\widetilde{a}_0$ and we will overestimate $C_N(t)$, and the opposite happens stopping at an odd order.

At any order we notice a degree of uncertainty around a mean value depending on the value of the box size used. Investigating fits of the NMSD obtained from simulations, knowing the self-propulsion velocity v shows that the value of the prefactor is roughly $\pm 2\%$ away from the expected prefactor depending on the box size (see the last point on the x axis "Simu" in Fig. 8). We can conclude that a roughly $\pm 2\%$ error is only a fitting error and that orders from $N \geqslant 3$ should provide satisfactory estimates of the self-propulsion speed.

We conclude by tempering the results of the fit in Fig. 8 by insisting that in order to correctly estimate dynamic coefficients one should fit the full expression instead of limiting regimes. We anticipate that the accuracy of order 3 for the NMSD should be enough in most situations to justify using it, as it admits an analytic result for the ISF and is thus easier to obtain for a wide range of values. Precise discussion on the use of our expressions at different orders to fit experimental data is, however, outside the scope of this paper and is left for future work.

V. CONCLUSION

We have here established theoretical expressions to quantify number fluctuations of three canonical models of self-propelled particles: ABPs, RTPs, and AOUPs. For all models, we relate number fluctuations to an integral of the ISFs through Eq. (6). AOUPs have Gaussian statistics and hence can be treated generically under a Gaussian framework; giving an ISF which can easily be integrated to give an explicit analytical expression Eq. (15) for the number correlations. For ABPs and RTPs, we have provided a novel derivation of the ISFs, allowing us to get both exact or approximate expressions—exact expressions are usually more precise but longer to evaluate numerically. Interestingly our derivation bridges together the methods consisting in solving the system through propagators and successive reorientation steps [21,24] or doing an expansion of the FPEs on the angular order of the perturbations to the density. The approach allows us to gain physical insights into the features of the ISFs and the number correlations: namely that oscillations originate from reorientation dynamics. All our theoretical expressions allow us to uncover three limiting regimes of the number mean squared difference in time: at short times diffusive, at intermediate times ballistic, and at long times diffusive with enhanced diffusion.

Here we conducted derivations for 2D systems and so our results are not straightforwardly applicable to, e.g., bacteria suspensions that are often found swimming in 3D [23]. However, similar formalisms may be used to obtain expressions for the ISFs in 3D [21,27] and so one can hope to use them to predict behavior of number fluctuations. While quantitatively these results might differ slightly, qualitatively we expect all features uncovered here to be consistently observed in 3D. Indeed, Gaussian models like AOUPs are similarly expressed in 3D as in 2D, with a slight modification of Eq. (15) into

$$C_N(t) = \langle N\rangle \left[f\left(\frac{\langle \Delta r^2(t)\rangle}{L^2} \right) \right]^d, \tag{38}$$

where d is the system's dimension. Therefore, we can expect all limiting regimes of the mean-squared displacement to arise

in the NMSD. Similarly, the derivations which provide the limiting laws Eqs. (37) and (33) can easily be extended to 3D. We can thus anticipate the use of the technique also in systems with 3D dynamics.

While we have focused here on noninteracting particle systems, it is natural to ask what kinds of behaviors would be observed at higher densities. Since real space number correlations provide easily physically interpretable signals, then we might expect they would also help to shed some light on collective behavior of active suspensions. Including a density-dependent self-propulsion for instance could be done following Ref. [26]. Remarkably, the investigation of *static* number fluctuations N in observation volumes has proven useful in quantifying *static* properties of dense nonequilibrium systems for over 30 years. For instance "giant" number fluctuations, where $\alpha > 0$ in the scaling $\langle N^2\rangle - \langle N\rangle^2 \sim N^{1+\alpha}$, indicate long-range organization in bacterial or synthetic active matter suspensions [63–74]. It is thus clear that investigating *dynamic* number fluctuations in virtual observation boxes like $N(t)$, is a promising route to quantify collective dynamic features of self-propelling suspensions.

ACKNOWLEDGMENTS

The authors acknowledge many fruitful discussions for the early elaboration of this work especially with Federico Paratore but also with Laura Alvarez, Lucio Isa, Ueli Töpfer, and Carolijn van Baalen. Naoufal Elaisati was especially helpful as he established some preliminary work on this topic. Further discussions with Maxime Déforêt, Timothée Gautry, Pierre Illien, Raphaël Jeanneret, Christina Kurzthaler, and Arnold Mathijssen were also helpful. We are also grateful to Adam Carter for sharing his code for making intermediate scattering functions.

T. Cerdin and S.M.: Conceptualization. T. Cerdin, S.M., and C.D.: Methodology. T. Cerdin: Formal analysis. T. Cerdin and S.M.: Overall investigation. T. Cerdin and T. Calazans: Investigation of limiting regimes. T. Cerdin: Visualization. C.D. and S.M.: Supervision. T. Cerdin, S.M., and C.D.: Writing—original draft. T. Cerdin, S.M., and C.D.: Writing—review and editing.

DATA AVAILABILITY

The data that support the findings of this article are openly available [75].

APPENDIX A: COMPUTATIONAL METHODS

For all three active particle models we simulate 1500 particles within a square simulation box of size $L_{\text{box}} = 250\,\mu\text{m}$ and periodic boundary conditions. As highlighted in Ref. [20], the periodic boundary conditions restrict the maximum size of boxes to be used for counting to about a third of the window size—otherwise periodic effects start to be seen. The dynamics are resolved using an Euler-Maruyama scheme to numerically integrate the stochastic equations of motion. In the case of RTPs we perform the random reorientation with a rate α using a Gillespie algorithm. To resolve the dynamics with sufficient accuracy over five orders of magnitude in time, we perform both short simulations with a very small time step $dt = 0.001\,\text{s}$ which is equal to the time step and longer simulations with time step $dt = 0.005$ while saving every 0.05 s. We then merge the results together on plots for $C_N(t)$ and $\langle \Delta N^2(t)\rangle$.

The ISF of simulations is computed on the long simulated datasets using the Python library [76] based on the direct method of calculation.

Numerical integration of the ISFs of order 3 and above are done using an inverse FFT with a custom routine in python. They are then saved for every set of dynamic parameters to be used to compute $C_N(t)$ at any box size.

Numerical integration of the $\boldsymbol{k}$ integrals in Eq. (6) for ABPs and RTPs are done following the scheme outlined in Ref. [19]. We carry out the $\boldsymbol{r}$ integrals:

$$C_N(t) = \langle N\rangle \int \frac{d\boldsymbol{k}}{(2\pi)^2} L^2 \prod_{i=x,y} \left(\frac{\sin\left(\frac{k_i L}{2}\right)}{k_i L/2}\right)^2 F(\boldsymbol{k},t). \quad \text{(A1)}$$

To facilitate the numerical integration, we change coordinates from Cartiesian $\{k_x L/2, kyL/2\}$ to rescaled spherical coordinates $\{K,\phi\}$. $C_N(t)$ can then be reexpressed as

$$C_N(t) = \langle N\rangle \int \frac{K dK}{\pi^2} f_v(K) F(2K/L,t), \quad \text{(A2)}$$

with

$$f_v(K) = \int_0^{2\pi} d\phi \left(\frac{\sin(K\cos\phi)}{K\cos\phi}\right)^2 \left(\frac{\sin(K\sin\phi)}{K\sin\phi}\right)^2. \quad \text{(A3)}$$

This separates the volumetric term in front of the integrals, which is quickly oscillating and complicates the numerical integration. As it is constant it can be computed once precisely and then reused subsequently in all Countoscope computations. Using it alongside the obtained values of $F(\boldsymbol{k},t)$ is how we speed up the final numerical integration.

All counting on simulated data to obtain $C_N(t)$ and the NMSD was done using the Countoscope code published in Ref. [77]. Upon publication, code for this specific investigation will be added on the same repository.

APPENDIX B: EQUIVALENCE BETWEEN HARMONIC COEFFICIENTS AND HYDRODYNAMIC FIELDS

To solve Eq. (16), an alternative method [27,28] is to use another decomposition of the probability distribution function (PDF) $P(\boldsymbol{r},\theta,t)$. Instead of using harmonic coefficients a_n, one can decompose the PDF in terms of hydrodynamic fields as

$$P(\boldsymbol{r},\theta,t) = \rho(\boldsymbol{r},t) + \boldsymbol{p}(\boldsymbol{r},t)\cdot\boldsymbol{u}(\theta) + \underline{\underline{q}}(\boldsymbol{r},t) : \left(\boldsymbol{uu} - \frac{\mathbb{I}}{2}\right). \quad \text{(B1)}$$

Writing out explicitly this expansion against the one we used [Eq. (17)] we have

$$\begin{aligned} P(\boldsymbol{r},\theta,t) &= \rho + p_x\cos\theta + p_y\sin\theta, \\ &\quad + \tfrac{1}{2}[(q_{xx}+q_{yy})\cos 2\theta + (q_{xy}+q_{yx})\sin 2\theta] \quad \text{(B2)} \\ &= a_0 + e^{i\theta}a_1 + e^{-i\theta}a_{-1} + e^{2i\theta}a_2 + e^{-2i\theta}a_{-2}, \quad \text{(B3)} \\ &= a_0 + (a_1+a_{-1})\cos\theta + i(a_1-a_{-1})\sin\theta \\ &\quad + [(a_2+a_{-2})\cos 2\theta + i(a_2-a_{-2})\sin 2\theta]. \quad \text{(B4)} \end{aligned}$$

From this we immediately deduce that we can express

$$\rho(\boldsymbol{r},t) = a_0(\boldsymbol{r},t), \tag{B5}$$

$$\boldsymbol{p}(\boldsymbol{r},t) = \begin{pmatrix} a_1 + a_{-1} \\ i(a_1 - a_{-1}) \end{pmatrix}. \tag{B6}$$

Since $\underline{\underline{q}}$ is chosen to be a symmetric traceless matrix we can find that $q_{xy} = q_{yx} = i(a_2 + a_{-2})$. For the diagonal components of $\underline{\underline{q}}$ we can compare Eq. (21) with the evolution equation obtained in Ref. [28]

$$\partial_t q_{ij} = -2B_{ijkl}\partial_l(v p_l) - (4D_r + \alpha)Q_{ij} + Dt\,\Delta Q_{ij} - \nabla_k \chi_{ijk},$$

where $B_{ijkl} = \frac{1}{4}(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk} - \delta_{ij}\delta_{kl})$ and summation over repeated indices are implied. Writing it out for the diagonal components and ignoring harmonic contributions of order 3 and above contained in χ_{ijk} we have

$$\partial_t q_{xx} = -\frac{v}{4}\nabla \cdot \left[\begin{pmatrix} 1 \\ -1 \end{pmatrix} \boldsymbol{p} \right] + D_t \Delta q_{xx} - (4Dr + \alpha) q_{xx}, \tag{B7}$$

$$\partial_t q_{yy} = -\frac{v}{4}\nabla \cdot \left[\begin{pmatrix} -1 \\ 1 \end{pmatrix} \boldsymbol{p} \right] + D_t \Delta q_{yy} - (4Dr + \alpha) q_{yy}. \tag{B8}$$

By direct comparison with Eq (21) and using the expression that we found for $\boldsymbol{p}$ we can show that

$$\underline{\underline{q}}(\boldsymbol{r},t) = \begin{pmatrix} \frac{1}{2}(a_2 + a_{-2}) & i(a_2 + a_{-2}) \\ i(a_2 + a_{-2}) & -\frac{1}{2}(a_2 + a_{-2}) \end{pmatrix}. \tag{B9}$$

The coefficients we derived can therefore be mapped onto hydrodynamic fields with clearer physical meaning.

Going at higher orders is not necessarily useful, since coefficients of order 3 and above are unclear in terms of physical meaning.

APPENDIX C: DETAILED SOLUTION AT ORDER 2

This Appendix details the calculation used to obtain the result for $F_2(k,t)$, the ISF obtained by truncating at order $N = 2$. We start from

$$F(\boldsymbol{k},t) = \int \frac{d\omega}{2\pi} e^{i\omega t} \widetilde{a}_{0,2}(\boldsymbol{k},\omega).$$

The denominator of $\widetilde{a}_{0,2}(\boldsymbol{k},\omega)$ is a polynomial of order 3 in ω that will be canceled by three roots. Introducing the notations

$$a = -8\alpha^3 - 60\alpha^2 D_r + 12\alpha D_r^2 + 280 D_r^3 + 81 D_r k^2 v^2$$
$$b = 4\alpha^2 + 20\alpha D_r + 52 D_r^2 - 9k^2 v^2,$$

we can find the roots to be

$$\omega_1 = \frac{1}{6} i \left((a + \sqrt{a^2 - b^3})^{1/3} + \frac{(a + \sqrt{a^2 - b^3})^{1/3}}{b} + 4\alpha + 10 D_r + 6 D_t k^2 \right)$$

$$\omega_2 = -\frac{1}{12} i \left((1 - i\sqrt{3})(a + \sqrt{a^2 - b^3})^{1/3} + (1 + i\sqrt{3}) \frac{(a + \sqrt{a^2 - b^3})^{1/3}}{b} - 8\alpha - 20 D_r - 12 D_t k^2 \right)$$

$$\omega_3 = -\frac{1}{12} i \left((1 + i\sqrt{3})(a + \sqrt{a^2 - b^3})^{1/3} + (1 - i\sqrt{3}) \frac{(a + \sqrt{a^2 - b^3})^{1/3}}{b} - 8\alpha - 20 D_r - 12 D_t k^2 \right).$$

These three poles are always in the superior part of the complex plane and therefore building a contour around it we get with the residue theorem

$$F_2(\boldsymbol{k},t) = i \left[\sum_{i=1}^{3} \mathrm{Res}[\widetilde{a}_{0,2}(\boldsymbol{k},\omega), \omega_i] \right]. \tag{C1}$$

The denominator of $\widetilde{a}_{0,2}(\boldsymbol{k},\omega)$ can be expressed in terms of these roots as

$$\begin{aligned} &(i\omega + D_t k^2) \Big[2(i\omega + D_t k^2 + 4D_r + \alpha) \\ &\quad \times (i\omega + D_t k^2 + D_r + \alpha) + \frac{v^2 k^2}{2} \Big] \\ &\quad + v^2 k^2 (i\omega + D_t k^2 + 4D_r + \alpha) \\ &= -2i \prod_{i=1}^{3} (\omega - \omega_i). \end{aligned}$$

The residues are then easily calculated by replacing the denominator, and introducing the numerator divided by 2 as $h(\omega) = (i\omega + D_t k^2 + 4D_r + \alpha)(i\omega + D_t k^2 + D_r + \alpha) + \frac{v^2 k^2}{4}$ we find the exact expression from the main text Eq. (30),

$$F_2(\boldsymbol{k},t) = -\sum_{i=1}^{3} \frac{h(\omega_i) e^{i\omega_i t}}{\prod_{j \neq i} (\omega_i - \omega_j)}. \tag{C2}$$

As mentioned, this method is general and can be applied to solve at order 1 easily and at order 3 to obtain a lengthier solution. At order 4 in the truncation, however, the denominator will be of degree 5 in ω and it will not be possible to obtain general analytical solutions of its roots and thus of the integral anymore.

APPENDIX D: ADDITIONAL FIGURES

These two figures illustrate the behavior of Number Correlations and NMSD at increasing orders in the truncation for ABPs (Fig. 9), and of Number Correlations at even orders in the truncation (Fig. 10).

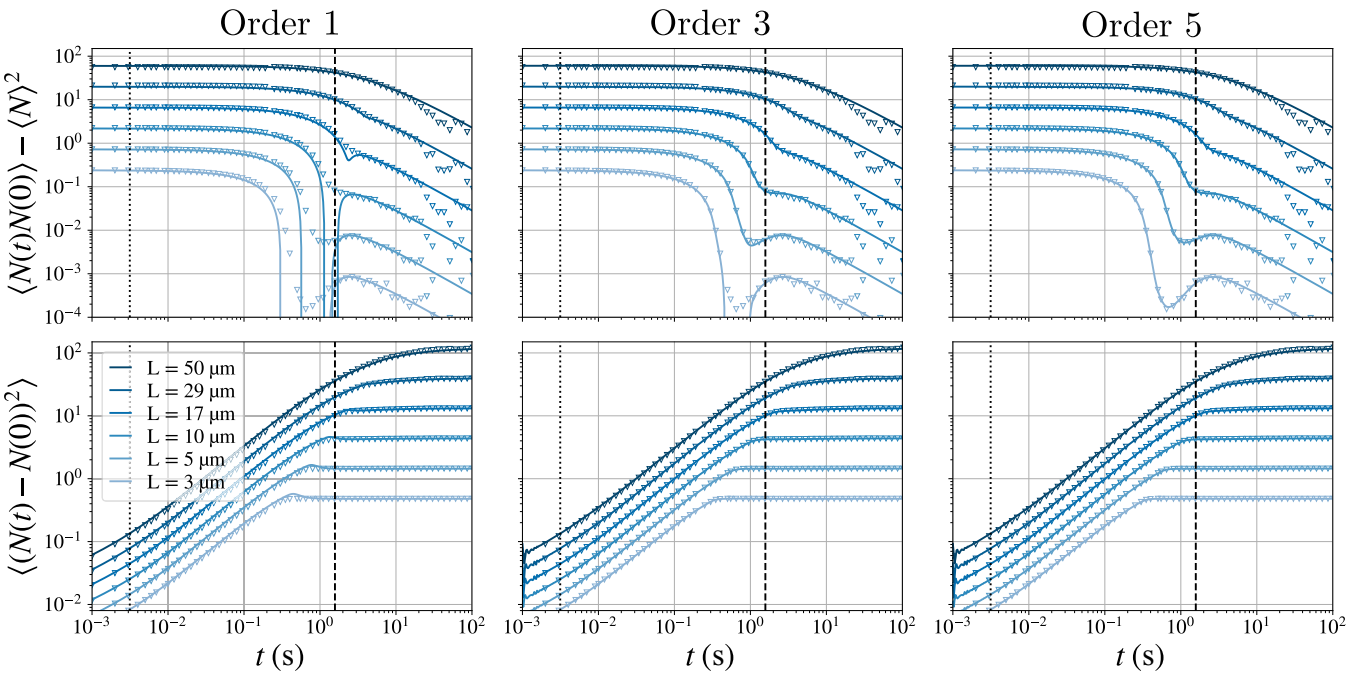


FIG. 9. Number correlations and NMSD for active Brownian particles at different orders in the truncation.

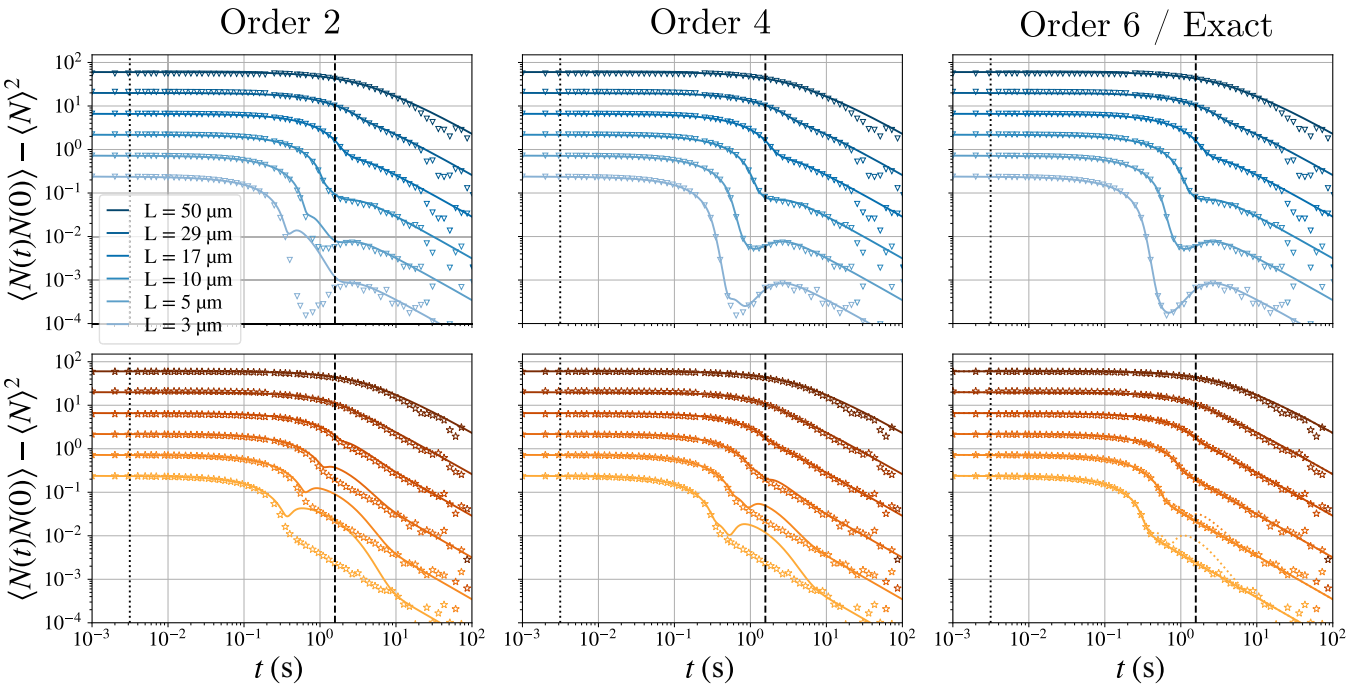


FIG. 10. Number correlations of active Brownian particles (top) and run-and-tumble particles (bottom) at pair orders in the truncation. Orders are increasing from left to right, with the rightmost column for RTP showing the exact expression in full lines and the order 6 result in dotted lines. Dynamic parameters are $v = 10\,\mu\mathrm{m\,s^{-1}}$, $D_t = 0.1\,\mu\mathrm{m^2\,s^{-1}}$, and $D_r = 1\,\mathrm{s^{-1}}$.

[1] M. Smoluchowski, Studien über molekularstatistik von emulsionen und deren zusammenhang mit der Brown'schen bewegung, Sitzungsber. Akad. Wiss. Wien **123**, 2381 (1914).

[2] M. von Smoluchowski, Uber die zeitliche veränderlichkeit der gruppierung von emulsionsteilchen und die reversibilität der diffusionserscheinungen, Phys. Z. **16**, 321 (1915).

[3] M. v. Smoluchowski, Studien über Kolloidstatistik und den Mechanismus der Diffusion, Kolloid-Z. **18**, 48 (1916).

[4] S. Chandrasekhar, Stochastic problems in physics and astronomy, Rev. Mod. Phys. **15**, 1 (1943).

[5] A. Westgren, Die veranderungsgeschwindigkeit der lokalen teilchenkonzentration in kolloiden systemen, I, Ark. Mat. Astron. Fys. **11**, 1 (1916).

[6] A. Westgren, Die veranderungsgeschwindigkeit der lokalen teilchenkonzentration in kolloiden systemen, II, Ark. Mat. Astron. Fys. **13**, 1 (1918).

[7] D. Lindley, The estimation of velocity distributions from counts, in *Proceedings of the International Congress of Mathematicians* (North-Holland, Amsterdam, 1954), pp. 427–444.

[8] L. Rothschild, A new method of measuring the activity of spermatozoa, J. Exp. Biol. **30**, 178 (1953).

[9] H. Ruben, Generalised concentration fluctuations under diffusion equilibrium, J. Appl. Probab. **1**, 47 (1964).

[10] D. Katz and H. Dott, Methods of measuring swimming speed of spermatozoa, Reproduction **45**, 263 (1975).

[11] W. Culling, Estimation of the mean velocity of particulate flows by counting, Earth Surf. Process. Landf. **10**, 569 (1985).

[12] M. t. Vrugt, B. Liebchen, and M. E. Cates, Colloquium: What do we mean by 'active matter'? Rev. Mod. Phys. **98**, 031001 (2026).

[13] F. Baker, R. Cragle, G. Salisbury, and N. VanDemark, Spermatozoan velocities in vitro. A simple method of measurement, Fertil. Steril. **8**, 149 (1957).

[14] B. J. Berne and R. Pecora, *Dynamic Light Scattering: With Applications to Chemistry, Biology, and Physics* (Courier Corporation, North Chelmsford, MA, 2000).

[15] E. L. Elson and D. Magde, Fluorescence correlation spectroscopy. I. Conceptual basis and theory, Biopolymers **13**, 1 (1974).

[16] J.-P. Günther, M. Börsch, and P. Fischer, Diffusion measurements of swimming enzymes with fluorescence correlation spectroscopy, Acc. Chem. Res. **51**, 1911 (2018).

[17] I. Santiago, Nanoscale active matter matters: Challenges and opportunities for self-propelled nanomotors, Nano Today **19**, 11 (2018).

[18] A. McGlasson and L. C. Bradley, Investigating time-dependent active motion of Janus micromotors using dynamic light scattering, Small **17**, 2104926 (2021).

[19] E. K. R. Mackay, S. Marbach, B. Sprinkle, and A. L. Thorneywork, The countoscope: Measuring self and collective dynamics without trajectories, Phys. Rev. X **14**, 041016 (2024).

[20] A. Carter, E. K. Mackay, B. Sprinkle, A. L. Thorneywork, and S. Marbach, Measuring collective diffusion coefficients by counting particles in boxes, Soft Matter **21**, 3991 (2025).

[21] K. Martens, L. Angelani, R. Di Leonardo, and L. Bocquet, Probability distributions for the run-and-tumble bacterial dynamics: An analogy to the Lorentz model, Eur. Phys. J. E **35**, 84 (2012).

[22] C. Kurzthaler, S. Leitmann, and T. Franosch, Intermediate scattering function of an anisotropic active Brownian particle, Sci. Rep. **6**, 36702 (2016).

[23] C. Kurzthaler, Y. Zhao, N. Zhou, J. Schwarz-Linek, C. Devailly, J. Arlt, J.-D. Huang, W. C. Poon, T. Franosch, J. Tailleur, and V. A. Martinez, Characterization and control of the run-and-tumble dynamics of *Escherichia coli*, Phys. Rev. Lett. **132**, 038302 (2024).

[24] Y. Zhao, C. Kurzthaler, N. Zhou, J. Schwarz-Linek, C. Devailly, J. Arlt, J.-D. Huang, W. C. K. Poon, T. Franosch, V. A. Martinez, and J. Tailleur, Quantitative characterization of run-and-tumble statistics in bulk bacterial suspensions, Phys. Rev. E **109**, 014612 (2024).

[25] D. Martin, J. O'Byrne, M. E. Cates, E. Fodor, C. Nardini, J. Tailleur, and F. van Wijland, Statistical mechanics of active Ornstein-Uhlenbeck particles, Phys. Rev. E **103**, 032607 (2021).

[26] A. Dinelli, J. O'Byrne, and J. Tailleur, Fluctuating hydrodynamics of active particles interacting via chemotaxis and quorum sensing: Static and dynamics, J. Phys. A: Math. Theor. **57**, 395002 (2024).

[27] T. Gautry, M. Deforet, and P. Illien, Closures of moment expansion of anisotropic active Brownian particles, J. Stat. Mech. (2026) 013207.

[28] A. P. Solon, M. E. Cates, and J. Tailleur, Active Brownian particles and run-and-tumble particles: A comparative study, Eur. Phys. J. Spec. Top. **224**, 1231 (2015).

[29] D. Nelson, *The Penguin Dictionary of Mathematics* (Penguin UK, London, 2008).

[30] E. K. Mackay, A. D. Young, A. Carter, S. Marbach, and A. L. Thorneywork, Collective hard core interactions leave multiscale signatures in number fluctuation spectra, arXiv:2512.17476.

[31] *E. coli in Motion*, edited by H. C. Berg, Biological and Medical Physics, Biomedical Engineering (Springer, New York, 2004).

[32] E. Clement, A. Lindner, C. Douarche, and H. Auradou, Bacterial suspensions under flow, Eur. Phys. J. Spec. Top. **225**, 2389 (2016).

[33] M. Polin, I. Tuval, K. Drescher, J. P. Gollub, and R. E. Goldstein, Chlamydomonas swims with two "gears" in a eukaryotic version of run-and-tumble locomotion, Science **325**, 487 (2009).

[34] C. Kurzthaler, C. Devailly, J. Arlt, T. Franosch, W. C. Poon, V. A. Martinez, and A. T. Brown, Probing the spatiotemporal dynamics of catalytic Janus particles with single-particle tracking and differential dynamic microscopy, Phys. Rev. Lett. **121**, 078001 (2018).

[35] A. Poncet, O. Bénichou, V. Démery, and D. Nishiguchi, Pair correlation of dilute active Brownian particles: From low-activity dipolar correction to high-activity algebraic depletion wings, Phys. Rev. E **103**, 012605 (2021).

[36] A. R. Sprenger, M. A. Fernandez-Rodriguez, L. Alvarez, L. Isa, R. Wittkowski, and H. Löwen, Active Brownian motion with orientation-dependent motility: Theory and experiments, Langmuir **36**, 7066 (2020).

[37] M. A. Fernandez-Rodriguez, F. Grillo, L. Alvarez, M. Rathlef, I. Buttinoni, G. Volpe, and L. Isa, Feedback-controlled active Brownian colloids with space-dependent rotational dynamics, Nat. Commun. **11**, 4223 (2020).

[38] J. R. Howse, R. A. L. Jones, A. J. Ryan, T. Gough, R. Vafabakhsh, and R. Golestanian, Self-motile colloidal particles: From directed propulsion to random walk, Phys. Rev. Lett. **99**, 048102 (2007).

[39] T. H. N. Minh, B. Rotenberg, and S. Marbach, Ionic fluctuations in finite volumes: Fractional noise and hyperuniformity, Faraday Discuss. **246**, 225 (2023).

[40] V. Démery and D. S. Dean, The conductivity of strong electrolytes from stochastic density functional theory, J. Stat. Mech. (2016) 023106.

[41] M. Jardat, V. Dahirel, and P. Illien, Diffusion of a tracer in a dense mixture of soft particles connected to different thermostats, Phys. Rev. E **106**, 064608 (2022).
[42] S. Mahdisoltani and R. Golestanian, Transient fluctuation-induced forces in driven electrolytes after an electric field quench, New J. Phys. **23**, 073034 (2021).
[43] O. Bernard, M. Jardat, B. Rotenberg, and P. Illien, On analytical theories for conductivity and self-diffusion in concentrated electrolytes, J. Chem. Phys. **159**, 164105 (2023).
[44] P. Illien, The Dean–Kawasaki equation and stochastic density functional theory, Rep. Prog. Phys. **88**, 086601 (2025).
[45] Y. Kuroda and K. Miyazaki, Microscopic theory for hyperuniformity in two-dimensional chiral active fluid, J. Stat. Mech. (2023) 103203.
[46] S. P. Das and A. Yoshimori, Coarse-grained forms for equations describing the microscopic motion of particles in a fluid, Phys. Rev. E **88**, 043008 (2013).
[47] M. A. Durán-Olivencia, P. Yatsyshin, B. D. Goddard, and S. Kalliadasis, General framework for fluctuating dynamic density functional theory, New J. Phys. **19**, 123022 (2017).
[48] S. K. Yadav and S. P. Das, Hydrodynamic equations for active matter: Consequences of Galilean transformations, J. Phys. Soc. Jpn. **91**, 084601 (2022).
[49] R. Kubo, Generalized cumulant expansion method, J. Phys. Soc. Jpn. **17**, 1100 (1962).
[50] R. Metzler and M. Weiss, Discriminating stochastic processes for the assessment of materials properties by diffusion measurements, Phys. Chem. Chem. Phys. **27**, 14350 (2025).
[51] A. Ahmadi, M. C. Marchetti, and T. B. Liverpool, Hydrodynamics of isotropic and liquid crystalline active polymer solutions, Phys. Rev. E **74**, 061913 (2006).
[52] M. C. Marchetti, J.-F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao, and R. A. Simha, Hydrodynamics of soft active matter, Rev. Mod. Phys. **85**, 1143 (2013).
[53] S. Bureković, F. De Luca, M. E. Cates, and C. Nardini, Active Cahn–Hilliard theory for non-equilibrium phase separation: Quantitative macroscopic predictions and a microscopic derivation, Phys. Rev. Res. **8**, 033141 (2026).
[54] G. Spera, C. Duclut, M. Durand, and J. Tailleur, Nematic torques in scalar active matter: When fluctuations favor polar order and persistence, Phys. Rev. Lett. **132**, 078301 (2024).
[55] J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids: With Applications to Soft Matter* (Academic Press, San Diego, CA, 2013).
[56] E. H. Hauge, Exact and Chapman-Enskog solutions of the Boltzmann equation for the Lorentz model, Phys. Fluids **13**, 1201 (1970).
[57] A. Plesner, H. H. B. Sørensen, and S. Hauberg, Accurate computation of the logarithm of modified Bessel functions on GPUs, in *Proceedings of the 38th ACM International Conference on Supercomputing, ICS '24* (Association for Computing Machinery, New York, 2024), pp. 213–224.
[58] P. Levitz, L. Michot, N. Malikova, M. Scheel, and T. Weitkamp, Probing particle dynamics in a fully opaque porous network using X-ray differential dynamic radiography (XDDR), Soft Matter **21**, 3067 (2025).
[59] M. Li, Y. Li, H. P. Zhang, and Y. Zhao, Detecting active lévy particles using differential dynamic microscopy, Phys. Rev. E **113**, 064403 (2026).
[60] T. Cerdin, S. Marbach, and C. Douarche, Number fluctuations distinguish different self-propelling dynamics, arXiv:2604.02872.
[61] M. R. Bailey, A. R. Sprenger, F. Grillo, H. Löwen, and L. Isa, Fitting an active Brownian particle's mean-squared displacement with improved parameter estimation, Phys. Rev. E **106**, L052602 (2022).
[62] S. Bröker, M. te Vrugt, J. Jeggle, J. Stenhammar, and R. Wittkowski, Pair-distribution function of active Brownian spheres in three spatial dimensions: Simulation results and analytical representation, Soft Matter **20**, 224 (2023).
[63] H. P. Zhang, A. Be'er, E.-L. Florin, and H. L. Swinney, Collective motion and density fluctuations in bacterial colonies, Proc. Natl. Acad. Sci. USA **107**, 13626 (2010).
[64] F. Peruani, J. Starruß, V. Jakovljevic, L. Søgaard-Andersen, A. Deutsch, and M. Bär, Collective motion and nonequilibrium cluster formation in colonies of gliding bacteria, Phys. Rev. Lett. **108**, 098102 (2012).
[65] Z. Liu, W. Zeng, X. Ma, and X. Cheng, Density fluctuations and energy spectra of 3D bacterial suspensions, Soft Matter **17**, 10806 (2021).
[66] Y. Fily and M. C. Marchetti, Athermal phase separation of self-propelled particles with no alignment, Phys. Rev. Lett. **108**, 235702 (2012).
[67] S. Dey, D. Das, and R. Rajesh, Spatial structures and giant number fluctuations in models of active matter, Phys. Rev. Lett. **108**, 238001 (2012).
[68] F. Alarcón, C. Valeriani, and I. Pagonabarraga, Morphology of clusters of attractive dry and wet self-propelled spherical particle suspensions, Soft Matter **13**, 814 (2017).
[69] H. Chaté, F. Ginelli, G. Grégoire, and F. Raynaud, Collective motion of self-propelled particles interacting without cohesion, Phys. Rev. E **77**, 046113 (2008).
[70] O. Chepizhko, D. Saintillan, and F. Peruani, Revisiting the emergence of order in active matter, Soft Matter **17**, 3113 (2021).
[71] F. Fadda, D. A. Matoz-Fernandez, R. van Roij, and S. Jabbari-Farouji, The interplay between chemo-phoretic interactions and crowding in active colloids, Soft Matter **19**, 2297 (2023).
[72] V. Narayan, S. Ramaswamy, and N. Menon, Long-lived giant number fluctuations in a swarming granular nematic, Science **317**, 105 (2007).
[73] J. Toner, Y. Tu, and S. Ramaswamy, Hydrodynamics and phases of flocks, Ann. Phys. **318**, 170 (2005).
[74] R. M. Navarro and S. M. Fielding, Clustering and phase behaviour of attractive active particles with hydrodynamics, Soft Matter **11**, 7525 (2015).
[75] T. Cerdin, S. Marbach, and C. Douarche, Active Countoscope [Computer software], Zenodo, 2026, https://doi.org/10.5281/zenodo.23082504.
[76] A. Carter and R. Fish, adamrobcarter/scattering_functions: v0.1.2 (Version v0.1.2) [Computer software], Zenodo, 2025, https://doi.org/10.5281/zenodo.17362168.
[77] S. Marbach, A. Carter, and B. Sprinkle, Countoscope/countoscope_prx: v1.0.1, (Version v1.0.1) [Computer software], Zenodo, 2025, https://doi.org/10.5281/zenodo.15000583.