


Quantitative kinetic theory of flocking with three-particle closure

Rüdiger Kürsten  and Thomas Ihle

Institut für Physik, Universität Greifswald, Felix-Hausdorff-Strasse 6, 17489 Greifswald, Germany

 (Received 25 February 2021; revised 9 July 2021; accepted 13 August 2021; published 3 September 2021)

We consider aligning self-propelled particles in two dimensions. Their motion is given by Langevin equations and includes nonadditive N -particle interactions. The qualitative behavior is as for the famous Vicsek model. We develop a kinetic theory of flocking beyond mean field. In particular, we self-consistently take into account the full pair correlation function. We find excellent quantitative agreement of the pair correlations with direct agent-based simulations within the disordered regime. Furthermore we use a closure relation to incorporate spatial correlations of three particles. In that way we achieve good quantitative agreement of the onset of flocking with direct simulations. Compared to mean-field theory, the flocking transition is shifted significantly toward lower noise because directional correlations favor disorder. We compare our theory with a recently developed Landau-kinetic theory.

DOI: [10.1103/PhysRevE.104.034604](https://doi.org/10.1103/PhysRevE.104.034604)

I. INTRODUCTION

Entities equipped with a propulsion mechanism, that is active matter, transfer free energy into directed motion. Such entities exist from micro- or even nanolength scales up to macroscopic size. Examples are manmade microswimmers, bacteria, insects, fish, mammals and robots.

Large groups of interacting active particles can exhibit complex emergent collective phenomena that differ fundamentally from their equilibrium counterparts. Some prominent examples are flocking [1,2], motility-induced phase separation [3,4], and bacterial turbulence [5]. An overview on the rapidly developing field of active matter and its application can be found, e.g., in the reviews in Refs. [6–14].

The Vicsek model [1] is historically one of the first and computationally one of the simplest active models that exhibits a flocking transition. Therefore, it is considered as one of the prototype models of active matter. In this model, point particles move in two dimensions at constant speed v in individual directions given by the polar angles ϕ_i , $i \in \{1, \dots, N\}$. At points in time $\tau, 2\tau, 3\tau, \dots$ all particle directions ϕ_i are changed due to interactions in the following way: Each particle takes the direction of the average velocity of all particles within distance R (including the particle itself), disturbed by a noise term. Thus, the interactions favor a local alignment of the particle velocities.

For large systems and periodic boundary conditions the Vicsek model is known in four phases [1,12,15–18]. For small noise or large particle densities, on average all particles move in a similar direction (i) [1].

Hence, we call the system polarly ordered. Furthermore, particles tend to cluster together locally. However, the clusters are distributed equally over space and thus the particle distribution is homogeneous on larger length scales. By increasing the noise strength or decreasing the particle density the system arranges in a “cross sea” phase (ii) [18], where high density

regions that look like crossing wave fronts are formed. This pattern moves through the system and the low density regions have almost no polar order. For even larger noise strength or smaller densities the system arranges in parallel nonintersecting traveling high density bands (iii) [15]. Again, there is almost no polar order in the low density regions between the high density bands. In phases (ii) and (iii) there is still an average polar order. The main contributions to the polar order come from the high density regions. For very large noise strength or very small particle densities there is no polar order (iv). That is, for large systems there is no motion of the center of mass and the particles are distributed homogeneously.

There have been qualitative descriptions of some of the aforementioned phases by field- or kinetic theories [2,19–25]. Those theories either do not reach quantitative agreement with direct simulations or only for very special parameters. In addition, most theories rely on the mean-field assumption. There are also kinetic theories of active systems that consider weak pair correlations to some extent; see, e.g., Refs. [26–29]. However, for the Vicsek model, recently it has been shown quantitatively, that correlations of two and more particles are important even in large parts of the disordered phase (iv) [30].

The aim of this paper is to provide a more quantitative theoretical description of the model. For technical reasons we are not studying the Vicsek model itself but a similar model that is believed to behave qualitatively equivalent to the Vicsek model. Major problems in kinetic theories of the standard Vicsek model are difficulties related to the finite time step as well as the presence of multiparticle collision integrals that are not analytically solvable; see, e.g., Ref. [22]. The latter was circumvented in Ref. [26] by using binary interactions with randomly selected interaction partners.

In this paper we develop a ring-kinetic theory. That means, we consider the dynamical equations for the one-particle distribution and for the two-particle correlation function explicitly. Higher order correlations are neglected in the first

step. However, the effects of the pair correlations on the one-particle distribution as well as on the pair correlation function itself, are fully taken into account. This concept has been applied in various fields; see, e.g., Refs. [31–34]. In this paper, we set up the ring-kinetic equations for N -particle interactions, a significant difference to most previous applications. That means the forces do not depend on the state of only two particles but on the positions of all particles.

Within the disordered parameter regime, we find excellent quantitative agreement between the ring-kinetic theory and agent-based simulations. For a moderate particle density, we find the threshold noise for the onset of flocking within the ring-kinetic theory to be very close to the value measured in agent-based simulations. In particular, the results are much more precise than the predictions of mean-field theory.

For larger particle densities, at the onset of flocking, higher-order correlations are more important. We extend the ring-kinetic theory and incorporate also spatial three-particle correlations via a closure ansatz. In that way, the flocking transition can be described also for larger densities, and the agreement between theory and simulations is further improved at moderate densities, where we reach quantitative agreement within the considered resolution. In addition, for higher densities, our theory is a significant improvement over mean field.

The paper is organized as follows. In Sec. II we define and discuss the Vicsek-like model studied here. In Sec. III we develop a homogeneous mean-field theory and calculate the critical noise strength of the flocking transition. In Sec. IV we introduce the notion of many particle correlations. In Sec. V we develop the full ring-kinetic theory, that is the dynamical equations for the one-particle distribution and the pair correlation function neglecting higher order correlations. In Sec. VI we present extensive quantitative comparisons between the solutions of the ring-kinetic equations and direct agent-based simulations. We identify the parameter region where the ring-kinetic theory is applicable. In Sec. VII we employ a closure relation to take into account spatial three-particle correlations. In that way we can significantly enlarge the applicability domain of the kinetic theory. In Sec. VIII we compare the results of the kinetic theory presented in this work with the simplified Landau kinetic theory of Patelli [29]. In Sec. IX we discuss our results and give an outlook to possible extensions of the method. In Appendix A we explicitly give the time evolution equations of the one-particle distribution and of the pair correlation function. In Appendix B we give all relevant equations in Fourier space that have been used to evaluate the kinetic equations numerically.

II. MODEL

We consider a Vicsek-like model in continuous time that was investigated as here, or in a similar form in many studies; see, e.g., Refs. [35–38]. The model is given by

$$\begin{aligned} \dot{x}_i &= v \cos(\phi_i), \\ \dot{y}_i &= v \sin(\phi_i), \\ \dot{\phi}_i &= w(|\Omega(i)|) \sum_{j \in \Omega(i)} \sin(\phi_j - \phi_i) + \sigma \xi_i, \quad i = 1, \dots, N, \end{aligned} \quad (1)$$

where $\mathbf{r}_i(t) = [x_i(t), y_i(t)]$ denote the particles positions and ϕ_i the directions of the particles velocities. The set

$$\Omega_i := \{j \in \{1, \dots, N\} : |\mathbf{r}_j - \mathbf{r}_i| \leq R\} \quad (2)$$

contains all particles that are within distance R to particle i . The $\xi_i(t)$ are independent Gaussian white noise terms satisfying

$$\langle \xi_i(t) \xi_j(s) \rangle = \delta_{ij} \delta(t - s). \quad (3)$$

The noise strength is given by σ , and $w(n)$ is an interaction weight function that depends on the number of neighbors of particle i (particles that are within distance R including particle i itself).

We consider the two-dimensional motion of N particles that move at constant speed v in individual directions ϕ_i . The directions of neighboring particles tend to align, however, they are disturbed by noise. In the following, we discuss the cases $w(n) = 1 = \text{const.}$ and $w(n) = 1/n$.

From a technical point of view, the case $w(n) = 1 = \text{const.}$ is the most desirable since in this case, the model includes only pair interactions. This has the advantage that, as in the regular BBGKY-hierarchy, three-particle correlations can be produced only from previously existing pair correlations, four-particle correlations can only be produced from three-particle correlations and so on. In contrast, if w is a function of n as, e.g., $w(n) = 1/n$ the interactions are in fact N -particle interactions. That means, all orders of correlations are produced immediately even if the model evolves from uncorrelated initial conditions.

Intuitively, one might think that such tiny details of the model are not that important and might lead to qualitatively equivalent results. The average number of neighbors is a constant anyway and one might hope that the fluctuations of the number of neighbors do not play a major role. Surprisingly, that is not the case and the two models differ qualitatively; see also [39] for a detailed analysis. For $w(n) = 1 = \text{const.}$ one still finds a homogeneous phase at large noise and a polarly ordered phase at smaller noise for systems of finite size. However, in the case of polar order, particles do not arrange in high density bands or cross sea patterns but they form high density clusters that contain almost all particles; see Fig. 1.

It is not even clear if there is a disordered phase at all in the thermodynamic limit. An alternative hypothesis is as follows. For every finite but possibly large noise strength one finds polar ordered clusters of very high densities for large enough systems. Due to the additive nature of the alignment interactions, such a cluster can remain stable even for large noise strength if the density is large enough. It could be that such a clustered state is the steady state for large systems at any noise strength. However, sophisticated investigations are necessary to answer this question. If one could show the existence of a disordered phase in the thermodynamic limit, and if one would be interested only in the behavior of the disordered phase (and not in the flocking transition), then one could also study the simpler model with additive interactions. Differences between additive and nonadditive interactions have been noticed already in Ref. [40], they have been studied in more detail in a very recent work [39]. Although the model with additive interactions is very interesting as well, we are mainly

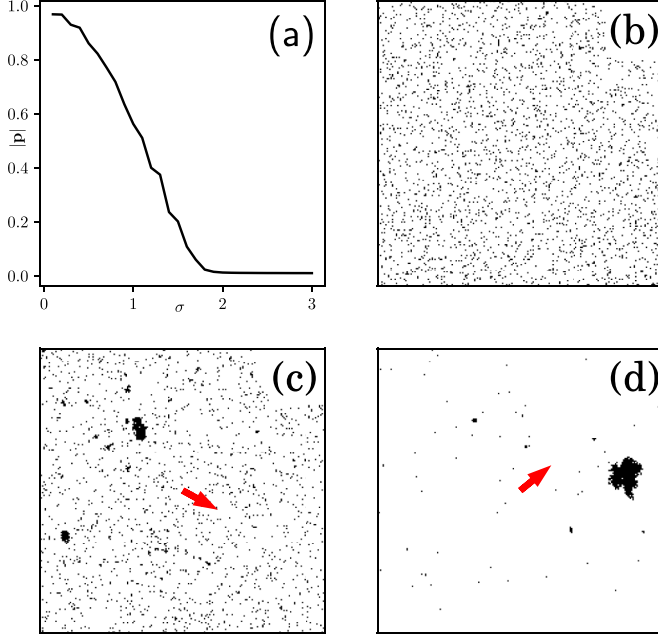


FIG. 1. Numerical results for the system (1) with $w(n) = 1 = \text{const.}$, simulated with an Euler-Maruyama-scheme [41] with $\Delta t = 0.01$. Parameters are $N = 10^4$, $L \approx 177.25$, $v = R = 1$, thus $C_1 := \rho_0 \pi R^2 = 1$. (a) Polar order parameter as a function of noise strength. (b–d) Snapshots after a thermalization time $T = 1000$ in the disordered phase (b) $\sigma = 2.2$, and in the polarly ordered phase (c) $\sigma = 1.5$ and (d) $\sigma = 0.3$.

interested in the study of models behaving qualitatively like the Vicsek model. Hence, we focus on the case

$$w(n) = 1/n \quad (4)$$

in this paper to reach quantitative agreement with agent-based simulation. However, throughout the paper, the theory is formulated for arbitrary weight functions $w(n)$.

III. KINETIC THEORY

Instead of investigating the Langevin equations, Eq. (1), we can equivalently study the associated N -particle Fokker-Planck equation [42]

$$\begin{aligned} \partial_t P_N(1, 2, \dots, N, t) &= - \sum_{i=1}^N \partial_{\phi_i} \left[\sum_{j=1}^N w(|\Omega(i)|) \theta_{ij} \right. \\ &\quad \times \sin(\phi_j - \phi_i) P_N(1, 2, \dots, N, t) \\ &\quad - \sum_{i=1}^N v \cos(\phi_i) \partial_{x_i} P_N(1, 2, \dots, N, t) \\ &\quad - \sum_{i=1}^N v \sin(\phi_i) \partial_{y_i} P_N(1, 2, \dots, N, t) \\ &\quad \left. + \sum_{i=1}^N \frac{\sigma^2}{2} \partial_{\phi_i}^2 P_N(1, 2, \dots, N, t) \right] \end{aligned} \quad (5)$$

$$\theta_{ij} := \theta \left[R - \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \right], \quad (6)$$

where we used the abbreviation of writing just 1 instead of x_1, y_1, ϕ_1 and similar for 2, 3, ... The probability density of finding the system in a phase space volume around the coordinates given by $1, 2, \dots, N$ at time t is denoted by $P_N(1, 2, \dots, N, t)$. This equation is exact.

We assume that the probability density is initially and hence at all times symmetric with respect to the permutation of coordinates $i \leftrightarrow j$. We denote the marginalized probability density by

$$P_k(1, 2, \dots, k, t) := \int P_N(1, 2, \dots, N, t) d(k+1) \dots dN, \quad (7)$$

where $d\mathbf{l}$ denotes $dx_l dy_l d\phi_l$ and integration is performed over the interval $[0, L]$ for x_l and y_l and over the interval $[0, 2\pi)$ for ϕ_l .

Using the Fokker-Planck Eq. (5) and the symmetry of permutations of coordinates we obtain the time evolution of the marginalized density

$$\begin{aligned} \partial_t P_1(1, t) &= - \partial_{\phi_1} \left[(N-1) \int w(|\Omega(1)|) d2 \dots dN \theta_{12} \right. \\ &\quad \times \sin(\phi_2 - \phi_1) P_N(1, 2, \dots, N, t) \\ &\quad - v \cos(\phi_1) \partial_{x_1} P_1(1, t) - v \sin(\phi_1) \partial_{y_1} P_1(1, t) \\ &\quad \left. + \frac{\sigma^2}{2} \partial_{\phi_1}^2 P_1(1, t) \right] \end{aligned} \quad (8)$$

We observe that the time evolution of the one-particle probability depends on the complete N particle probability due to the fact that the number of neighbors $|\Omega(1)|$ of the first particle depends on the positions of all the other particles.

For spatially homogeneous solutions, P_1 depends only on ϕ and t , hence we can define

$$p_\phi(\phi_1, t) := L^2 P_1(1, t) \quad (9)$$

such that Eq. (8) simplifies to

$$\begin{aligned} \partial_t p_\phi(\phi_1, t) &= - \partial_{\phi_1} \left[(N-1) \int w(|\Omega(1)|) d1 \dots dN \theta_{12} \right. \\ &\quad \times \sin(\phi_2 - \phi_1) P_N(1, 2, \dots, N, t) \\ &\quad \left. + \frac{\sigma^2}{2} \partial_{\phi_1}^2 p_\phi(\phi_1, t) \right] \end{aligned} \quad (10)$$

Considering the noise strength σ as control parameter, in the thermodynamic limit $N \rightarrow \infty$, $\frac{N}{L^2} = \text{const.}$, the system exhibits a transition from disorder to polar order. In the disordered phase, all particles directions are distributed randomly. In the polar ordered phase there is a global preferred direction of motion.

In the next section we calculate the threshold noise strength where the transition takes place, within the mean-field theory.

A. Mean-field Theory

Solving Eq. (10) exactly is very difficult or even impossible. Therefore, we are looking for the stationary solution under the mean-field assumption

$$P_N(1, 2, \dots, N) = P_1(1)P_1(2) \dots P_1(N), \quad (11)$$

which simplifies with Eq. (9) to

$$P_N(1, 2, \dots, N) = \frac{1}{L^{2N}} p_\phi(\phi_1) p_\phi(\phi_2) \dots p_\phi(\phi_N). \quad (12)$$

To simplify the interaction term in Eq. (10) it is reasonable to split the spatial integration domain $\mathbb{R}^{2(N-1)}$ into pieces such that the number of neighbors of particle one is constant on each piece. Thus, for a fixed position of particle one, we label subsets of $\mathbb{R}^{2(N-1)}$ on which particle one has exactly n neighbors by n . This has the advantage that $w(n)$ is constant on each subset. The union (over n) of all those pieces restores the original spatial integration domain in Eq. (10). When there are no neighbors of particle one ($n = 0$) there is no contribution to the interaction term. Summarizing all nonzero interaction terms, Eq. (10) becomes

$$\begin{aligned} \partial_t p_\phi(\phi_1, t) = & -\partial_{\phi_1} \left[\frac{N-1}{L^{2(N-1)}} \sum_{n=1}^{N-1} \int w(n+1) d2 d\mathbf{r}_3 \dots d\mathbf{r}_N \theta_{12} \dots \theta_{1(n+1)} (1 - \theta_{1(n+2)}) \dots (1 - \theta_{1N}) \sin(\phi_2 - \phi_1) \right. \\ & \left. \times p_\phi(\phi_1, t) p_\phi(\phi_2, t) + \text{permutations} \right] + \frac{\sigma^2}{2} \partial_{\phi_1}^2 p_\phi(\phi_1, t), \end{aligned} \quad (13)$$

where $\mathbf{r}_i = (x_i, y_i)$. In the above integral $\mathbf{r}_2, \dots, \mathbf{r}_{n+1}$ are integrated over the inside of the disk with radius R around particle one and $\mathbf{r}_{n+2}, \dots, \mathbf{r}_N$ are integrated over the outside of the disk around particle one. That means n is the number of neighbors of particle one. By permutations we mean all other choices of $n-1$ particles from $\{3, \dots, N\}$ to be integrated over the inside. If we denote the indexes of particles chosen to be integrated over the inside by $\pi(3), \dots, \pi(n+1)$ and the indexes of the particles chosen to be integrated over the outside by $\pi(n+2), \dots, \pi(N)$, then the product of Heaviside functions reads $\theta_{12} \theta_{1\pi(3)} \dots \theta_{1\pi(n+1)} (1 - \theta_{1\pi(n+2)}) \dots (1 - \theta_{1\pi(N)})$. All those permutations π give the same contribution, hence we can take care of them by a combinatorial factor to arrive at

$$\begin{aligned} \partial_t p_\phi(\phi_1, t) = & -\partial_{\phi_1} \left[\frac{N-1}{L^{2(N-1)}} \sum_{n=1}^{N-1} \binom{N-2}{n-1} \int w(n+1) d2 d\mathbf{r}_3 \dots d\mathbf{r}_N \theta_{12} \dots \theta_{1(n+1)} (1 - \theta_{1(n+2)}) \dots (1 - \theta_{1N}) \right. \\ & \left. \times \sin(\phi_2 - \phi_1) p_\phi(\phi_1, t) p_\phi(\phi_2, t) \right] + \frac{\sigma^2}{2} \partial_{\phi_1}^2 p_\phi(\phi_1, t). \end{aligned} \quad (14)$$

Here, in the two-dimensional case, the C_1 coefficient is calculated as

$$C_1 := \frac{N}{L^2} \int \theta_{12} d\mathbf{r}_2, \quad (15)$$

that is the expectation value of the number of particles within a circle of radius R . Since we are interested in the thermodynamic limit $N \rightarrow \infty$ it suffices to approximate

$$(N-1) \binom{N-2}{n-1} \approx \frac{N^n}{(n-1)!}. \quad (16)$$

Inserting Eqs. (15) and (16) into Eq. (14) we obtain

$$\partial_t p_\phi(\phi_1, t) = - \sum_{n=1}^{N-1} \frac{C_1^n}{(n-1)!} \left(1 - \frac{C_1}{N}\right)^{N-1-n} w(n+1) \partial_{\phi_1} \left[\int d\phi_2 \sin(\phi_2 - \phi_1) p_\phi(\phi_1, t) p_\phi(\phi_2, t) \right] + \frac{\sigma^2}{2} \partial_{\phi_1}^2 p_\phi(\phi_1, t). \quad (17)$$

Taking the limit $N \rightarrow \infty$ and substituting $k = n - 1$ we arrive at

$$\partial_t p_\phi(\phi_1, t) = - C_1 \sum_{k=0}^{\infty} \frac{C_1^k}{k!} \exp(-C_1) w(k+2) \partial_{\phi_1} \left[\int d\phi_2 \sin(\phi_2 - \phi_1) p_\phi(\phi_1, t) p_\phi(\phi_2, t) \right] + \frac{\sigma^2}{2} \partial_{\phi_1}^2 p_\phi(\phi_1, t). \quad (18)$$

One way of treating Eq. (18) is to study its full Fourier transform. Since here we are only interested in the critical noise strength it suffices to consider only the zeroth and the first order in the Fourier transform. Rotating the coordinate system in an appropriate way we can use the ansatz

$$p_\phi(\phi) = \frac{1}{2\pi} + \varepsilon \cos(\phi). \quad (19)$$

The isotropic distribution corresponding to $\varepsilon = 0$ is always a solution of Eq. (18). At the critical noise strength it changes stability.

Inserting the ansatz Eq. (19) into Eq. (18), multiplying by $\cos(\phi_1)/\pi$ integrating over ϕ_1 and neglecting terms of order ε^2 we obtain by partial integration

$$\begin{aligned} \partial_t \varepsilon &= -\frac{1}{\pi} C_1 \sum_{k=0}^{\infty} \frac{C_1^k}{k!} \exp(-C_1) w(k+2) \left\{ \int_0^{2\pi} d\phi_1 d\phi_2 \sin(\phi_1) \sin(\phi_2 - \phi_1) \left[\frac{1}{4\pi^2} + \frac{\varepsilon \cos(\phi_1)}{2\pi} + \frac{\varepsilon \cos(\phi_2)}{2\pi} \right] \right\} - \frac{\sigma^2}{2} \varepsilon \\ &= \left[\frac{C_1}{2} \sum_{k=0}^{\infty} \frac{C_1^k}{k!} \exp(-C_1) w(k+2) - \frac{\sigma^2}{2} \right] \varepsilon. \end{aligned} \quad (20)$$

Hence, the mean-field critical noise strength is

$$\sigma_c = \sqrt{C_1 \sum_{k=0}^{\infty} \frac{C_1^k}{k!} \exp(-C_1) w(k+2)}. \quad (21)$$

In case of the constant weight function $w(n) = 1$ this reduces to

$$\sigma_c = \sqrt{C_1}. \quad (22)$$

For a Poisson distribution we have

$$\begin{aligned} \left\langle \frac{1}{k+1} \right\rangle &= \sum_{k=0}^{\infty} \frac{1}{k+1} \frac{C_1^k}{k!} \exp(-C_1) \\ &= \frac{1}{C_1} \sum_{k=0}^{\infty} \frac{C_1^{k+1}}{(k+1)!} \exp(-C_1) \\ &= \frac{1}{C_1} \sum_{\bar{k}=1}^{\infty} \frac{C_1^{\bar{k}}}{\bar{k}!} \exp(-C_1) \\ &= \frac{1}{C_1} \left[\sum_{\bar{k}=0}^{\infty} \frac{C_1^{\bar{k}}}{\bar{k}!} \exp(-C_1) - \exp(-C_1) \right] \\ &= \frac{1}{C_1} [1 - \exp(-C_1)] \end{aligned} \quad (23)$$

and similar

$$\begin{aligned} \left\langle \frac{1}{k+2} \right\rangle &= \left\langle \frac{1}{k+1} \right\rangle - \left\langle \frac{1}{(k+1)(k+2)} \right\rangle \\ &= \left\langle \frac{1}{k+1} \right\rangle - \frac{1}{C_1^2} \sum_{\bar{k}=2}^{\infty} \frac{C_1^{\bar{k}}}{\bar{k}!} \exp(-C_1) \\ &= \left\langle \frac{1}{k+1} \right\rangle - \frac{1}{C_1^2} [1 - (1+C_1)\exp(-C_1)]. \end{aligned} \quad (24)$$

Hence, for the weight function $w(n) = 1/n$ we obtain the critical noise strength using Eqs. (21), (23), and (24):

$$\begin{aligned} \sigma_c &= \sqrt{\left[1 - \exp(-C_1) - \frac{1}{C_1} [1 - (1+C_1)\exp(-C_1)] \right]} \\ &= \sqrt{1 - \frac{1}{C_1} [1 - \exp(-C_1)]}. \end{aligned} \quad (25)$$

IV. CORRELATIONS

In general, the mean-field assumption Eq. (11) is not valid but must be corrected by terms containing correlations of various order. The correlation functions G_k are defined recursively

by [31,43,44]

$$G_1 \equiv P_1, \quad (26)$$

$$G_k(1, \dots, k)$$

$$\begin{aligned} &:= P_k(1, \dots, k) - \sum_{\sigma} \sum_{l=1}^{k-1} \frac{1}{(l-1)!} \frac{1}{(k-l)!} \\ &\quad \times G_l[1, \sigma(2), \dots, \sigma(l)] P_{k-l}[\sigma(l+1), \dots, \sigma(k)], \end{aligned} \quad (27)$$

where \sum_{σ} denotes the sum over all permutations of the elements $\{2, \dots, k\}$. We can rewrite Eq. (27) as $P_l(\dots) = G_l(\dots) + \dots$ and insert it recursively for all P_l of order $l < k$ into Eq. (27) and eventually replace P_1 by G_1 . Performing such an expansion, only P_k and G -functions remain on the right-hand side of Eq. (27). It follows inductively from Eq. (27) that the indexes of all correlation functions on the right-hand side are ordered, that is for each term $G_l(i_1, \dots, i_l)$ appearing in the expansion of Eq. (27) it holds $i_1 < i_2 < \dots < i_l$. Thus, instead of Eq. (27) we might alternatively write

$$\begin{aligned} G_k(1, \dots, k) &= P_k(1, \dots, k) - \sum \{ \text{over all possible} \\ &\quad \text{products of } G\text{-functions such that each of the arguments} \\ &\quad 1, \dots, k \text{ appears exactly once and for each } G\text{-function} \\ &\quad \text{the arguments are ordered} \}. \end{aligned} \quad (28)$$

For example, the two-, three-, and four-particle correlation functions are given explicitly as

$$G_2(1, 2) = P_2(1, 2) - P_1(1)P_1(2) \quad (29)$$

$$\begin{aligned} G_3(1, 2, 3) &= P_3(1, 2, 3) - P_1(1)P_1(2)P_1(3) \\ &\quad - P_1(1)G_2(2, 3) - P_1(2)G_2(1, 3) \\ &\quad - P_1(3)G_2(1, 2), \end{aligned} \quad (30)$$

$$\begin{aligned} G_4(1, 2, 3, 4) &= P_4(1, 2, 3, 4) - P_1(1)P_1(2)P_1(3)P_1(4) \\ &\quad - G_2(1, 2)P_1(3)P_1(4) - G_2(1, 3)P_1(2)P_1(4) \\ &\quad - G_2(1, 4)P_1(2)P_1(3) - G_2(2, 3)P_1(1)P_1(4) \\ &\quad - G_2(2, 4)P_1(1)P_1(3) - G_2(3, 4)P_1(1)P_1(2) \\ &\quad - G_2(1, 2)G_2(3, 4) - G_2(1, 3)G_2(2, 4) \\ &\quad - G_2(1, 4)G_2(2, 3) - G_3(1, 2, 3)P_1(4) \\ &\quad - G_3(1, 2, 4)P_1(3) - G_3(1, 3, 4)P_1(2) \\ &\quad - G_3(2, 3, 4)P_1(1). \end{aligned} \quad (31)$$

We want to mention some important properties of the correlation functions G_k . Since we assume that the N -particle

probability distribution is symmetric with respect to particle exchanges, also the correlation functions follow the same symmetry. Furthermore, for $k \geq 2$ it holds

$$\int G_k(1, \dots, k) dk = 0, \quad (32)$$

where the integration is performed over the whole space and all possible orientations of particle k . This property follows by induction over k from Eq. (27). It shows that G_k contains only information about k -particle correlations and not about lower order correlations. As a simple consequence it follows that

$$\int_X G_k(1, \dots, k) dk = - \int_{X^C} G_k(1, \dots, k) dk, \quad (33)$$

where X is some subset of the set of all possible configurations of particle k and X^C is the complement of X .

V. RING-KINETIC THEORY

The mean-field assumption Eq. (11) is equivalent to vanishing correlation functions $G_i \equiv 0$ for $i \geq 2$. In this paper we go one step further and take two-particle correlations into account. However, we still assume that higher-order correlations

can be neglected. That is, we consider

$$\begin{aligned} P_1(1, t), \quad G_2(1, 2, t) &:= P_2(1, 2, t) - P_1(1, t)P_1(2, t), \\ G_i(1, \dots, i, t) &\equiv 0 \quad \text{for } i \geq 3. \end{aligned} \quad (34)$$

Similar assumptions have been made in Ref. [26] but also in a very recent kinetic Landau theory [29] on the model with additive interactions, that is with $w(n) \equiv 1$. Therein, it was assumed additionally that G_2 is small and furthermore that also noise and coupling are small. It should be mentioned that according to [30] and to the results presented in this work, it follows that in the vicinity of the flocking transition and even in a considerable part of the disordered phase, the assumption of negligible G_3 and small G_2 is not justifiable. However, far in the disordered regime it is a reasonable approximation. In Sec. VIII we present a detailed comparison of the Landau theory of [29] applied to the model of nonadditive interactions with our full ring-kinetic theory.

In this paper we impose no restrictions on G_2 , that means we allow large pair correlations. As we assume spatial homogeneity, the one-particle probability density P_1 is independent on the position (x_1, y_1) and it suffices to consider the angular dependence p_ϕ given by Eq. (9). Similarly, the two-particle probability density P_2 depends only on the angles ϕ_1, ϕ_2 and on the distance between the spatial coordinates $\Delta_x := x_2 - x_1, \Delta_y := y_2 - y_1, \Delta = (\Delta_x, \Delta_y)$. Hence, it is reasonable to introduce the reduced probability density

$$\begin{aligned} p_2(\phi_1, \phi_2, \Delta) &:= L^2 \int P_2(\phi_1, \phi_2, x_1, x_2, y_1, y_2) \delta[\Delta_x - (x_2 - x_1)] \delta[\Delta_y - (y_2 - y_1)] dx_1 dx_2 dy_1 dy_2 \\ &= L^4 P_2(\phi_1, \phi_2, x_1, x_2 = x_1 + \Delta_x, y_1, y_2 = y_1 + \Delta_y) \end{aligned} \quad (35)$$

and correlation functions

$$\begin{aligned} g_2(\phi_1, \phi_2, \Delta) &:= L^2 \int G_2(\phi_1, \phi_2, x_1, x_2, y_1, y_2) \delta[\Delta_x - (x_2 - x_1)] \delta[\Delta_y - (y_2 - y_1)] dx_1 dx_2 dy_1 dy_2 \\ &= L^4 G_2(\phi_1, \phi_2, x_1, x_2 = x_1 + \Delta_x, y_1, y_2 = y_1 + \Delta_y). \end{aligned} \quad (36)$$

We obtain the time evolution equations of the reduced one- and two-particle probability density functions, that are the first two equations of the BBGKY-hierarchy, from the Fokker-Planck Eq. (5) using the marginalization Eq. (7), plugging in the cluster expansion Eq. (27) and using the ring-kinetic ansatz Eq. (34) as well as properties Eqs. (32) and (33) of the correlation functions. For example, the derivation of the time evolution of the single particle distribution can be explained in four steps. (i) In Eq. (5), particle one interacts with particles $2, \dots, N$. After integrating over the degrees of freedom of particles $2, \dots, N$ all these interaction terms are equal due to the symmetry of P_N under permutation of the particles. Thus, it suffices to consider interactions between particle one and another arbitrarily chosen particle when the interaction term is multiplied by $N - 1$. Here, we choose particle three as interaction partner. (ii) Plugging in the cluster expansion Eq. (27) for P_N , we have a sum of products of P_1 's and G_2 's. We can sort the summands by the terms depending on particle one and on particle three. These are the terms that are involved in the interaction. For example, one type of terms

is $P_1(1)P_1(3)$ multiplied with arbitrary combinations of P_1 's and G_2 's. Another example is $P_1(1)G_2(3, 7)$, etc. (iii) For each of the terms classified in step (ii) we split the spatial integration of the involved particles in an integral over the interior of the circle around particle one and an integral over the exterior of a circle around particle one. However, particle three must be integrated over the interior of the circle only, because particles one and three are interacting according to step (i). Note that for particles that appear as arguments of G_2 the value of the integral over the interior of the circle is exactly minus the value of the integral over the exterior of the circle due to Eq. (33). The classes of terms specified in step (iii) correspond to the diagrams in the following equation. (iv) In a last step we consider the integration over all particles that are not directly involved in the interaction, that means, that are not the arguments of the terms specified in step (ii). Depending on whether these particles are in the circle around particle one or not, they effect the weight function $w(n)$. Otherwise, they do not take part in the interaction and they are not correlated (in the considered terms) with the particles involved in the inter-

action. Thus, the average over the remaining particles results in an expectation value of the weight function $w(n)$ that can be calculated explicitly by methods of Ref. [30] in the limit $N \rightarrow \infty$. We use the following diagrammatic representation of the appearing collision integrals

$$1 \bullet \longrightarrow 2 = \sin(\phi_2 - \phi_1)\theta(R - |\mathbf{r}_2 - \mathbf{r}_1|), \quad (37)$$

$$1 \bullet \text{---} \bullet 2 = \theta(R - |\mathbf{r}_2 - \mathbf{r}_1|), \quad (38)$$

$$1 \bullet \text{---} \text{---} \bullet 2 = 1 - \theta(R - |\mathbf{r}_2 - \mathbf{r}_1|), \quad (39)$$

$$1 \bullet \text{---} \text{---} \text{---} \bullet 2 = g_2(1, 2), \quad (40)$$

$$\bullet k = p(\phi_k), \text{ if } k \text{ is not connected via a coiled line,} \quad (41)$$

$$1 \odot = \partial_{\phi_1}. \quad (42)$$

All terms representing a symbol in a diagram are meant to be multiplied and integrated over all degrees of freedom that do not occur on the left-hand side of the equation in which the diagram appears. As an example, we give the meaning of the following integral:

$$1 \odot \text{---} \text{---} \bullet 2 = \int \partial_{\phi_1} g_2(\phi_1, \phi_2, \Delta) \sin(\phi_2 - \phi_1) \times \theta(R - |\Delta|) d\phi_2 d\Delta. \quad (43)$$

For the one-particle angular distribution we find within our diagrammatic notation

$$\begin{aligned} \partial_t p(\phi_1) = & -\rho w_2 \left[1 \odot \text{---} \text{---} \bullet 3 + 1 \odot \longrightarrow \bullet 3 \right] \\ & - \rho^2 (w_3 - w_2) \left[1 \odot \begin{array}{c} \bullet 3 \\ \text{---} \\ \bullet 4 \end{array} + 1 \odot \begin{array}{c} \bullet 3 \\ \text{---} \\ \bullet 4 \end{array} \right] \\ & - \rho^3 (w_4 - 2w_3 + w_2) 1 \odot \begin{array}{c} \bullet 3 \\ \text{---} \\ \bullet 5 \\ \text{---} \\ \bullet 4 \end{array} + \frac{\sigma^2}{2} \partial_{\phi_1}^2 p(\phi_1). \end{aligned} \quad (44)$$

See Eq. (A1) in Appendix A for Eq. (44) without diagrammatic notation. Here, $\rho := N/L^2$ denotes the average particle density and w_k the expectation value of the weight function for $k + l$ neighbors with respect to the distribution of the number of particles found within a circle of radius R , $q(l)$, that is

$$w_k := \langle w(k + l) \rangle = \sum_{l=0}^{\infty} w(k + l) q(l). \quad (45)$$

The distribution of the number of correlated particles that are placed within a circle of radius R was recently derived in Ref. [30]. In case that only two-particle correlations are present this distribution depends only on parameter

$$C_2 := N^2 \int G_2(1, 2) \theta_{12} d1 d2, \quad (46)$$

where

$$\theta_i := \theta(R - \sqrt{x_i^2 + y_i^2}), \quad (47)$$

with the Heaviside function θ . A similar quantity that effects the number of neighbor distribution [30] is

$$D_2 := N \int G_2(1, 2) \theta_{12} d1 d2, \quad (48)$$

where θ_{ij} is given by Eq. (6) The distribution of the number of particles within a circle of radius R is given in Ref. [30] as

$$\begin{aligned} q(l) = & (C_1 - C_2)^l \exp(C_2/2 - C_1) \\ & \times \sum_{k=0}^{\infty} \left[\frac{C_2}{2(C_1 - C_2)^2} \right]^k \frac{1}{k!(l - 2k)!}. \end{aligned} \quad (49)$$

As before, $C_1 = \frac{N}{L^2} \pi R^2$ is just the average number of neighbors. Alternatively to the representation by an infinite sum Eq. (49) one can give a simple expression for the characteristic function of $q(l)$ [30]

$$\chi(u) = \exp \left[\sum_{l=1}^2 \sum_{t=0}^l (-1)^{l+t} \frac{C_l}{l!} \binom{l}{t} \exp(itu) \right]. \quad (50)$$

Coming back to Eq. (44), we give a guide to explain all interaction integrals. Those are all possible diagrams where each particle is connected to 1 via a solid line or an arrow (that means all particles are neighbors of particle 1) such that every point is connected to 1 also via a path consisting only of arrows and/or coils and there is exactly one arrow that starts from 1. Particles that are not connected to 1 via a path of arrows and/or coils do not directly take part in the interactions and thus do not appear in the diagrams. Because there is only one arrow in the diagrams and each point can be connected to only one coil (that represents g_2) the diagrams can have no more than four points. Each diagram has a prefactor of $-\rho^{k-1}$, where k is the number of particles involved in the diagram. This prefactor is a combination of a combinatorial factor and $1/L^{2k}$ from Eq. (9) or Eq. (36). Furthermore, there is another prefactor of w_k that takes into account the expectation value of the weight function when an average over all particles not involved in the diagram is performed. In principle, we would have additional diagrams where points not connected via an arrow are connected to 1 via dashed lines, that means those particles are not a neighbor of particle 1. However, we can replace each dashed line by a solid line and multiply with -1 to compensate it due to property Eq. (33). However, for these additional diagrams we have different prefactors of w_{k-1} or w_{k-2} if one or two dashed lines are involved, respectively, due to the different number of neighbors of particle 1. Considering all the aforementioned interaction terms, angular diffusion and streaming we arrive at Eq. (44). Note that for pair interactions, $w(n) = \text{const.}$, we have $w_2 = w_3 = w_4$ and only the first two interaction integrals remain.

Analogously to Eq. (44), we find starting from Eq. (5), the time evolution equation for the two-particle probability

distribution

$$\begin{aligned}
 \partial_t p_2(\phi_1, \phi_2, \Delta) = & -w_2 \left[\text{diagram 1} + \text{diagram 2} \right] - \rho(w_3 - w_2) \left[\text{diagram 3} + \text{diagram 4} \right] \\
 & - \rho^2(w_4 - 2w_3 + w_2) \left[\text{diagram 5} \right] - \rho w_3 \left[\text{diagram 6} + \text{diagram 7} \right] \\
 & - \rho^2(w_4 - w_3) \left[\text{diagram 8} + \text{diagram 9} + \text{diagram 10} + \text{diagram 11} + \text{diagram 12} + \text{diagram 13} \right] \\
 & - \rho^3(w_5 - 2w_4 + w_3) \left[\text{diagram 14} + \text{diagram 15} + \text{diagram 16} \right] - \rho^4(w_6 - 3w_5 + 3w_4 - w_3) \left[\text{diagram 17} \right] \\
 & - \rho w_2 \left[\text{diagram 18} + \text{diagram 19} + \text{diagram 20} + \text{diagram 21} \right] \\
 & - \rho^2(w_3 - w_2) \left[\text{diagram 22} + \text{diagram 23} + \text{diagram 24} + \text{diagram 25} + \text{diagram 26} + \text{diagram 27} \right] \\
 & - \rho^3(w_4 - 2w_3 + w_2) \left[\text{diagram 28} + \text{diagram 29} + \text{diagram 30} \right] - \rho^4(w_5 - 3w_4 + 3w_3 - w_2) \left[\text{diagram 31} \right] \\
 & + \frac{\sigma^2}{2} \partial_{\phi_1}^2 [p(\phi_1)p(\phi_2) + g_2(\phi_1, \phi_2, \Delta)] + 1 \leftrightarrow 2 \\
 & + v(\cos \phi_1 - \cos \phi_2) \partial_{\Delta} g_2(\phi_1, \phi_2, \Delta) + v(\sin \phi_1 - \sin \phi_2) \partial_{\Delta} g_2(\phi_1, \phi_2, \Delta),
 \end{aligned} \tag{51}$$

where $1 \leftrightarrow 2$ is an abbreviation for all previous terms with particles one and two interchanged.

Employing Eq. (29), in our notation Eqs. (9), (35), and (36),

$$g_2(\phi_1, \phi_2, \Delta) = p_2(\phi_1, \phi_2, \Delta) - L^2 p_{\phi}(\phi_1)p_{\phi}(\phi_2), \tag{52}$$

we obtain the time evolution equation of the pair correlation function

$$\begin{aligned}
 \partial_t g_2(\phi_1, \phi_2, \Delta) = & \partial_t p_2(\phi_1, \phi_2, \Delta) - L^2 p_{\phi}(\phi_1) \partial_t p_{\phi}(\phi_2) \\
 & - L^2 p_{\phi}(\phi_2) \partial_t p_{\phi}(\phi_1).
 \end{aligned} \tag{53}$$

Inserting Eqs. (44) and (51) yields

$$\begin{aligned}
 \partial_t g_2(\phi_1, \phi_2, \Delta) = & -w_2 \left[\text{diagram 1} + \text{diagram 2} \right] - \rho(w_3 - w_2) \left[\text{diagram 3} + \text{diagram 4} \right] \\
 & - \rho^2(w_4 - 2w_3 + w_2) \left[\text{diagram 5} \right] - \rho w_3 \left[\text{diagram 6} + \text{diagram 7} \right] \\
 & - \rho(w_3 - w_2) \left[\text{diagram 8} + \text{diagram 9} \right] - \rho^2(w_4 - w_3) \left[\text{diagram 10} + \text{diagram 11} + \text{diagram 12} + \text{diagram 13} \right] \\
 & - \rho^2(w_4 - 2w_3 + w_2) \left[\text{diagram 14} + \text{diagram 15} \right] - \rho^3(w_5 - 2w_4 + w_3) \left[\text{diagram 16} + \text{diagram 17} \right] \\
 & - \rho^3(w_5 - 3w_4 + 3w_3 - w_2) \left[\text{diagram 18} \right] - \rho^4(w_6 - 3w_5 + 3w_4 - w_3) \left[\text{diagram 19} \right]
 \end{aligned}$$

$$\begin{aligned}
 & -\rho w_2 \left[1 \text{---} \textcircled{1} \text{---} \textcircled{2} \text{---} \textcircled{3} + 1 \text{---} \textcircled{1} \text{---} \textcircled{2} \text{---} \textcircled{3} \right] - \rho^2 (w_3 - w_2) \left[1 \text{---} \textcircled{1} \text{---} \textcircled{2} \text{---} \textcircled{3} \text{---} \textcircled{4} + 1 \text{---} \textcircled{1} \text{---} \textcircled{2} \text{---} \textcircled{3} \text{---} \textcircled{4} + 1 \text{---} \textcircled{1} \text{---} \textcircled{2} \text{---} \textcircled{3} \text{---} \textcircled{4} + 1 \text{---} \textcircled{1} \text{---} \textcircled{2} \text{---} \textcircled{3} \text{---} \textcircled{4} \right] \\
 & - \rho^3 (w_4 - 2w_3 + w_2) \left[1 \text{---} \textcircled{1} \text{---} \textcircled{2} \text{---} \textcircled{3} \text{---} \textcircled{4} \text{---} \textcircled{5} + 1 \text{---} \textcircled{1} \text{---} \textcircled{2} \text{---} \textcircled{3} \text{---} \textcircled{4} \text{---} \textcircled{5} \right] - \rho^4 (w_5 - 3w_4 + 3w_3 - w_2) \left[1 \text{---} \textcircled{1} \text{---} \textcircled{2} \text{---} \textcircled{3} \text{---} \textcircled{4} \text{---} \textcircled{5} \text{---} \textcircled{6} \right] \\
 & + \frac{\sigma^2}{2} \partial_{\phi_1}^2 g_2(\phi_1, \phi_2, \Delta) + 1 \leftrightarrow 2 \\
 & + v(\cos \phi_1 - \cos \phi_2) \partial_{\Delta} g_2(\phi_1, \phi_2, \Delta) + v(\sin \phi_1 - \sin \phi_2) \partial_{\Delta} g_2(\phi_1, \phi_2, \Delta).
 \end{aligned} \tag{54}$$

See Eq. (A2) in Appendix A for Eq. without diagrammatic notation. Note that, for example, the multiplication

$$\begin{aligned}
 [\partial_t p_\phi(\phi_1)] \times p_\phi(\phi_2) &= [\partial_t p_\phi(\phi_1)] \text{---} \textcircled{1} \text{---} \bullet 2 \\
 &+ [\partial_t p_\phi(\phi_1)] \text{---} \text{---} \text{---} \bullet 2
 \end{aligned} \tag{55}$$

can be done symbolically. A number of collision integrals in $\partial_t p_2$ and $-L^2 p_\phi(\phi_1) \partial_t p_\phi(\phi_2)$ or $-L^2 p_\phi(\phi_2) \partial_t p_\phi(\phi_1)$ cancel

such as for example $1 \text{---} \textcircled{1} \text{---} \textcircled{2} \text{---} \textcircled{3}$. However, their counterparts with particles one and two being neighbors, such as for

example $1 \text{---} \textcircled{1} \text{---} \textcircled{2} \text{---} \textcircled{3}$ do not cancel. This is an effect of the N -particle interactions. As a result these terms appear with different sign but also with different prefactors and thus they do not cancel. This effect is responsible for a number of terms that are not present for pair interactions, $w(n) = \text{const}$.

VI. COMPARISON OF RING-KINETIC THEORY AND AGENT-BASED SIMULATIONS

Recently, a quantitative numerical method to predict the minimal required order of correlations has been introduced in Ref. [30]. It is based on the measurement of the number of neighbors distribution (neighbors are particles that are closer than R) of a randomly selected particle. Besides this measurement, the number of neighbors distribution is also calculated under the assumption that only correlations up to a given order l_{max} are present, that is $G_l \equiv 0$ for all $l > l_{\text{max}}$. If both distributions agree reasonably well, then the correlation order l_{max} is considered to be sufficient; see Ref. [30] for details.

In Fig. 2 we show a correlation map that we obtained by this method for large systems of $N = 10^4$ particles. It shows, depending on parameters, which order of correlations is required. We expect excellent quantitative predictions of the ring-kinetic theory in a parameter regime where pair correlations are sufficient. Strictly speaking, this is the case in the disordered phase only. However, we find that a reasonable ring-kinetic description of the system is still possible if the influence of higher-order correlations is already measurable but still not dominant.

We solve the time evolution equations for p_ϕ and g_2 , Eqs. (44) and , numerically in Fourier space; see Appendix B. As a measurable quantity we consider the standard radial

distribution function $g(r)$ that is defined by

$$g(r) = \frac{L^2}{N(N-1)} \sum_{i \neq j} \frac{1}{2\pi r} \langle \delta(r - |\mathbf{r}_i - \mathbf{r}_j|) \rangle. \tag{56}$$

It is related to the correlation function g_2 given in Eq. (36) by

$$\begin{aligned}
 g(r) &= 1 + \int_0^{2\pi} d\phi_1 \int_0^{2\pi} d\phi_2 \int_0^L d\Delta_x \int_0^L d\Delta_y \\
 &\times g_2(\phi_1, \phi_2, \Delta) \frac{1}{2\pi r} \delta(r - |\Delta|)
 \end{aligned} \tag{57}$$

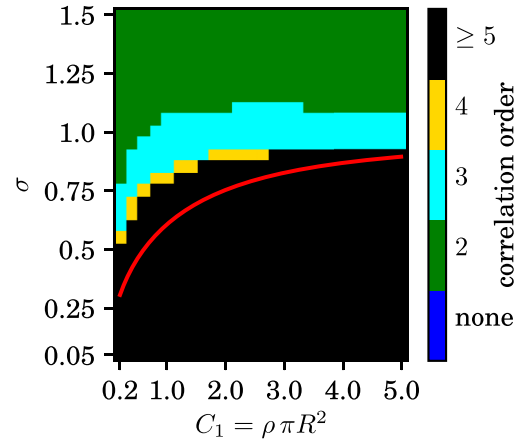


FIG. 2. Correlation map of the Vicsek-like model given by Eq. (1) with $w(n) = 1/n$. Colors encode the necessary order of correlations (e.g., cyan \sim three-particle correlations, etc.) for a quantitative description of the system. The correlation map was obtained by a recently introduced method [30] comparing the measured number of neighbor distribution to a theoretically predicted one. One considers the number of neighbor distribution to be acceptably close if the corresponding Kullback-Leibler divergence is less than 10^{-3} . We simulated 24 realizations for each noise strength $\sigma \in \{0.05, 0.1, 0.15, \dots, 1.5\}$ and each particle density $C_1 = \rho \pi R^2 \in \{0.2, 0.4, 0.6, \dots, 5.0\}$ with parameters $v = R = 1$, $N = 10^4$. For this system size, typically, the number of neighbor distribution has achieved its $N \rightarrow \infty$ limit at least in the disordered phase. The correlation parameters have been measured for a time interval of 10^4 after a thermalization time of 10^3 . Equation (1) was integrated using an Euler-Maruyama scheme with time step $\Delta t = 10^{-2}$. The analysis has been done as in Ref. [30]. The red curve shows the mean-field transition line given by Eq. (25).

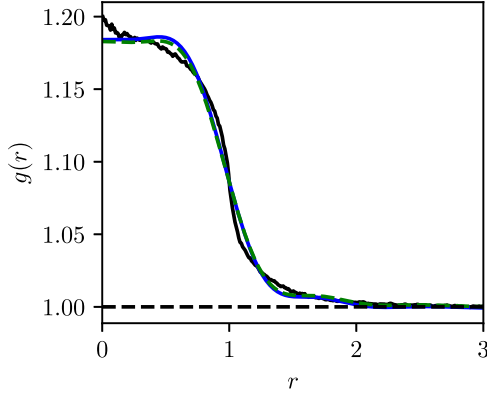


FIG. 3. Radial distribution function $g(r)$ obtained from ring-kinetic theory (blue solid line) and sampled directly from agent-based simulations according to Eq. (56) (black solid line). For the ring-kinetic theory we integrate the time evolution Eqs. (44) and in Fourier space (see Appendix B) with an Euler scheme with step size $\Delta t = 10^{-2}$ for an absolute time of 100, starting initially uncorrelated and disordered. For the agent-based simulation we integrate Eq. (1) with an Euler-Maruyama scheme with step size $\Delta t = 10^{-2}$. After a thermalization time of 10^3 we average over a time of 10^4 and over 24 realizations. Parameters are $\sigma = 1.5$, $v = R = 1$, $N = 509$, $C_1 = \rho\pi R^2 = 1$, $L \approx 40$. The ring-kinetic theory uses Fourier modes F_{klmn} according to Eq. (B2) with angular indexes $k, l \in \{-2, \dots, 2\}$ and spatial indexes $m, n \in \{-48, \dots, 48\}$. Performing a spatial Fourier transform of the measured curve (black line) with indexes from the same range $m, n \in \{-48, \dots, 48\}$ and transforming back into real space we obtain the dashed green line. The dashed solid line at $y = 1$ serves as a guide to the eyes.

and can be calculated analytically from the Fourier representation; see Appendix B.

In Fig. 3 we compare the radial distribution function measured in agent-based simulations with the results from ring-kinetic theory. Overall they agree very well, however, there are minimal deviations. Those deviations can be explained by the finite resolution used in the Fourier transform. We evaluated the ring-kinetic equations in Fourier space using a minimal spatial wave length of 0.83. Since the radial distribution functions $g(r)$ drops rapidly at about $r \approx 1$ it is not perfectly resolved, causing small deviations. To test this hypothesis, we Fourier transform the pair correlation function corresponding to the measured $g(r)$ with the same spatial resolution and Fourier transform it back into real space. The resulting curve deviates by less than one percent from the ring-kinetic result; see Fig. 3. The ring-kinetic theory predicts the full two-particle correlation function $g_2(\phi_1, \phi_2, \Delta)$ and not only the radial distribution function $g(r)$ that is an integral of it; see Eq. (57). Without spontaneous symmetry breaking, that is in the disordered phase, the system is isotropic. In that case, the pair correlation function depends only on three independent arguments and not on four, such as ϕ_1 , ϕ_2 , Δ_x , and Δ_y . We choose those three degrees of freedom as the length of the vector Δ , $r := |\Delta|$, the difference between the orientations of the two particles, $\Delta\phi := \phi_2 - \phi_1$, and the difference of the polar angle of the vector Δ and the orientation of the first particle $\alpha := \gamma - \phi_1$, where γ is the polar angle of the vector Δ ; see Fig. 4. Depending on those arguments we define the

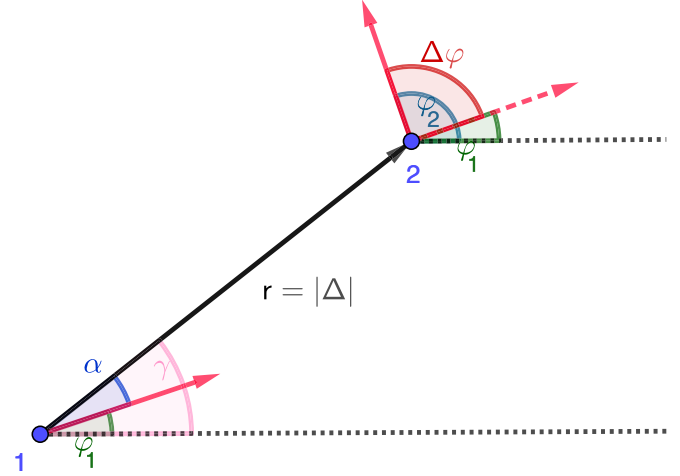


FIG. 4. Sketch of the three arguments r , α , and $\Delta\phi$ of the pair correlation function h for homogeneous and isotropic systems; see Eq. (58) for the definition of $h(r, \alpha, \Delta\phi)$. The blue points denote the positions of particles one and two. Their distance is denoted by r and their direction of motion (red arrows) is given by the angles ϕ_1 and ϕ_2 . The difference between those two angles is denoted by $\Delta\phi$. The difference between the polar angle γ of the vector pointing from particle one to particle two and ϕ_1 is denoted by α .

function

$$h(r, \alpha, \Delta\phi) := 2\pi \int_0^{2\pi} d\phi_1 \int_0^{2\pi} d\phi_2 \int_0^{2\pi} d\gamma \\ \times g_2(\phi_1, \phi_2, \Delta_x = r \cos \gamma, \Delta_y = r \sin \gamma) \\ \times \delta(\alpha - \gamma + \phi_1) \delta(\Delta\phi - \phi_2 + \phi_1). \quad (58)$$

For translational and rotational invariant systems it contains all information about two-particle correlations. It can be directly sampled from numerical data according to

$$h(r, \alpha, \Delta\phi) = \frac{L^2(2\pi)^2}{N(N-1)} \frac{1}{2\pi r} \sum_{i \neq j} \langle \delta(r - |\mathbf{r}_i - \mathbf{r}_j|) \\ \times \delta(\alpha - \gamma_{ij} + \phi_i) \delta(\Delta\phi - \phi_j + \phi_i) \rangle - 1. \quad (59)$$

A. Dependence of $h(r, \alpha, \Delta\phi)$ on $\Delta\phi$

We consider the dependence of the correlation function $h(r, \alpha, \Delta\phi)$ on the difference in orientation of the two particles $\Delta\phi$ for fixed r and α . In Fig. 5 we fix $r = 0.5$ and $\alpha = 0, \pi/2, \pi, 3\pi/2$. That means that the two particles are in the interior of each others interaction circle. We see that the correlations are highest for $\Delta\phi = 0$ showing that nearby particles align. For $\Delta\phi = \pi$ the correlations are even slightly negative. Rotations of one particle around the other seem to be of no particular importance at the considered distance as we see almost the same picture for different values of α . In Fig. 6 we fix $r = 1.0$. That means both particles are exactly at the boundary of each others interaction region. For $\alpha = 0$ and $\alpha = \pi$, that means if particle two is in the front or in the back of particle one (looking in the direction of motion of particle one), the distribution of $h(\Delta\phi)$ is still symmetric with a maximum at $\Delta\phi = 0$. If particle two is on the left or on

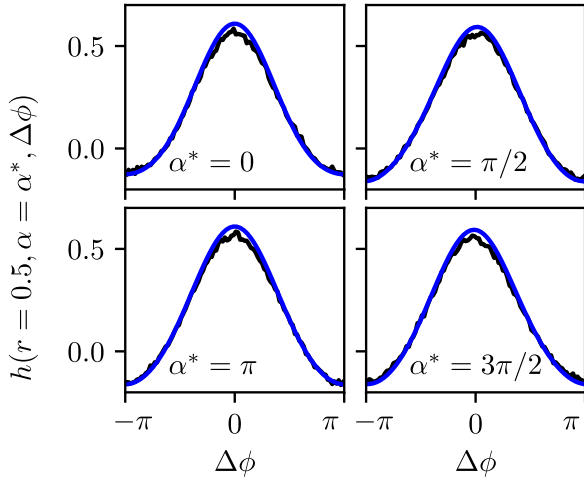


FIG. 5. Dependence of the correlation function $h(r, \alpha, \Delta\phi)$ on difference of velocity directions $\Delta\phi$ for fixed values of $r = 0.5$ and α (value given in the plot). Results of the ring-kinetic theory (blue line) are compared to direct measurements of agent-based simulations (black line). Parameters as described in the caption of Fig. 3.

the right of particle one, then it is slightly more likely that particle two points a bit outward the interaction region than to be perfectly aligned with particle one. This is reasonable due to the following argument. When particle two is placed at distance $r = 1$ at a given time, it is more likely that it was at $r < 1$ than that it was at $r > 1$ shortly before, because the pair correlations are much higher for $r < 1$.

In Fig. 7 we fix $r = 1.5$. Thus, the two particles are not directly interacting with each other. We see that the strength of the correlations is much smaller in this case. Also the fluctuations in the measurements of agent-based simulations are larger, because there are less events sampled. Quantitatively, the correlations are similar to the case of $r = 1.0$.

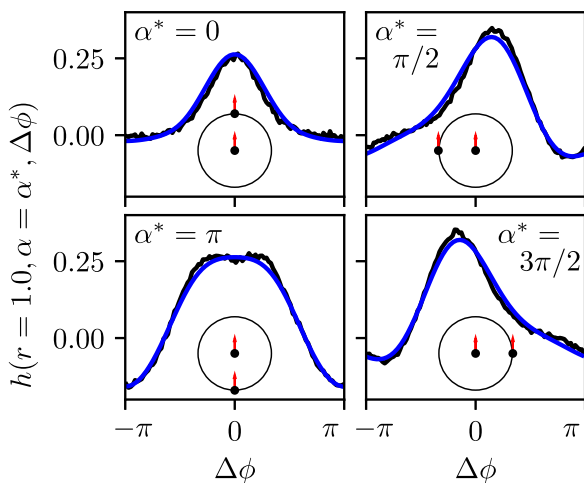


FIG. 6. Two-particle correlation function h as a function of the difference of the velocity directions of both particles as in Fig. 5, but here for $r = 1$. The black bullets sketch the relative positions of particle one (in the center of the circle) and particle two (on the circumference of the circle). The red arrows indicate the direction of the particle velocities at $\Delta\phi = 0$.

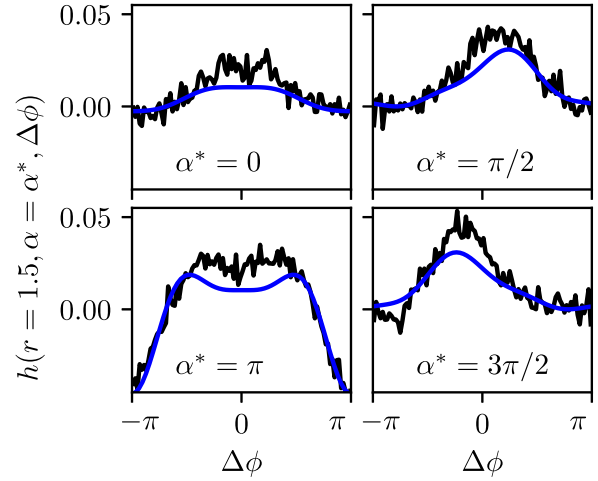


FIG. 7. Two-particle correlation function h as a function of the difference of the velocity directions of both particles as in Fig. 5, but here for $r = 1.5$.

B. Dependence of $h(r, \alpha, \Delta\phi)$ on α

In Fig. 8 we consider the α -dependence of the correlation function h for $r = 0.5$ and different values of $\Delta\phi$. As discussed in the previous subsection, there is almost no α -dependence. The value of h is largest for $\Delta\phi = 0$ and smallest and even negative for $\Delta\phi = \pi$ which is in agreement with Fig. 5.

In Fig. 9 we fix $r = 1$. For aligned particles, $\Delta\phi = 0$, the α -dependence of the correlation function is still very small with a slight preference of particle two being right or left of particle one compared to a placement in the front or back of particle one. For antialigned particles, $\Delta\phi = \pi$, it is unlikely that particle two is in the back of particle one. In that case, the particles would move away from each other. That means

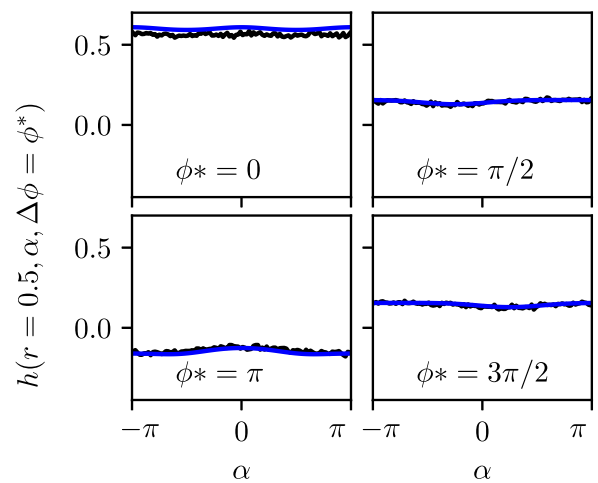


FIG. 8. Dependence of the correlation function $h(r, \alpha, \Delta\phi)$ on the angle α giving the rotation of particle two around particle with respect to the direction of motion of particle one. Fixed values of $r = 0.5$ and $\Delta\phi$ (value given in the plot) are considered. Results of the ring-kinetic theory (blue line) are compared to direct measurements of agent-based simulations (black line). Parameters as described in the caption of Fig. 3.

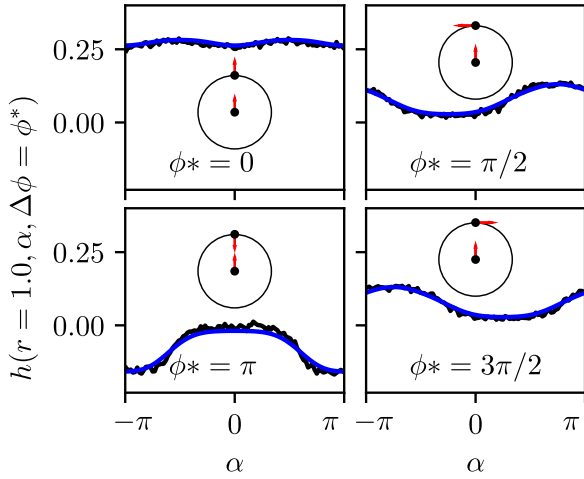


FIG. 9. Two-particle correlation function h as a function of the relative orientation α as in Fig. 8, but here for $r = 1$. The black bullets sketch the relative positions of particle one (in the center of the circle) and particle two (on the circumference of the circle) at $\alpha = 0$. The red arrows indicate the direction of the particle velocities.

that they have been closer to each other in the past which makes it unlikely that they are antialigned. For $\Delta\phi = \pi/2$ and $\Delta\phi = 3\pi/2$, there is a complex α -dependence with maximal correlations at $\alpha \approx 0.75\pi$ and $\alpha \approx -0.75\pi$, respectively. The value of α with minimal correlations is shifted by π with respect to the values of maximal correlations.

In Fig. 10 we fix $r = 1.5$. The correlation function behaves qualitatively similar to the case of $r = 1$, however here, the correlations are much smaller.

C. Dependence of $h(r, \alpha, \Delta\phi)$ on r

In this subsection we consider the r -dependence of the correlation function h for fixed values of α and $\Delta\phi$. We find the same qualitative behavior for all considered values of α ; see Figs. 11–14. For $\Delta\phi = 0, \pi/2$ and $3\pi/2$, for small r the function h shows a plateau of high correlations and decreases

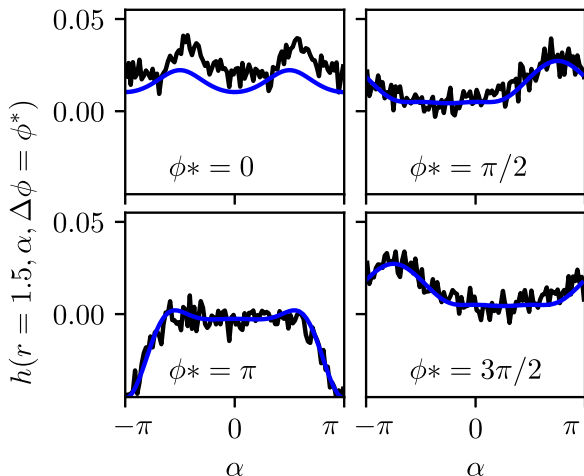


FIG. 10. Two-particle correlation function h as a function of the relative orientation α as in Fig. 8, but here for $r = 1.5$.

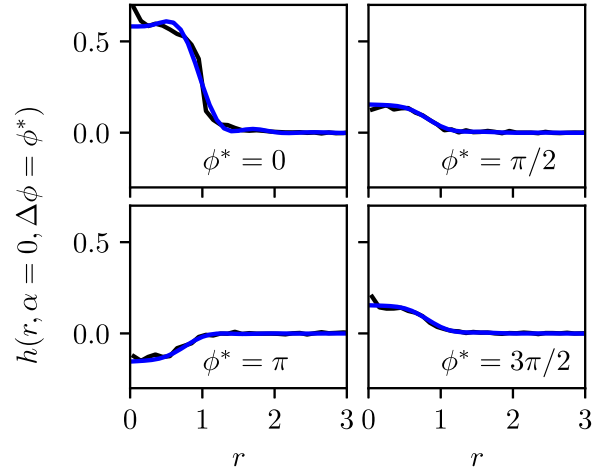


FIG. 11. Dependence of the correlation function $h(r, \alpha, \Delta\phi)$ on the distance r for fixed values of $\alpha = 0$ and $\Delta\phi$ (value given in the plot). Results of the ring-kinetic theory (blue line) are compared to direct measurements of agent-based simulations (black line). Parameters as described in the caption of Fig. 3.

relatively fast at about $r = 1$ toward zero. For $\Delta\phi = \pi$, the plateau for small r is negative and the decay toward zero at about $r = 1$ is still present. This shows that nearby antialigned particles are unlikely.

D. Applicability of the ring-kinetic theory

In the previous subsections we have seen that for the considered parameters, the ring-kinetic theory agrees very well with direct agent-based simulations. There are only very small deviations partially caused by the finite resolution in the Fourier transform in the numerical implementation of the ring-kinetic equations. In this subsection we study the applicability of the ring-kinetic theory depending on parameters. We have seen in the previous subsections that the spatial pair correlations decay rather rapidly for $r \lesssim R$. Therefore, we compare the integrated connected spatial pair correlations C_2 and D_2 defined in Eq. (46) between ring-kinetic theory and

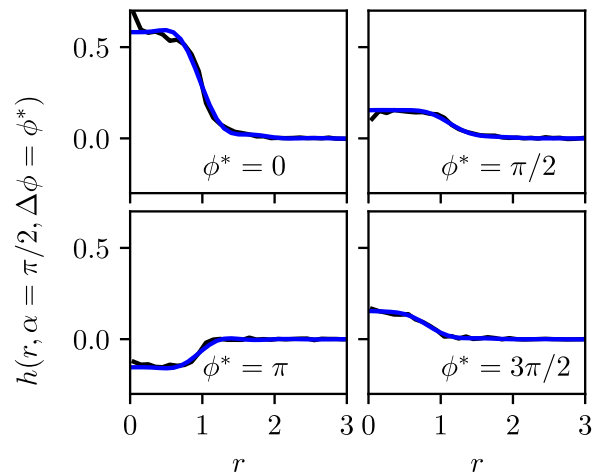


FIG. 12. Two-particle correlation function h as a function of distance r as in Fig. 11, but here for $\alpha = \pi/2$.

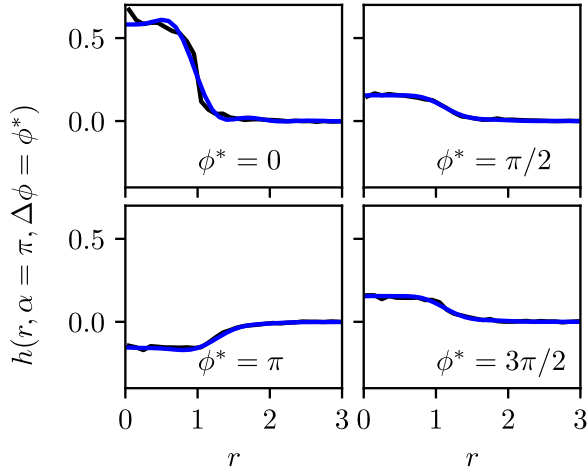


FIG. 13. Two-particle correlation function h as a function of distance r as in Fig. 11, but here for $\alpha = \pi$.

direct simulations in Fig. 15. We see very good agreement between ring-kinetic theory and simulations for noise strengths $\sigma \gtrsim 0.68$. For smaller noise strengths we find serious deviations.

For the considered parameter set, homogeneous mean-field theory, Eq. (25), predicts the onset of collective motion at $\sigma_c \approx 0.607$. In the direct simulations we determined the onset of flocking at $\sigma_{c,\text{sim}} = 0.53(1)$ from fluctuations of the polar order parameter $p = |\mathbf{p}| = \sqrt{p_x^2 + p_y^2}$, with $p_x = 1/N \sum_{i=1}^N \cos \phi_i$, $p_y = 1/N \sum_{i=1}^N \sin \phi_i$; see Fig. 16.

Returning to the deviations of the spatial correlations predicted by ring-kinetic theory in Fig. 15, we find that they are significant already before the onset of flocking. Nevertheless, the predicted correlations are not completely wrong, but have the correct order of magnitude. Therefore, we can still obtain the onset of collective motion using the ring-kinetic theory.

To analyze fluctuations or susceptibilities we would require to analyze stationary states also within the ordered phase. However, the solutions of the ring-kinetic equations require

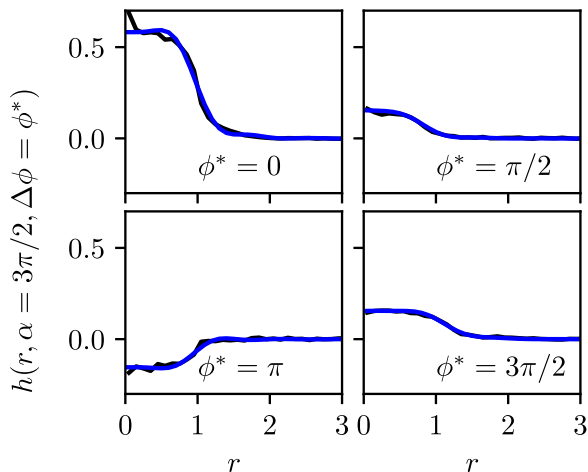


FIG. 14. Two-particle correlation function h as a function of distance r as in Fig. 11, but here for $\alpha = 3\pi/2$.

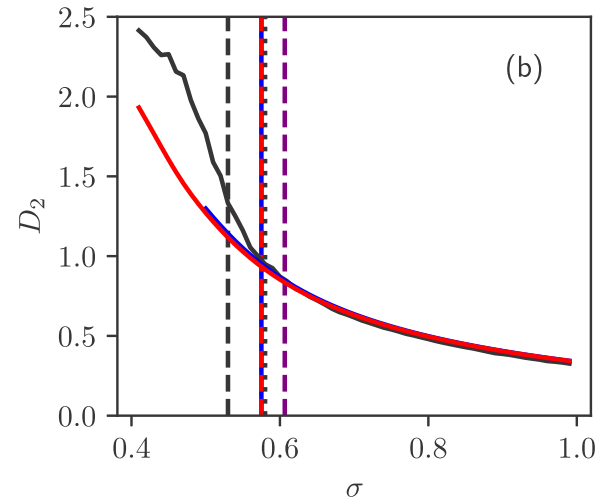
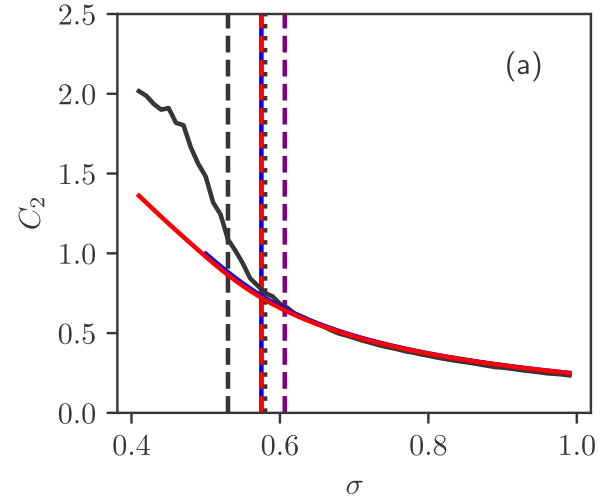


FIG. 15. Local spatial pair correlation parameters C_2 (a) and D_2 (b) as defined in Eq. (46) compared for direct simulations (black line), ring-kinetic theory (blue line), and a kinetic theory including a g_3 -closure (red line). The dashed vertical lines show the onset of flocking measured in agent-based simulations at $\sigma_{c,\text{sim}} = 0.53$ (black), in the ring-kinetic theory at $\sigma_{c,rk} = 0.575$ (blue), ring-kinetic theory with closure at $\sigma_{c,\text{closure}} = 0.575$ (red) and in mean-field theory at $\sigma_{c,\text{mf}} = 0.607$ (purple). System parameters are: $C_1 = \rho\pi R^2 = 1$, $R = v = 1$, $N = 509$. For comparison we give also the measured flocking transition in the thermodynamic limit $\sigma_{c,\text{sim},N \rightarrow \infty} = 0.58$ (dotted black line) obtained for $N = 10^4$ in Ref. [30].

a very long time to become stationary in the ordered regime (when started with only a minimal polar order). Therefore, we determine the onset of flocking within the ring-kinetic theory by a numerical stability analysis of the disordered state. In the numerical solution of the ring-kinetic equations we start with some (very small) initial polar order because the disordered states are always (for small noise unstable) stationary solutions. We started with an initial polar order of $p_0 = 2\pi \times 10^{-6}$ and considered the state as polarly ordered if the final polar order (after a total time of 250 or 1000) is larger than p_0 and disordered if $p < p_0$ in the final state; see Fig. 17. Typically we find also a decreasing trend of p at the end of the observation time if $p < p_0$ and an increasing trend if $p > p_0$.

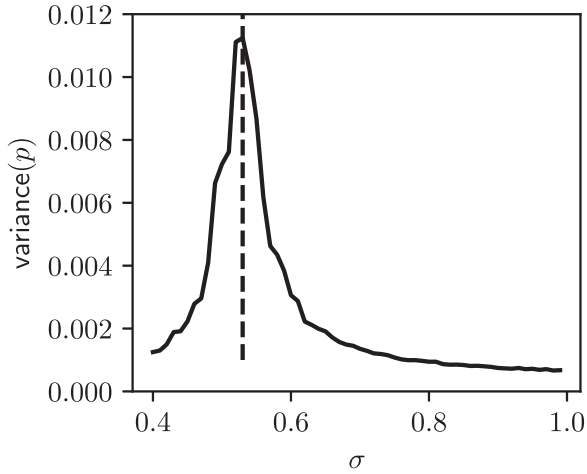


FIG. 16. Fluctuations of the polar order parameter p , measured in agent-based simulations. The maximum at $\sigma = 0.53$ indicates the onset of collective motion. Parameters are as described in the caption of Fig. 15.

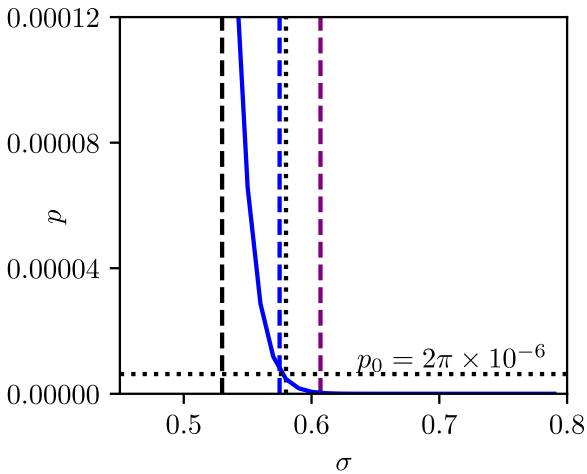


FIG. 17. Polar order parameter p obtained with ring-kinetic theory. The system was initiated without correlations and with a minimal polar order of $p_0 = 2\pi \times 10^{-6} \approx 6.3 \times 10^{-6}$. The graph shows the polar order after a waiting time of $T = 250$. Far in the disordered regime the system is already in steady state after this time, that means the correlations are stationary. However, close to the transition and also in the polarly ordered regime the systems needs much longer to become stationary. In particular at small noise, the polar order is still much below its steady state value. Therefore, we estimate the transition by a numerical stability analysis of the disordered state: we consider the system as disordered if the polar order parameter after time T is below its initial value p_0 . In that case we typically observe that the polar order parameter is still decreasing at the end of the observation time. However, we consider the system as ordered when the final value p is larger than p_0 . In that case we also typically observe that p is still increasing at the end of the observation time. The transition noise strength between polar order and disorder is displayed as the blue vertical dashed line at $\sigma = 0.505$. The black and purple vertical dashed lines show the transition noise strengths obtained in agent-based simulations and in mean-field theory, respectively. Parameters are as described in the caption of Fig. 15.

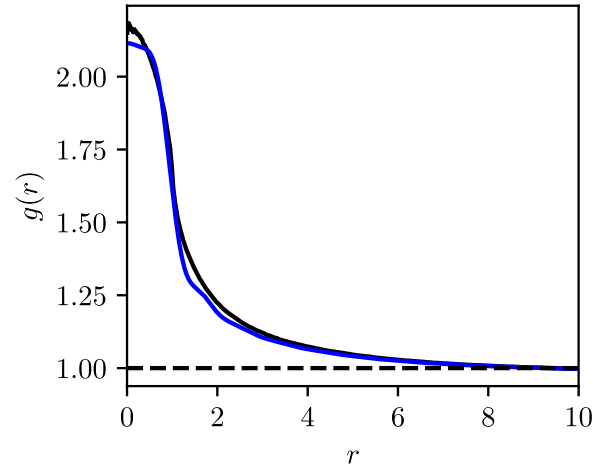


FIG. 18. Radial distribution function $g(r)$ obtained from ring-kinetic theory (blue solid line) and sampled directly from agent-based simulations according to Eq. (56) (black solid line) for $\sigma = 0.58$ shortly before the onset of flocking. Other parameters are as described in the caption of Fig. 3.

In that way, the ring-kinetic theory predicts the flocking transition at $\sigma_{c, \text{rk}} = 0.575(5)$. This result is an improvement over mean field for the given system size. Furthermore, it coincides within the measurement uncertainty with the transition noise strength for large systems. Within the ring-kinetic theory we employed the limit $N \rightarrow \infty$ to derive the time evolution equations. However, we solve the ring-kinetic equations on a finite domain, which can in principle introduce a finite size effect. Here, however, the pair correlation functions decay to zero at distances much smaller than the considered system sizes; see Fig. 18. Thus, we can expect the ring-kinetic theory to resemble large systems.

Similar to phase transitions in equilibrium spin systems, correlations shift the flocking transition toward smaller noise, compared to mean-field theory. That means correlations favor disorder. This can be understood qualitatively as follows. If we consider a particle that moves not in the direction of the majority, then without correlations it would be convinced to join the majority due to interactions with other particles (that on average behave as the majority). If correlations are present, then the interaction partners can have the same direction as the considered particle (different from the majority) due to correlations. That means a particle that is oriented differently from the global average, is likely to be accompanied by other particles that differ from the global average if strong angular correlations are present. This correlation effect weakens the alignment mechanism compared to mean field.

To confirm the applicability of the ring-kinetic theory, we apply it also at different densities. We formally assumed the limit $N \rightarrow \infty$ in the derivation of the kinetic equations, in particular in the distribution of the number particles within a circle. For small densities we require large systems to obtain reasonably large particle numbers N . However, for those large systems the numerical solution of the ring-kinetic equations becomes computationally too expensive with the Fourier techniques we are using. Therefore, we focus on larger densities $C_1 = 3$ and $C_1 = 5$. For those parameters, the ring-kinetic

theory becomes unstable when approaching the flocking transition. The reason is that the distribution of the number of particles within a circle [Eq. (49)] becomes unphysical, producing negative probabilities when C_2 approaches C_1 . This artefact is caused by the fact that in reality, higher-order correlations have to be taken into account to predict the correct distribution of the number of particles within a circle. Ignoring them can lead to negative prefactors w_k in the ring-kinetic Eqs. (44) and that make the equation unstable. To fix this problem we require some information about higher-order correlations, such that a reasonable (physical) distribution of the number of particles within a circle can be used. In particular we need the local integrals over the correlation functions:

$$C_k := N^k \int G_k(1, \dots, k) \prod_{l=1}^k \theta(R - |\mathbf{r}_l|) dl. \quad (60)$$

Other relevant quantities that affect the number of neighbor distribution are

$$D_k := N^{k-1} \int G_k(1, 2, \dots, k) d1 \prod_{l=2}^k \theta(R - |\mathbf{r}_1 - \mathbf{r}_l|) dl; \quad (61)$$

see Ref. [30] for details. We are estimating the next order coefficients C_3 and D_3 using a closure ansatz presented in the next section.

VII. CLOSURE RELATION

To obtain an estimate for the three-particle correlation function G_3 we employ the closure ansatz

$$P_3(1, 2, 3) = \Psi(1, 2)\Psi(2, 3) + \Psi(1, 3)\Psi(2, 3) + \Psi(1, 3)\Psi(1, 2), \quad (62)$$

where $\Psi(1, 2)$ is a symmetric function

$$\Psi(1, 2) = \Psi(2, 1) \quad (63)$$

that satisfies

$$\int d2 \Psi(1, 2) = \Gamma = \text{const.} \quad (64)$$

We assume furthermore that $\Psi(1, 2)$ is translational invariant. Similar closures have been used in astronomy and plasma physics; see, e.g., Refs. [45–47]. Here, the ansatz is motivated by the limit of small noise. There, nearby particles are strongly aligned. If particle one is aligned with particle two and particle two is aligned with particle three, then particle one will be also aligned with particle three. For finite noise however, the quality of the ansatz is not evident *a priori*. It can be justified only when it leads to reasonable results.

Integrating Eq. (62) over all degrees of freedom fixes Γ as

$$\Gamma = \frac{1}{L\sqrt{6\pi}}. \quad (65)$$

Integrating Eq. (62) over the degrees of freedom of particle three leads to

$$\Psi(1, 2) = \frac{L\sqrt{6\pi}}{2} [P_2(1, 2) - \int d3 \Psi(1, 3)\Psi(2, 3)]. \quad (66)$$

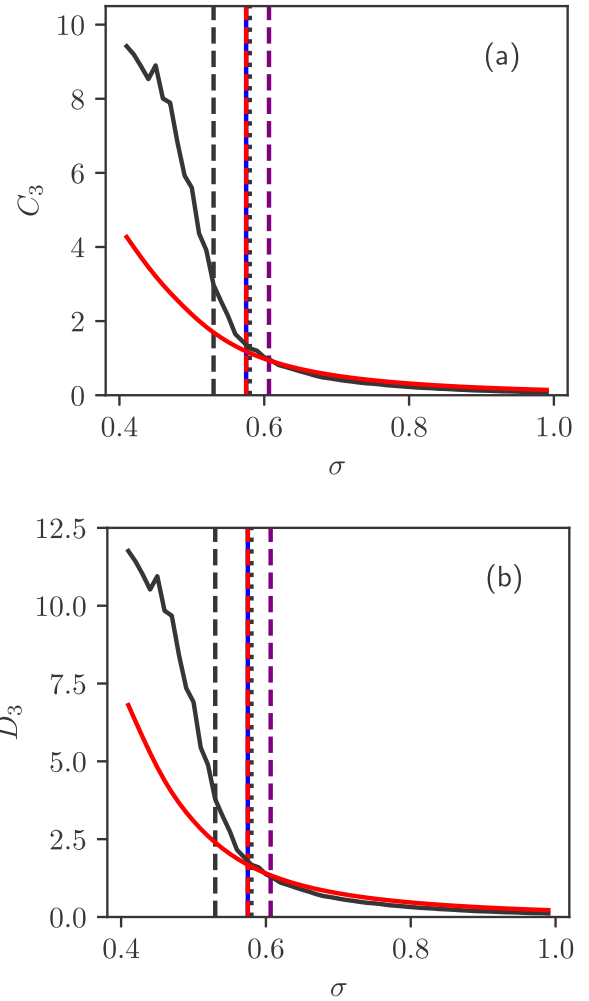


FIG. 19. Local spatial three-particle correlation parameters C_3 (a) and D_3 (b) as defined in Eq. (46) compared for direct simulations (black line) and a kinetic theory including a g_3 -closure (red line). The dashed vertical lines shows the onset of flocking measured in direct simulations (black), in the ring-kinetic theory (blue), ring-kinetic theory with closure (red), in mean-field theory (purple) and measured in [30] for much larger systems (black dotted). System parameters are as described in the caption of Fig. 15.

For a given P_2 we can iterate this equation to solve for Ψ . Starting with the initial guess $\Psi_0 = \text{const.} = \frac{1}{\sqrt{32\pi}L^2}$, the iterative procedure converges very fast, usually within one or two iterations. Having solved for Ψ , we obtain P_3 from the ansatz Eq. (62) and G_3 according to Eq. (30). In principle, we could calculate an additional collision term with G_3 in the kinetic equation of g_2 Eq. . Here, however, we neglect this term, assuming that the angular dependence of G_3 is sufficiently small. We consider only the effect of G_3 on the number of neighbor distribution. That means we calculate C_3 and D_3 according to Eqs. (60) and (61) and use the number of neighbor distribution of Ref. [30] to calculate the prefactors in the kinetic Eqs. (44) and .

Considering the parameters of Sec. VI, with density $C_1 = 1$ we compare the measured values of the three-particle correlation parameters C_3 and D_3 with the results of the kinetic theory in Fig. 19. For large noise, the relative difference between

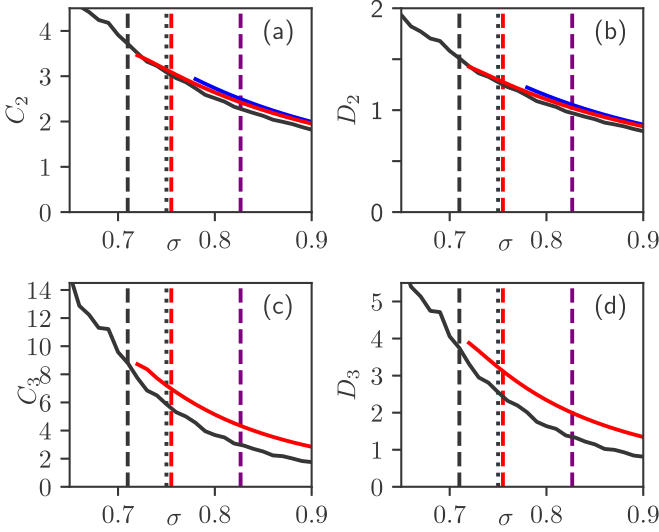


FIG. 20. Spatial pair correlation parameters C_2 (a) and D_2 (b) and three-particle correlation parameters C_3 (c) and D_3 (d) measured in agent-based simulations (black line), within pure ring-kinetic theory (blue line) and within kinetic theory including a three-particle closure (red line). The vertical dashed lines display the onset of flocking in kinetic theory with closure at $\sigma_{c,\text{closure}} = 0.755$ (red), agent-based simulations at $\sigma_{c,\text{sim}} = 0.71$ (black) and mean-field theory at $\sigma_{c,\text{mf}} = 0.827$ (purple). System parameters are $C_1 = \rho\pi R^2 = 3$, $v = R = 1$, $N = 300$. For comparison we give also the measured flocking transition in the thermodynamic limit $\sigma_{c,\text{sim},N \rightarrow \infty} = 0.75$ (dotted black line) obtained for $N = 10^4$ in Ref. [30].

theory and simulation is large. As discussed above, the closure ansatz Eq. (62) is motivated mainly for strong alignment. Thus, we do not expect that it works well for large noise. However, in that case the influence of three-particle correlations is negligible any way. For smaller noise, coming closer to the flocking transition, the closure becomes better and agrees very well with the direct simulations. Even closer to the transition, the closure clearly underestimates the correlations. However, the agreement is satisfactory almost up to the transition. The introduction of the closure relations also improves the agreement of pair correlations with the simulation, cf. Fig. 15. Furthermore, the prediction of the transition noise strength is further improved to $\sigma_{c,\text{closure}} = 0.525(5)$. It agrees with the value from direct simulations within the reached uncertainty.

A. Density dependence

For larger densities $C_1 = 3$ and $C_1 = 5$ we find similar results. The pair correlations predicted by the kinetic theory agree within about 10% with the direct simulations as long as the kinetic theory is stable; cf. Figs. 20(a) and 20(b) and Figs. 21(a) and 21(b). Also the three-particle correlations predicted by the closure ansatz have the correct order of magnitude; see Figs. 20(c) and 20(d) and Figs. 21(c) and 21(d). As discussed above, the closure ansatz is designed for aligning particles and not for strong noise. Thus, far in the disordered regime, the pure ring-kinetic treatment outperforms the kinetic theory with three-particle closure.

For $C_1 = 3$ we find the transition noise strength at $\sigma_{c,\text{closure}} = 0.755(5)$ compared to the value measured in direct

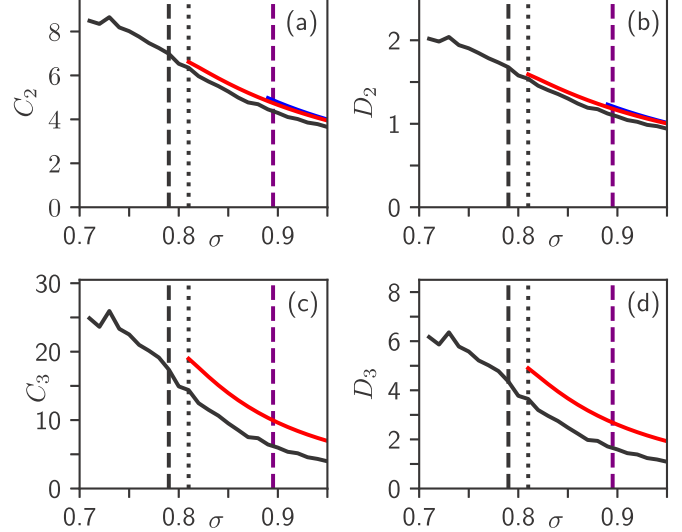


FIG. 21. Spatial pair correlation parameters C_2 (a) and D_2 (b) and three-particle correlation parameters C_3 (c) and D_3 (d) measured in agent-based simulations (black line), within pure ring-kinetic theory (blue line) and within kinetic theory including a three-particle closure (red line). The vertical dashed lines display the onset of flocking in agent-based simulations at $\sigma_{c,\text{sim}} = 0.79$ (black) and mean-field theory at $\sigma_{c,\text{mf}} = 0.895$ (purple). System parameters are $C_1 = \rho\pi R^2 = 5$, $v = R = 1$, $N = 500$. For comparison we give also the measured flocking transition in the thermodynamic limit $\sigma_{c,\text{sim},N \rightarrow \infty} = 0.81$ (dotted black line) obtained for $N = 10^4$ in Ref. [30].

simulations for $N = 300$ $\sigma_{c,\text{sim}} = 0.71(1)$, for $N = 10000$ $\sigma_{c,\text{sim},N \rightarrow \infty} = 0.75(1)$ [30] and the mean-field value $\sigma_{c,\text{mf}} = 0.827$. For $C_1 = 5$ the kinetic theory breaks down shortly above the transition because pair and three-particle correlations are no longer sufficient to reproduce the correct distribution of the number of particles within a circle. Instead, unphysical, negative probabilities are produced when higher-order correlations are neglected, similar to the pure ring-kinetic treatment at $C_1 = 3$. The transition measured is for $N = 500$ $\sigma_{c,\text{sim}} = 0.79(1)$, for $N = 10000$ $\sigma_{c,\text{sim},N \rightarrow \infty} = 0.81(1)$ [30] and the mean-field value is $\sigma_{c,\text{mf}} = 0.895$. Although the kinetic theory breaks down shortly before the onset of flocking, the predicted two-particle correlations are still very close to the measured values and the three-particle correlations have the correct order of magnitude. In summary, we find that the role of three-particle correlations increases for larger densities at the flocking transition. Therefore, one expects that also higher-order correlations become important. This explains that deviations between theory and agent-based simulations increase for large densities. Nevertheless, we still find satisfactory quantitative agreement with the simulations at $C_1 = 3$. The introduction of a closure relation for spatial three-particle correlations enlarged the parameter region of applicability of the kinetic theory significantly compared to the pure ring-kinetic theory without closure.

B. Velocity dependence

We study the parameter set of Fig. 15 at $C_1 = \rho\pi R^2 = 1$ for different velocities. For a smaller velocity of $v = 0.5$ we

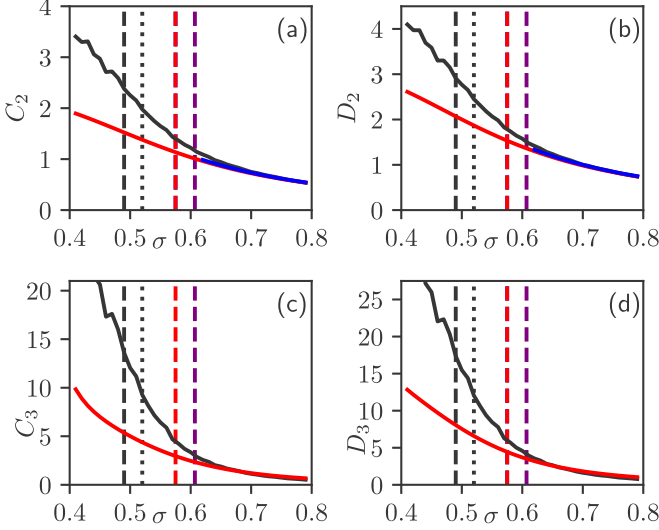


FIG. 22. Spatial pair correlation parameters C_2 (a) and D_2 (b) and three particle correlation parameters C_3 (c) and D_3 (d) measured in agent-based simulations (black line), within pure ring-kinetic theory (blue line) and within kinetic theory including a three particle closure (red line). The vertical dashed lines display the onset of flocking in ring-kinetic theory $\sigma_{c, rk} = 0.575$ (red), kinetic theory with closure at $\sigma_{c, closure} = 0.575$ (red), agent-based simulations at $\sigma_{c, sim} = 0.49$ for $N = 509$ (dashed black), for $N = 10\,000$ (dotted black) and mean-field theory at $\sigma_{c, mf} = 0.607$ (purple). System parameters are $C_1 = \rho\pi R^2 = 1$, $v = 0.5$, $R = 1$, $N = 509$.

find larger correlations; see Fig. 22. This is expected because particles have more time to interact and align at smaller velocities. The predictions of the pure ring-kinetic theory do not agree as well as for $v = 1$ similar as for larger densities. These deviations are very likely caused by higher-order correlations that need to be taken into account. As for larger densities, the ring-kinetic theory is unstable for small noise strengths and the description breaks down already above the flocking transition as soon as $C_2 = C_1$. Including the closure improves the kinetic theory significantly and extends its range of applicability. However, there are significant deviations of pair and three-particle correlations when the flocking transition is approached. Nevertheless, the flocking transition predicted by our kinetic theory is a clear improvement compared to mean-field theory.

For a larger velocity, $v = 1.5$ we find smaller spatial correlations as expected; see Fig. 23. The spatial pair correlations agree also very well between ring-kinetic theory and agent-based simulations almost until the onset of flocking. Here, the ring-kinetic theory is stable and predicts the flocking the transition at slightly too small noise. The closure improves the results only marginally. In principle, we would expect better agreement due to the smaller correlations. However, only the pure spatial correlations decrease for larger velocities, but the angular correlations increase because particles can only stay close to each other for a long time if they are aligned at high velocities. For even larger velocity (e.g., $v = 2$) we observe very strong angular correlations that suppress the onset of flocking. That means we find no onset of flocking within the ring-kinetic theory (with or without closure) although the

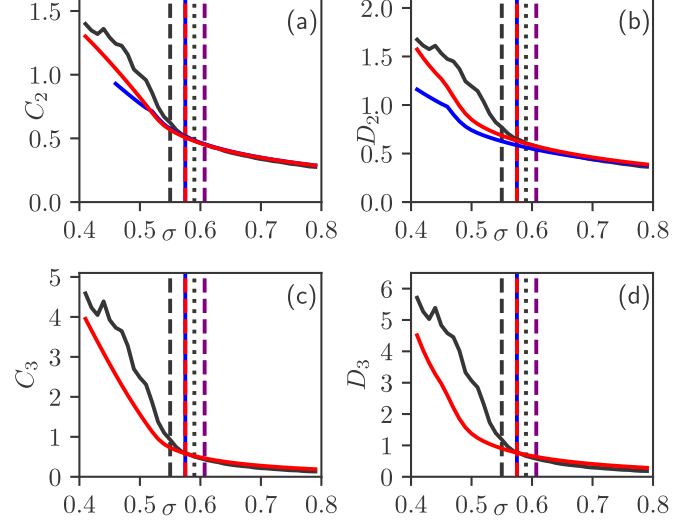


FIG. 23. Spatial pair correlation parameters C_2 (a) and D_2 (b) and three particle correlation parameters C_3 (c) and D_3 (d) measured in agent-based simulations (black line), within pure ring-kinetic theory (blue line) and within kinetic theory including a three particle closure (red line). The vertical dashed lines display the onset of flocking in pure ring-kinetic theory at $\sigma_{c, rk} = 0.575$ (blue), kinetic theory with closure at $\sigma_{c, closure} = 0.575$ (red), agent-based simulations at $\sigma_{c, sim} = 0.55$ for $N = 509$ (dashed black), for $N = 10\,000$ (dotted black) and mean-field theory at $\sigma_{c, mf} = 0.607$ (purple). System parameters are $C_1 = \rho\pi R^2 = 1$, $v = 1.5$, $R = 1$, $N = 509$.

spatial correlations are predicted reasonably well. We believe that higher-order angular correlations need to be taken into account for high velocities. The next step in this direction will be the incorporation of g_3 -collision integrals into Eq. , where g_3 is still determined by the closure Eq. (62).

VIII. COMPARISON BETWEEN FULL RING-KINETIC THEORY AND LANDAU KINETIC THEORY

The Landau theory of Ref. [29] considers the model with additive interactions, $w(n) \equiv 1$. Here, we adopt the theory to the case of nonadditive interactions, $w(n) = 1/n$ to compare it with the present approach. The equivalent of Eq. (17) of Ref. [29] is in our notation

$$\begin{aligned} \partial_t g_2(\phi_1, \phi_2, \Delta) = & -w_2 \left[1 \textcircled{\bullet} \textcircled{\bullet} \textcircled{\bullet} 2 + 2 \textcircled{\bullet} \textcircled{\bullet} \textcircled{\bullet} 1 \right] \\ & + v(\cos \phi_1 - \cos \phi_2) \partial_{\Delta_x} g_2(\phi_1, \phi_2, \Delta) \\ & + v(\sin \phi_1 - \sin \phi_2) \partial_{\Delta_y} g_2(\phi_1, \phi_2, \Delta). \end{aligned} \quad (67)$$

That is, compared to Eq. , all collision integrals on the right-hand side that contain g_2 are neglected because g_2 and the coupling constant are both assumed to be small. Furthermore, the angular diffusion term $\frac{\sigma^2}{2}(\partial_{\phi_1}^2 + \partial_{\phi_2}^2)g_2(\phi_1, \phi_2, \Delta)$ is neglected because both, σ^2 and g_2 , are assumed to be small. However, we could not make sense out of Eq. (67) without the angular diffusion term. Intuitively, it is clear that the correlations are going to diverge according to Eq. (67) because there is a source term but no damping term. In fact, we find indeed a diverging g_2 if we integrate Eq. (67) numerically in

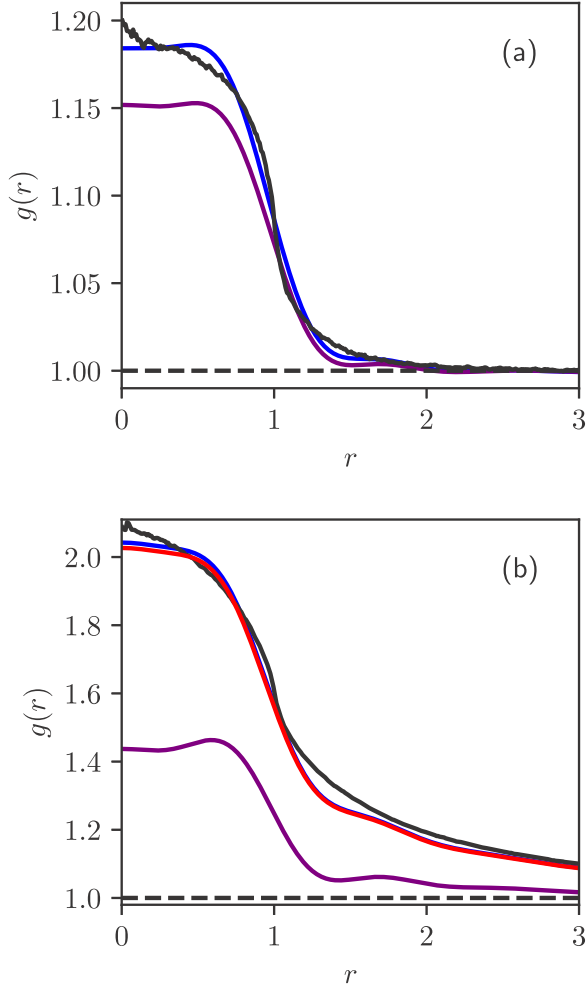


FIG. 24. Radial distribution function $g(r)$ obtained from agent-based simulations (black line), ring-kinetic theory (blue line, almost hidden below the red line), kinetic theory with closure (red line) and Landau-kinetic theory according to Eq. (68) (purple line). In panel (a) we used the parameters described in the caption of Fig. 3; in panel (b) the noise is smaller $\sigma = 0.6$ but still in the disordered phase, other parameters are the same.

Fourier space. Furthermore, we do not see how the principal problem of diverging g_2 can be prevented in any parameter limit if the noise term is neglected. Therefore, we did not neglect the angular diffusion and considered instead the time evolution equation

$$\begin{aligned} \partial_t g_2(\phi_1, \phi_2, \Delta) = & -w_2 \left[1 \begin{array}{c} \odot \\ \longrightarrow \\ \bullet \end{array} 2 + 2 \begin{array}{c} \odot \\ \longrightarrow \\ \bullet \end{array} 1 \right] \\ & + v(\cos \phi_1 - \cos \phi_2) \partial_{\Delta_x} g_2(\phi_1, \phi_2, \Delta) \\ & + v(\sin \phi_1 - \sin \phi_2) \partial_{\Delta_y} g_2(\phi_1, \phi_2, \Delta) \\ & + \frac{\sigma^2}{2} (\partial_{\phi_1}^2 + \partial_{\phi_2}^2) g_2(\phi_1, \phi_2, \Delta). \end{aligned} \quad (68)$$

In Fig. 24(a) we show the pair correlation function $g(r)$ obtained from the Landau kinetic theory (purple line) compared to the results of the full ring-kinetic theory (blue line) and direct agent-based simulations (black line) for the parameters of Fig. 3 far in the disordered phase at $\sigma = 1.5$ (the flocking

transition occurs at $\sigma = 0.53$). We find that the qualitative behavior of the pair correlation function is predicted correctly, but it is too small by about 20%. We expect that the results of the Landau kinetic theory improve if much larger noise strengths are considered. In that case however, the correlations decrease and become less important.

In Fig. 24(b) we display the pair correlation function for the same parameters but smaller noise, $\sigma = 0.6$. The system is still disordered but closer to the flocking transition. We added also the curve of our kinetic theory including a three-particle closure (red line). Here, we see that the Landau theory is off by more than 50%, the ring-kinetic theory agrees quantitatively very well with the agent-based simulations. We conclude that the simplified Landau kinetic theory of Ref. [29] is a good approximation far in the disordered regime. However, it is not suitable to describe the system in the vicinity of the flocking transition.

IX. DISCUSSION

We consider polarly aligning self-propelled point particles in two dimensions. The investigated models are in the spirit of the famous Vicsek models. However, they follow a continuous time dynamics given by a system of Langevin equations. For technical reasons it would be desirable to introduce only pair interactions. But in that case the model behaves qualitatively different from the Vicsek model: in the ordered phase, no bands or cross sea patterns are observed. Instead, strongly aligned high density clusters are formed.

To observe a behavior qualitatively equivalent to the Vicsek model, we need to introduce N -particle-, that is nonadditive interactions as pointed out recently in Refs. [39,40].

The presence of N -particle interactions seriously complicates matters. As a consequence, the kinetic equations contain weight factors that are expectation values of the distribution of the number of particles within a circle. It was discovered recently how this distribution can be calculated exactly even in the presence of many particle correlations [30]. Employing this predecessor work, we were able to handle the complications arising from the nonadditive interactions. From a technical point of view, the incorporation of N -particle interactions into ring-kinetic theories is a significant development that can be of interest also for other (possibly passive) systems with interactions of this type.

Truncating the BBGKY-hierarchy after the second equation we obtain the time evolution equations for the one-particle distribution and the pair correlation function, neglecting higher-order correlations. We solve those equations numerically, transforming both spatial and angular dependence to Fourier space. We compare the resulting steady states with direct, agent-based simulations of the Langevin equations. In the disordered phase and not too close to the onset of flocking, we find excellent agreement between ring-kinetic theory and direct simulations.

Reducing the noise strength, we find the onset of flocking in the ring-kinetic theory. Compared to the homogeneous mean-field theory, the transition is shifted toward smaller noise. This shift is caused by positive angular pair correlations that stabilize the disordered phase. The same effect is seen in direct simulations. For $C_1 = \rho\pi R^2 = 1$ we find good quanti-

tative agreement of the onset of flocking between ring-kinetic theory and agent-based simulations of large systems. For larger densities, close to the onset of flocking, the ring-kinetic equations become unstable because the weights depending on the distribution of the number of particles within a circle are predicted wrong by the ring-kinetic theory. This is not too surprising as it is known that higher-order correlations are required to predict the correct distribution of the number of particles within a circle and thus the correct weights within the ring-kinetic equations [30].

To enlarge the applicability of the kinetic theory we introduce a closure ansatz. With that relation we calculate the three-particle correlation function g_3 from the single particle distribution and the pair correlations. In this work, we do not consider the full effect of the three-particle correlations, but only its influence on the distribution of the number of particles within a circle and hence on the weights in the kinetic equations. Phrased differently, we consider only the effect of three-particle correlations on the spatial distribution and we neglect angular three-particle correlations that are entering the kinetic equations by additional collision integrals.

Far in the disordered phase, for large noise, the closure ansatz does not agree very well with direct simulations. However, in that region three-particle correlations can be ignored as discussed above. At about the onset of flocking, the closure ansatz for the spatial three-particle correlations agrees within about 20% with direct simulations. That means it is a considerable improvement compared to the neglect of those correlations. Furthermore, the introduction of the closure significantly enlarges the parameter regime where spatial pair correlations agree quantitatively with direct simulations.

Employing the closure ansatz we are able to describe the onset of flocking within kinetic theory also for larger densities such as $C_1 = \rho\pi R^2 = 3$ where the onset of flocking agrees quantitatively with the measured value for large systems within measurement uncertainty. In comparison, the deviations of the transition noise in mean-field theory are 16% and 13%, respectively.

We also consider the dependence on the particle velocity and find good agreement between ring-kinetic theory and simulations at large noise for all considered velocities. Close to the onset of flocking, small velocities increase in particular spatial correlation, whereas high velocities suppress spatial correlations but favor angular correlations. This is intuitive because at high speed, particles remain close to each other for long times only if they have roughly the same velocity, and for small velocities, particles that interact have a long time to align. Both effects lead to deviations of the predicted onset of flocking in the kinetic theory with three-particle closure. However, the results of the kinetic theory are still a major improvement over mean-field theory. There is further potential to improve the theory for high velocities by incorporating g_3 collision integrals in the g_2 equation by means of the presented closure ansatz.

In general, it can be seen that the ring-kinetic theory (possibly extended by a closure relation estimating g_3) gives quantitatively good results as long as higher-order correlations are not too large. Far in the disordered phase (for large noise), this is always the case. Close to the flocking transition we observe larger spatial correlations for higher particle densities or small velocities and large angular correlations for high velocities.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the GWK support for funding this project by providing computing time through the Center for Information Services and HPC (ZIH) at TU Dresden on the HRSK-II. The authors gratefully acknowledge the Universitätsrechenzentrum Greifswald for providing computing time. We thank Aurelio Patelli, Fernando Peruani, and Sven Stroteich for valuable discussions.

APPENDIX A: TIME EVOLUTION EQUATIONS

The explicit expression for the time evolution equation of the one-particle angular distribution Eq. (44) reads

$$\begin{aligned}
 \partial_t p(\phi_1) = & -\rho w_2 \partial_{\phi_1} \int \sin(\phi_2 - \phi_1) [g_2(\phi_1, \phi_2, \Delta) + p(\phi_1)p(\phi_2)] \theta(R - |\Delta|) d\phi_2 d\Delta - \rho^2 [w_3 - w_2] \partial_{\phi_1} \\
 & \times \left\{ \int \sin(\phi_2 - \phi_1) p(\phi_1) g_2(\phi_2, \phi_3, \tilde{\Delta}) \theta(R - |\Delta|) \theta(R - |\Delta + \tilde{\Delta}|) d\Delta d\tilde{\Delta} d\phi_2 d\phi_3 \right. \\
 & \left. + \int \sin(\phi_2 - \phi_1) p(\phi_2) g_2(\phi_1, \phi_3, \Delta + \tilde{\Delta}) \theta(R - |\Delta|) \theta(R - |\Delta + \tilde{\Delta}|) d\Delta d\tilde{\Delta} d\phi_2 d\phi_3 \right\} - \rho^3 [w_4 - 2w_3 + w_2] \partial_{\phi_1} \\
 & \times \left\{ \int \sin(\phi_2 - \phi_1) g_2(\phi_1, \phi_3, \Delta_{13}) g_2(\phi_2, \phi_4, \Delta_{24}) \theta_{12} \theta_{13} \theta_{14} d2 d3 d4 \right\} + \frac{\sigma^2}{2} \partial_{\phi_1}^2 p(\phi_1). \tag{A1}
 \end{aligned}$$

Similar, one finds for the time evolution equation of the pair correlation function Eq.

$$\begin{aligned}
 \partial_t g_2(\phi_1, \phi_2, \Delta) = & -w_2 \partial_{\phi_1} \theta_{12} \sin(\phi_2 - \phi_1) [g_2(\phi_1, \phi_2, \Delta) + p(\phi_1)p(\phi_2)] \\
 & - \theta_{12} \rho (w_3 - w_2) \partial_{\phi_1} \int d3 \theta_{13} \sin(\phi_2 - \phi_1) p(\phi_2) g_2(\phi_1, \phi_3, \Delta_{13}) \\
 & - \theta_{12} \rho (w_3 - w_2) \partial_{\phi_1} \int d3 \theta_{13} \sin(\phi_2 - \phi_1) p(\phi_1) g_2(\phi_2, \phi_3, \Delta_{23}) \\
 & - \theta_{12} \rho^2 (w_4 - 2w_3 + w_2) \partial_{\phi_1} \int d3 d4 \theta_{13} \theta_{14} \sin(\phi_2 - \phi_1) g_2(\phi_1, \phi_3, \Delta_{13}) g_2(\phi_2, \phi_4, \Delta_{24})
 \end{aligned}$$

$$\begin{aligned}
& -\theta_{12} \left\{ \rho w_3 \partial_{\phi_1} \int d^3\theta_{13} \sin(\phi_3 - \phi_1) [p(\phi_1)g_2(\phi_2, \phi_3, \Delta_{23}) + p(\phi_3)g_2(\phi_1, \phi_2, \Delta_{12})] \right. \\
& + \rho [w_3 - w_2] \partial_{\phi_1} \int d^3\theta_{13} \sin(\phi_3 - \phi_1) [p(\phi_1)p(\phi_2)p(\phi_3) + p(\phi_2)g_2(\phi_1, \phi_3, \Delta_{13})] \\
& + \rho^2 [w_4 - w_3] \partial_{\phi_1} \int \sin(\phi_3 - \phi_1) d^3 d^4\theta_{13}\theta_{14} [g_2(\phi_1, \phi_2, \Delta_{12})g_2(\phi_3, \phi_4, \Delta_{34}) \\
& + g_2(\phi_1, \phi_4, \Delta_{14})g_2(\phi_2, \phi_3, \Delta_{23}) + g_2(\phi_1, \phi_3, \Delta_{13})g_2(\phi_2, \phi_4, \Delta_{24}) + p(\phi_1)p(\phi_3)g_2(\phi_2, \phi_4, \Delta_{24})] \\
& + \rho^2 [w_4 - 2w_3 + w_2] \partial_{\phi_1} \int \sin(\phi_3 - \phi_1) d^3 d^4\theta_{13}\theta_{14} \\
& \times [+p(\phi_1)p(\phi_2)g_2(\phi_3, \phi_4, \Delta_{34}) + p(\phi_2)p(\phi_3)g_2(\phi_1, \phi_4, \Delta_{14})] + \rho^3 [w_5 - 2w_4 + w_3] \\
& \times \partial_{\phi_1} \int \sin(\phi_3 - \phi_1) d^3 d^4 d^5\theta_{13}\theta_{14}\theta_{15} [g_2(\phi_1, \phi_4, \Delta_{14})g_2(\phi_2, \phi_5, \Delta_{25})p(\phi_3) \\
& + p(\phi_1)g_2(\phi_2, \phi_4, \Delta_{24})g_2(\phi_3, \phi_5, \Delta_{35})] + \rho^3 [w_5 - 3w_4 + 3w_3 - w_2] \\
& \times \partial_{\phi_1} \int \sin(\phi_3 - \phi_1) d^3 d^4 d^5\theta_{13}\theta_{14}\theta_{15} g_2(\phi_1, \phi_4, \Delta_{14})p(\phi_2)g_2(\phi_3, \phi_5, \Delta_{35}) \\
& + \rho^4 [w_6 - 3w_5 + 3w_4 - w_3] \partial_{\phi_1} \int \sin(\phi_3 - \phi_1) d^3 d^4 d^5 d^6\theta_{13}\theta_{14}\theta_{15}\theta_{16} g_2(\phi_1, \phi_4, \Delta_{14}) \\
& \times g_2(\phi_2, \phi_5, \Delta_{25})g_2(\phi_3, \phi_6, \Delta_{36}) \left. \right\} - (1 - \theta_{12}) \left\{ \rho w_2 \partial_{\phi_1} \int d^3\theta_{13} \sin(\phi_3 - \phi_1) \right. \\
& \times [p(\phi_1)g_2(\phi_2, \phi_3, \Delta_{23}) + p(\phi_3)g_2(\phi_1, \phi_2, \Delta_{12})] \\
& + \rho^2 [w_3 - w_2] \partial_{\phi_1} \int \sin(\phi_3 - \phi_1) d^3 d^4\theta_{13}\theta_{14} [g_2(\phi_1, \phi_2, \Delta_{12})g_2(\phi_3, \phi_4, \Delta_{34}) \\
& + g_2(\phi_1, \phi_4, \Delta_{14})g_2(\phi_2, \phi_3, \Delta_{23}) + g_2(\phi_1, \phi_3, \Delta_{13})g_2(\phi_2, \phi_4, \Delta_{24}) + p(\phi_1)p(\phi_3)g_2(\phi_2, \phi_4, \Delta_{24})] \\
& + \rho^3 [w_4 - 2w_3 + w_2] \partial_{\phi_1} \int \sin(\phi_3 - \phi_1) d^3 d^4 d^5 \\
& \times \theta_{13}\theta_{14}\theta_{15} [g_2(\phi_1, \phi_4, \Delta_{14})g_2(\phi_2, \phi_5, \Delta_{25})p(\phi_3) + p(\phi_1)g_2(\phi_2, \phi_4, \Delta_{24})g_2(\phi_3, \phi_5, \Delta_{35})] \\
& + \rho^4 [w_5 - 3w_4 + 3w_3 - w_2] \partial_{\phi_1} \int \sin(\phi_3 - \phi_1) d^3 d^4 d^5 d^6\theta_{13}\theta_{14}\theta_{15}\theta_{16} g_2(\phi_1, \phi_4, \Delta_{14}) \\
& \times g_2(\phi_2, \phi_5, \Delta_{25})g_2(\phi_3, \phi_6, \Delta_{36}) \left. \right\} + 1 \leftrightarrow 2, \\
& - v[\cos(\phi_1) - \cos(\phi_2)] \partial_{\Delta_x} g_2(\phi_1, \phi_2, \Delta) - v[\sin(\phi_1) - \sin(\phi_2)] \partial_{\Delta_y} g_2(\phi_1, \phi_2, \Delta) \\
& + \frac{\sigma^2}{2} (\partial_{\phi_1}^2 + \partial_{\phi_2}^2) g_2(\phi_1, \phi_2, \Delta), \tag{A2}
\end{aligned}$$

where $+1 \leftrightarrow 2$ means that all terms above are added with indexes one and two interchanged.

APPENDIX B: FOURIER TRANSFORM

1. Time evolution equations

We solve the time evolution Eqs. (44) and (51) numerically in Fourier space, transforming both spatial and angular coordinates. This has the huge advantage that all appearing integrals (also high-dimensional ones) can be solved analytically. We

employ the following Fourier ansatz for the one-particle distribution and the pair correlation function:

$$p(\phi) = \sum_k A_k \exp(ik\phi), \tag{B1}$$

$$\begin{aligned}
g_2(\phi_1, \phi_2, \Delta) &= \sum_{k,l,m,n} F_{k,l,m,n} \exp(ik\phi_1) \exp(il\phi_2) \\
&\times \exp(im\Delta_x 2\pi/L) \exp(in\Delta_y 2\pi/L). \tag{B2}
\end{aligned}$$

The number of neighbor distribution that determines the weights in Eqs. (44) and (51) is known analytically up to a discrete Fourier transform [30]. It depends only on the integrals Eq. (46) when three-particle and higher-order correlations are

neglected. Those integrals can be evaluated in Fourier space. It is useful to introduce the abbreviations

$$K_{m,n} := \frac{1}{L^2} \int_{-L/2}^{L/2} \int_{-L/2}^{L/2} e^{im\frac{2\pi}{L}x} e^{in\frac{2\pi}{L}y} \times \theta(R - |\mathbf{x}|) dx dy$$

$$= \frac{R}{L\sqrt{m^2 + n^2}} \times J_1\left(\frac{2\pi}{L}R\sqrt{m^2 + n^2}\right), \tag{B3}$$

$$Q_{klmn}^\pm := \sum_{s,t} F_{klst} K_{m\pm s, n\pm t}, \tag{B4}$$

and

$$R_{klmn}^\pm := \sum_{s,t} F_{klst} K_{m\pm s, n\pm t} K_{s,t}, \tag{B5}$$

where J_1 is the Bessel function of the first kind. With those abbreviations one obtains

$$C_2 = 4\pi^2 N^2 R_{0000}^+, \tag{B6}$$

$$D_2 = 4\pi^2 N Q_{0000}^+. \tag{B7}$$

Thus, with the results of Ref. [30], the weights in Eqs. (44) and (51) can be expressed explicitly in terms of the Fourier modes F_{klmn} . Fourier transforming those time evolution equations results in

$$\begin{aligned} \partial_t A_k &= N w_2 k \pi (Q_{k-1,1,0,0}^+ - Q_{k+1,-1,0,0}^+) \\ &+ C_1 w_2 k \pi (A_{k-1} A_1 - A_{k+1} A_{-1}) \\ &+ N^2 2\pi^2 k (w_3 - w_2) (A_{k-1} R_{1,0,0,0}^+ - A_{k+1} R_{-1,0,0,0}^+) \\ &+ C_1 N 2\pi^2 k (w_3 - w_2) (A_1 Q_{k-1,0,0,0}^+ - A_{-1} Q_{k+1,0,0,0}^+) \\ &+ N^3 4\pi^3 k (w_4 - 2w_3 + w_2) \\ &\times (Q_{k-1,0,0,0}^+ R_{1,0,0,0}^+ - Q_{k+1,0,0,0}^+ R_{-1,0,0,0}^+) \\ &- \frac{\sigma^2}{2} k^2 A_k \end{aligned} \tag{B8}$$

and

$$\partial_t F_{klmn} = \sum_{i=1}^{39} \textcircled{i} + \{k \leftrightarrow l, m \leftrightarrow -m, n \leftrightarrow -n\}, \tag{B9}$$

where the terms \textcircled{i} are given by

$$\textcircled{1} = w_2 \frac{k}{2} (Q_{k-1,l+1,m,n}^- - Q_{k+1,l-1,m,n}^-), \tag{B10}$$

$$\textcircled{2} = w_2 \frac{k}{2} K_{mn} (A_{k-1} A_{l+1} - A_{k+1} A_{l-1}), \tag{B11}$$

$$\textcircled{3} = N\pi (w_3 - w_2) k K_{mn} \times (Q_{k-1,0,0,0}^+ A_{l+1} - Q_{k+1,0,0,0}^+ A_{l-1}), \tag{B12}$$

$$\textcircled{4} = N\pi (w_3 - w_2) k \times (R_{l+1,0,m,n}^+ A_{k-1} - R_{l-1,0,m,n}^+ A_{k+1}), \tag{B13}$$

$$\textcircled{5} = N^2 2\pi^2 (w_4 - 2w_3 + w_2) k \times (R_{l+1,0,m,n}^+ Q_{k-1,0,0,0}^+ - R_{l-1,0,m,n}^+ Q_{k+1,0,0,0}^+), \tag{B14}$$

$$\textcircled{6} = N k w_3 \pi (A_{k-1} R_{l,1,m,n}^+ - A_{k+1} R_{l,-1,m,n}^+), \tag{B15}$$

$$\textcircled{7} = C_1 \pi k w_3 (Q_{k-1,l,m,n}^- A_1 - Q_{k+1,l,m,n}^- A_{-1}), \tag{B16}$$

$$\textcircled{8} = N\pi k (w_3 - w_2) K_{mn} A_l (Q_{k-1,1,0,0}^+ - Q_{k+1,-1,0,0}^+), \tag{B17}$$

$$\textcircled{9} = N^2 (w_4 - w_3) k 2\pi^2 \times (Q_{k-1,l,m,n}^- R_{1,0,0,0}^+ - Q_{k+1,l,m,n}^- R_{-1,0,0,0}^+), \tag{B18}$$

$$\textcircled{10} = N^2 (w_4 - w_3) 2\pi^2 k \times (R_{l,1,m,n}^+ Q_{k-1,0,0,0}^+ - R_{l,-1,m,n}^+ Q_{k+1,0,0,0}^+), \tag{B19}$$

$$\textcircled{11} = N^2 (w_4 - w_3) k 2\pi^2 \times R_{l,0,m,n}^+ (Q_{k-1,1,0,0}^+ - Q_{k+1,-1,0,0}^+), \tag{B20}$$

$$\textcircled{12} = C_1 N (w_4 - w_3) k 2\pi^2 \times R_{l,0,m,n}^+ (A_{k-1} A_1 - A_{k+1} A_{-1}), \tag{B21}$$

$$\textcircled{13} = N^2 (w_4 - 2w_3 + w_2) k 2\pi^2 A_l K_{mn} \times (A_{k-1} R_{1,0,0,0}^+ - A_{k+1} R_{-1,0,0,0}^+), \tag{B22}$$

$$\textcircled{14} = C_1 N k 2\pi^2 (w_4 - 2w_3 + w_2) A_l K_{mn} \times (Q_{k-1,0,0,0}^+ A_1 - Q_{k+1,0,0,0}^+ A_{-1}), \tag{B23}$$

$$\textcircled{15} = N^2 C_1 4\pi^3 k (w_5 - 2w_4 + w_3) R_{l,0,m,n}^+ \times (Q_{k-1,0,0,0}^+ A_1 - Q_{k+1,0,0,0}^+ A_{-1}), \tag{B24}$$

$$\textcircled{16} = N^3 k 4\pi^3 (w_5 - 2w_4 + w_3) R_{l,0,m,n}^+ \times (A_{k-1} R_{1,0,0,0}^+ - A_{k+1} R_{-1,0,0,0}^+), \tag{B25}$$

$$\textcircled{17} = N^3 k 4\pi^3 (w_5 - 3w_4 + 3w_3 - w_2) K_{mn} A_l \times (Q_{k-1,0,0,0}^+ R_{1,0,0,0}^+ - Q_{k+1,0,0,0}^+ R_{-1,0,0,0}^+), \tag{B26}$$

$$\textcircled{18} = N^4 8\pi^4 k (w_6 - 3w_5 + 3w_4 - w_3) R_{l,0,m,n}^+ \times (Q_{k-1,0,0,0}^+ R_{1,0,0,0}^+ - Q_{k+1,0,0,0}^+ R_{-1,0,0,0}^+), \tag{B27}$$

$$\textcircled{19} = N k \pi w_2 K_{mn} \times (A_{k-1} F_{l,1,-m,-n} - A_{k+1} F_{l,-1,-m,-n}), \tag{B28}$$

$$\textcircled{20} = -N k \pi w_2 (A_{k-1} R_{l,1,m,n}^+ - A_{k+1} R_{l,-1,m,n}^+) \tag{B29}$$

$$\textcircled{21} = C_1 w_2 k \pi (F_{k-1,l,m,n} A_1 - F_{k+1,l,m,n} A_{-1}), \tag{B30}$$

$$\textcircled{22} = -C_1 w_2 k \pi (Q_{k-1,l,m,n}^- A_1 - Q_{k+1,l,m,n}^- A_{-1}), \tag{B31}$$

$$\textcircled{23} = N^2 2\pi^2 k (w_3 - w_2) \times (R_{1,0,0,0}^+ F_{k-1,l,m,n} - R_{-1,0,0,0}^+ F_{k+1,l,m,n}), \tag{B32}$$

$$\textcircled{24} = -N^2 2\pi^2 k (w_3 - w_2) \times (Q_{k-1,l,m,n}^- R_{1,0,0,0}^+ - Q_{k+1,l,m,n}^- R_{-1,0,0,0}^+), \tag{B33}$$

$$\textcircled{25} = N^2 k 2\pi^2 (w_3 - w_2) K_{mn} \times (Q_{k-1,0,0,0}^+ F_{l,1,-m,-n} - Q_{k+1,0,0,0}^+ F_{l,-1,-m,-n}), \tag{B34}$$

$$\textcircled{26} = -N^2 k 2\pi^2 (w_3 - w_2) \times (Q_{k-1,0,0,0}^+ R_{l,1,m,n}^+ - Q_{k+1,0,0,0}^+ R_{l,-1,m,n}^+), \quad (\text{B35})$$

$$\textcircled{27} = N^2 k 2\pi^2 (w_3 - w_2) K_{mn} \times F_{l,0,-m,-n} (Q_{k-1,1,0,0}^+ - Q_{k+1,-1,0,0}^+), \quad (\text{B36})$$

$$\textcircled{28} = -N^2 k 2\pi^2 (w_3 - w_2) R_{l,0,m,n}^+ \times (Q_{k-1,1,0,0}^+ - Q_{k+1,-1,0,0}^+), \quad (\text{B37})$$

$$\textcircled{29} = C_1 N k 2\pi^2 (w_3 - w_2) K_{mn} \times F_{l,0,-m,-n} (A_{k-1} A_1 - A_{k+1} A_{-1}), \quad (\text{B38})$$

$$\textcircled{30} = -C_1 N k 2\pi^2 (w_3 - w_2) \times R_{l,0,m,n}^+ (A_{k-1} A_1 - A_{k+1} A_{-1}), \quad (\text{B39})$$

$$\textcircled{31} = C_1 N^2 k 4\pi^3 (w_4 - 2w_3 + w_2) K_{mn} \times F_{l,0,-m,-n} (Q_{k-1,0,0,0}^+ A_1 - Q_{k+1,0,0,0}^+ A_{-1}), \quad (\text{B40})$$

$$\textcircled{32} = -C_1 N^2 k 4\pi^3 (w_4 - 2w_3 + w_2) \times R_{l,0,m,n}^+ (Q_{k-1,0,0,0}^+ A_1 - Q_{k+1,0,0,0}^+ A_{-1}), \quad (\text{B41})$$

$$\textcircled{33} = N^3 k 4\pi^3 (w_2 - 2w_3 + w_2) K_{mn} \times F_{l,0,-m,-n} (A_{k-1} R_{1,0,0,0}^+ - A_{k+1} R_{-1,0,0,0}^+), \quad (\text{B42})$$

$$\textcircled{34} = -N^3 k 4\pi^3 (w_2 - 2w_3 + w_2) \times R_{l,0,m,n}^+ (A_{k-1} R_{1,0,0,0}^+ - A_{k+1} R_{-1,0,0,0}^+), \quad (\text{B43})$$

$$\textcircled{35} = N^4 k 8\pi^4 (w_5 - 3w_4 + 3w_3 - w_2) F_{l,0,-m,-n} \times K_{mn} (Q_{k-1,0,0,0}^+ R_{1,0,0,0}^+ - Q_{k+1,0,0,0}^+ R_{-1,0,0,0}^+), \quad (\text{B44})$$

$$\textcircled{36} = -N^4 k 8\pi^4 (w_5 - 3w_4 + 3w_3 - w_2) R_{l,0,m,n}^+ \times (Q_{k-1,0,0,0}^+ R_{1,0,0,0}^+ - Q_{k+1,0,0,0}^+ R_{-1,0,0,0}^+), \quad (\text{B45})$$

$$\textcircled{37} = v m \frac{i\pi}{L} (F_{k-1,l,m,n} + F_{k+1,l,m,n}), \quad (\text{B46})$$

$$\textcircled{38} = v n \frac{\pi}{L} (F_{k-1,l,m,n} - F_{k+1,l,m,n}), \quad (\text{B47})$$

$$\textcircled{39} = -\frac{\sigma^2}{2} k^2 F_{klmn}, \quad (\text{B48})$$

and $\{k \leftrightarrow l, m \leftrightarrow -m, n \leftrightarrow -n\}$ refers to the same terms but with k and l , m and $-m$, n and $-n$ interchanged, respectively.

2. Closure relation

To evaluate the closure ansatz Eq. (62) we consider the function Ψ in Fourier space

$$\Psi(\phi_1, \phi_2, \Delta) = \sum_{klmn} U_{klmn} \exp(ik\phi_1) \times \exp(il\phi_2) \exp(im\Delta_x 2\pi/L) \exp(in\Delta_y 2\pi/L). \quad (\text{B49})$$

It is reasonable to consider scaled modes

$$u_{stuv} = L^3 U_{stuv}. \quad (\text{B50})$$

From Eqs. (64) and (65) it follows that

$$u_{s000} = \delta_{s,0} u_{0000} = \delta_{s,0} \frac{L\Gamma}{2\pi} = \delta_{s,0} \frac{1}{2\pi\sqrt{6\pi}}. \quad (\text{B51})$$

Fourier transforming the iteration Eq. (66) we obtain

$$u_{klmn} = \frac{1}{2u_{0000}} \left[\frac{1}{8\pi^3} \delta_{k,0} \delta_{l,0} \delta_{m,0} \delta_{n,0} - \sum_s u_{ksmn} u_{-slmn} + \frac{1}{2\pi} F_{klmn} \right]. \quad (\text{B52})$$

Note that this equation is compatible with the normalization condition Eq. (B51). In practice, we start with $u_{klmn} = 0$ except for u_{0000} that is given by Eq. (B51). Then we iterate Eq. (B52) to calculate u_{klmn} , that is the closure ansatz function Ψ . In the parameter regime we investigated, the recursion converges very fast. It almost reaches its fixed point already after two iterations. We used five iterations after which we reach perfect convergence within our numerical accuracy.

Once Ψ is known, the spatial three-particle correlation parameters C_3 and D_3 can be calculated via

$$\begin{aligned} C_3 + 3C_1 C_2 + C_1^3 &= N^3 \int P_3(1, 2, 3) d1 d2 d3 \\ &= 3N^3 \int \Psi(1, 2) \Psi(2, 3) \theta_1 \theta_2 \theta_3 d1 d2 d3 \\ &= 3\rho^3 \int L^6 \Psi(1, 2) \Psi(2, 3) \theta_1 \theta_2 \theta_3 d1 d2 d3 \\ &= 3\rho^3 \int \sum_{stuv} \exp(is\phi_1) \exp(it\phi_2) \exp(iu(x_2 - x_1) 2\pi/L) \end{aligned}$$

$$\begin{aligned}
 & \times \exp(i(y_2 - y_1)2\pi/L) \sum_{klmn} \exp(ik\phi_2) \exp(il\phi_3) \\
 & \times \exp(im(x_3 - x_2)2\pi/L) \\
 & \times \exp(in(y_3 - y_2)2\pi/L) \theta_1 \theta_2 \theta_3 d1 d2 d3 \\
 & = 3N^3 (2\pi)^3 \sum_{tuvmn} u_{0,t,u,v} u_{-t,0,m,n} K_{u,v} K_{u-m,v-n} K_{m,n}
 \end{aligned} \tag{B53}$$

and similar

$$\begin{aligned}
 D_3 + 2D_2C_1 + C_2 + C_1^2 &= N^2 \int \theta_{12}\theta_{13}G_3(1, 2, 3) d1 d2 d3 \\
 &= (2\pi)^3 N^2 \sum_{lmnuv} u_{0lmn} u_{-l0uv} K_{m-u,n-v} K_{u,v} \\
 &+ (2\pi)^3 N^2 \sum_{lmnuv} u_{0lmn} u_{0-luv} K_{m+u,n+v} K_{u,v} \\
 &+ (2\pi)^3 N^2 \sum_{lmnuv} u_{-l0mn} u_{l0uv} K_{m,n} K_{u,v}.
 \end{aligned} \tag{B54}$$

Note that $C_1 := \frac{N}{L^2} \pi R^2$ and $D_1 := 1$ per definition.

In summary, the numerical time evolution works as follows. We calculate the pair correlation coefficients C_2 and D_2 according to Eq. (B7). Next, we calculate the closure ansatz function with the iteration Eq. (B52) and then the three-particle correlation coefficients C_3 and D_3 according to Eqs. (B53) and (B54). Having calculated the correlation coefficients we compute the distribution of the number of particles in a circle depending on C_2 and C_3 according to Eq. (50), cf. also Ref. [30] for details. With this distribution we obtain the weights w_k according to Eq. (45). Given the weights we can eventually time evolve the time evolution Eqs. (B8) and (B9) with a simple Euler scheme.

3. Comparison to agent-based simulations

The correlation functions $g(r)$ and $h(r, \alpha, \Delta\phi)$ considered in Sec. VI can be sampled in agent-based simulations according to Eqs. (56) and (59). However, these functions can be calculated from the Fourier modes of g_2 according to

$$g(r) = 1 + 4\pi^2 \sum_{m,n} F_{00mn} J_0\left(\frac{2\pi r}{L} \sqrt{m^2 + n^2}\right), \tag{B55}$$

$$h(r, \alpha, \Delta\phi) = (2\pi)^2 \sum_{klmn} F_{klmn} \exp(il\Delta\phi) \exp[-i(k+l)\alpha] \times \left(\frac{n-im}{\sqrt{n^2+m^2}}\right)^{k+l} J_{-(k+l)}\left(\frac{2\pi r}{L} \sqrt{m^2 + n^2}\right). \tag{B56}$$

Assuming rotational symmetry, it is also possible to revert Eq. (B55) that is to calculate the spatial Fourier modes F_{00mn} from the function $g(r)$ according to

$$F_{00mn} = \frac{1}{L^2 2\pi} \int [g(r) - 1] r J_0\left(\frac{2\pi r}{L} \sqrt{m^2 + n^2}\right) dr. \tag{B57}$$

We use this relation to see the impact of the resolution in the Fourier transform on $g(r)$ by Fourier transforming the measured function according to Eq. (B57) and transforming it back according to Eq. (B55). In that way we achieved the green dashed line in Fig. 3. The solid black line in this figure is a histogram of agent-based simulations according to Eq. (56) and the solid blue line is calculated according to Eq. (B55), where the Fourier modes F_{klmn} are the result of the ring-kinetic theory.

-
- [1] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, Novel Type of Phase Transition in a System of Self-Driven Particles, *Phys. Rev. Lett.* **75**, 1226 (1995).
 - [2] J. Toner and Y. Tu, Long-Range Order in a Two-Dimensional Dynamical xy Model: How Birds Fly Together, *Phys. Rev. Lett.* **75**, 4326 (1995).
 - [3] Y. Fily and M. C. Marchetti, Athermal Phase Separation of Self-Propelled Particles with no Alignment, *Phys. Rev. Lett.* **108**, 235702 (2012).
 - [4] I. Buttinoni, J. Bialké, F. Kümmel, H. Löwen, C. Bechinger, and T. Speck, Dynamical Clustering and Phase Separation in Suspensions of Self-Propelled Colloidal Particles, *Phys. Rev. Lett.* **110**, 238301 (2013).
 - [5] C. W. Wolgemuth, Collective swimming and the dynamics of bacterial turbulence, *Biophys. J.* **95**, 1564 (2008).
 - [6] J. Toner, Y. Tu, and S. Ramaswamy, Hydrodynamics and phases of flocks, *Ann. Phys.* **318**, 170 (2005).

- [7] S. Ramaswamy, The mechanics and statistics of active matter, *Annu. Rev. Condens. Matter Phys.* **1**, 323 (2010).
- [8] T. Vicsek and A. Zafeiris, Collective motion, *Phys. Rep.* **517**, 71 (2012).
- [9] P. Romanczuk, M. Bär, W. Ebeling, B. Lindner, and L. Schimansky-Geier, Active Brownian particles, *Eur. Phys. J.: Spec. Top.* **202**, 1 (2012).
- [10] D. Klotsa, As above, so below, and also in between: Mesoscale active matter in fluids, *Soft Matter* **15**, 8946 (2019).
- [11] O. Dauchot and H. Löwen, Chemical physics of active matter, *J. Chem. Phys.* **151**, 114901 (2019).
- [12] H. Chaté, Dry aligning dilute active matter, *Annu. Rev. Condens. Matter Phys.* **11**, 189 (2020).
- [13] M. R. Shaebani, A. Wysocki, R. G. Winkler, G. Gompper, and H. Rieger, Computational models for active matter, *Nature Rev. Phys.* **2**, 181 (2020).
- [14] M. Bär, R. Großmann, S. Heidenreich, and F. Peruani, Self-propelled rods: Insights and perspectives for active matter, *Annu. Rev. Condens. Matter Phys.* **11**, 441 (2020).
- [15] G. Grégoire and H. Chaté, Onset of Collective and Cohesive Motion, *Phys. Rev. Lett.* **92**, 025702 (2004).
- [16] H. Chaté, F. Ginelli, G. Grégoire, and F. Raynaud, Collective motion of self-propelled particles interacting without cohesion, *Phys. Rev. E* **77**, 046113 (2008).
- [17] A. P. Solon, H. Chaté, and J. Tailleur, From Phase to Microphase Separation in Flocking Models: The Essential Role of Nonequilibrium Fluctuations, *Phys. Rev. Lett.* **114**, 068101 (2015).
- [18] R. Kürsten and T. Ihle, Dry Active Matter Exhibits a Self-Organized Cross Sea Phase, *Phys. Rev. Lett.* **125**, 188003 (2020).
- [19] J. Toner and Y. Tu, Flocks, herds, and schools: A quantitative theory of flocking, *Phys. Rev. E* **58**, 4828 (1998).
- [20] E. Bertin, M. Droz, and G. Grégoire, Boltzmann and hydrodynamic description for self-propelled particles, *Phys. Rev. E* **74**, 022101 (2006).
- [21] E. Bertin, M. Droz, and G. Grégoire, Hydrodynamic equations for self-propelled particles: Microscopic derivation and stability analysis, *J. Phys. A* **42**, 445001 (2009).
- [22] T. Ihle, Kinetic theory of flocking: Derivation of hydrodynamic equations, *Phys. Rev. E* **83**, 030901(R) (2011).
- [23] T. Ihle, Invasion-wave-induced first-order phase transition in systems of active particles, *Phys. Rev. E* **88**, 040303(R) (2013).
- [24] A. Peshkov, E. Bertin, F. Ginelli, and H. Chaté, Boltzmann-Ginzburg-Landau approach for continuous descriptions of generic Vicsek-like models, *Eur. Phys. J.: Spec. Top.* **223**, 1315 (2014).
- [25] T. Ihle, Chapman-enskog expansion for the vicsek model of self-propelled particles, *J. Stat. Mech.* (2016) 083205.
- [26] Y.-L. Chou and T. Ihle, Active matter beyond mean field: Ring-kinetic theory for self-propelled particles, *Phys. Rev. E* **91**, 022103 (2015).
- [27] J. Stenhammar, C. Nardini, R. W. Nash, D. Marenduzzo, and A. Morozov, Role of Correlations in the Collective Behavior of Microswimmer Suspensions, *Phys. Rev. Lett.* **119**, 028005 (2017).
- [28] V. Škultéty, C. Nardini, J. Stenhammar, D. Marenduzzo, and A. Morozov, Swimming Suppresses Correlations in Dilute Suspensions of Pusher Microorganisms, *Phys. Rev. X* **10**, 031059 (2020).
- [29] A. Patelli, Landau kinetic equation for dry aligning active models, *Journal of Statistical Mechanics: Theory and Experiment* (2020) 033210.
- [30] R. Kürsten, S. Stroteich, M. Z. Hernández, and T. Ihle, Multiple Particle Correlation Analysis of Many-Particle Systems: Formalism and Application to Active Matter, *Phys. Rev. Lett.* **124**, 088002 (2020).
- [31] M. Ernst and E. Cohen, Nonequilibrium fluctuations in μ space, *J. Stat. Phys.* **25**, 153 (1981).
- [32] E. Leutheusser, Self-consistent kinetic theory for the Lorentz gas, *Phys. Rev. A* **28**, 1762 (1983).
- [33] H. Bussemaker, M. Ernst, and J. Dufty, Generalized Boltzmann equation for lattice gas automata, *J. Stat. Phys.* **78**, 1521 (1995).
- [34] T. Van Noije, M. Ernst, and R. Brito, Ring kinetic theory for an idealized granular gas, *Physica A* **251**, 266 (1998).
- [35] F. Peruani, A. Deutsch, and M. Bär, A mean-field theory for self-propelled particles interacting by velocity alignment mechanisms, *Eur. Phys. J.: Spec. Top.* **157**, 111 (2008).
- [36] F. Peruani, L. Schimansky-Geier, and M. Bär, Cluster dynamics and cluster size distributions in systems of self-propelled particles, *Eur. Phys. J.: Spec. Top.* **191**, 173 (2010).
- [37] F. D. C. Farrell, M. C. Marchetti, D. Marenduzzo, and J. Tailleur, Pattern Formation in Self-Propelled Particles with Density-Dependent Motility, *Phys. Rev. Lett.* **108**, 248101 (2012).
- [38] B. Liebchen and D. Levis, Collective Behavior of Chiral Active Matter: Pattern Formation and Enhanced Flocking, *Phys. Rev. Lett.* **119**, 058002 (2017).
- [39] O. Chepizhko, D. Saintillan, and F. Peruani, Revisiting the emergence of order in active matter, *Soft Matter* **17**, 3113 (2021).
- [40] S. Stroteich, Linear response and closure methods: A computer-assisted search for a quantitative theory of correlated active particle systems, Master's thesis, University of Greifswald, 2019.
- [41] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer-Verlag, Berlin, 1992).
- [42] H. Risken, *The Fokker-Planck Equation*, 2nd ed. (Springer-Verlag, Berlin, 1989).
- [43] H. D. Ursell, The evaluation of Gibbs' phase-integral for imperfect gases, *Math. Proc. Cambridge* **23**, 685 (1927).
- [44] J. E. Mayer and E. Montroll, Molecular distribution, *J. Chem. Phys.* **9**, 2 (1941).
- [45] M. Davis and P. Peebles, On the integration of the BBGKY equations for the development of strongly nonlinear clustering in an expanding universe, *Astrophys. J. Suppl. S.* **34**, 425 (1977).
- [46] S. D. M. White, The hierarchy of correlation functions and its relation to other measures of galaxy clustering, *Mon. Not. R. Astr. Soc.* **186**, 145 (1979).
- [47] T. O'Neil and N. Rostoker, Triplet correlation for a plasma, *Phys. Fluids* **8**, 1109 (1965).