A kinetic model and scaling properties of non-equilibrium clustering of self-propelled particles

Fernando Peruani\textsuperscript{1} and Markus Bär\textsuperscript{2}

\textsuperscript{1} Laboratoire J A Dieudonné, Université de Nice Sophia Antipolis, UMR 7351, CNRS, Parc Valrose, F-06108 Nice Cedex 02, France
\textsuperscript{2} Physikalisch-Technische Bundesanstalt, Abbestrasse 2-12, D-10587 Berlin, Germany
E-mail: peruani@unice.fr and makus.baer@ptb.de

New Journal of Physics 15 (2013) 065009 (15pp)
Received 3 March 2013
Published 18 June 2013
Online at \url{http://www.njp.org/}
doi:10.1088/1367-2630/15/6/065009

Abstract. We demonstrate that the clustering statistics and the corresponding phase transition to non-equilibrium clustering found in many experiments and simulation studies with self-propelled particles (SPPs) with alignment can be obtained by a simple kinetic model. The key elements of this approach are the scaling of the cluster cross-section with cluster size—described by an exponent \( \alpha \)—and the scaling of the cluster perimeter with cluster size—described by an exponent \( \beta \). The analysis of the kinetic approach reveals that the SPPs exhibit two phases: (i) an individual phase, where the cluster size distribution (CSD) is dominated by an exponential tail that defines a characteristic cluster size, and (ii) a collective phase characterized by the presence of a non-monotonic CSD with a local maximum at large cluster sizes. Through a finite-size study of the kinetic model, we show that the critical point \( P_c \) that separates the two phases scales with the system size \( N \) as \( P_c \propto N^{-\xi} \), while the CSD \( p(m) \), at the critical point \( P_c \), is always a power law such that \( p(m) \propto m^{-\gamma} \), where \( m \) is the cluster size. Our analysis shows that the critical exponents \( \xi \) and \( \gamma \) are a function of \( \alpha \) and \( \beta \), and even provides the relationship between them. Furthermore, the kinetic approach suggests that in the thermodynamic limit, a genuine clustering phase transition, in two and three dimensions, requires that \( \alpha = \beta \). Interestingly, the critical exponent \( \gamma \) is found to be in the range \( 0.8 < \gamma < 1.5 \) in line with the observations from experiments and simulations.
1. Introduction

Many experimental self-propelled particle (SPP) systems [1, 2], from actin filaments driven by molecular motors [3, 4], gliding and swimming bacteria [5–7] to active colloidal particles [8, 9], exhibit remarkably rich cluster dynamics. In particular, the formation of large moving polar clusters has been observed in most of these examples. Such clusters are formed by particles that move roughly in the same direction. Simulations of simple models of SPPs have revealed that similar clustering dynamics are observed for SPPs with both polar [10–13] and nematic [14–16], alignment interactions. More detailed models display similar clustering statistics, e.g. in simulations of self-propelled rods [17–19], self-propelled discs [20, 21], particles with hydrodynamic coupling [22] and swimming particles with flagella, e.g. in sperm cells [23].

In many of these systems, a peculiar phase characterized by the existence of remarkably large moving clusters has been observed. We designate this phase of collective motion as non-equilibrium clustering. This phase often appears as an intermediate phase between a completely disordered phase, with homogeneous density, and a phase with global orientational order. An experimental case (myxobacteria) and a simulation example (self-propelled rods) are shown in figure 1. The occurrence of large clusters is strongly correlated with a crossover in the shape of the cluster size distributions (CSD) (see figures 1(b) and (d)). Figure 2 illustrates recent findings from simulations of self-propelled hard rods [24, 25] and from simulations with a modified Vicsek model (VM) with nematic alignment [14–16]. The collective clustering phase is characterized by a non-monotonic CSD with a characteristic peak at large cluster sizes, see CSD for high densities in figure 1. At the onset of the collective clustering phase, the CSD follows a power law with a characteristic exponent. It has been observed in SPP experiments [5, 6] and simulations [17, 18] that the CSD can be power law distributed with the exponent typically varying between 0.85 and 1.35. While the range of exponents indicates the absence of a universal scaling at the onset of the collective clustering phase, it is not yet clear what determines the actual value of the exponent.

Note that the collective clustering phase can be observed in the absence of global orientational order, as reported in myxobacterial experiments [6] and self-propelled rod simulations [17] (figure 1), as well as in the presence of global orientational order [11–13, 26]. The connection between the collective clustering phase and the onset of orientational order, in a system with proven genuine long-range (orientational) order, is not known. Empirically, we observe that the onset of the collective clustering phase is typically found in between the disordered and the ordered phases, as illustrated in figure 2. We note that most investigations with
Figure 1. Examples of non-equilibrium cluster formation in self-propelled particle systems. (a) Large moving clusters in experiments with myxobacteria, where it has been found that the cluster size distribution (CSD) is a function of the cell density as shown in (b) (see [6]). (c) Snapshot of moving clusters in self-propelled rod simulations, where the CSD is known to be a function of the particle density, as illustrated in (d) (see [17]).

Figure 2. The onset of the collective (clustering) phase (blue) is often found in between a disordered and an ordered (orange) phase. The control parameter depends on the model we look at: in (a), \( \eta \) denotes the fraction of covered area, while in (b), \( \epsilon \) is angular noise. Panel (a) corresponds to simulations with self-propelled hard rods in two dimensions with aspect ratio \( \kappa = 13 \) [24, 25]. Panel (b) shows a modified Vicsek model with nematic alignment with density \( \rho = 0.25 \) [14]. The noise value \( \epsilon \) for the transition to global order was taken from a simulation with \( > 10^6 \) self-propelled particles [15], while the noise value for the onset of clustering in a system with 16,384 self-propelled particles is based on the results of [16].

New Journal of Physics 15 (2013) 065009 (http://www.njp.org/)
a focus on the cluster statistics (both in experiments and simulations) were carried out with small
to intermediate numbers (around $10^2$–$10^3$) of self-propelled particles, whereas investigations on
the onset of global order were conducted with rather large numbers ($>10^5$). Hence, little is
known about the onset of the collective clustering phase in systems (and simulations) with a
large number of particles.

Here, we focus on the emergence of the collective clustering phase in the absence of global
order, i.e. we assume that cluster velocities are uncorrelated. We propose a kinetic clustering
theory based on the approach presented by us in [17], and extend our earlier treatment to
coagulation and fragmentation kernels that depend, respectively, on the scaling of the cluster
cross-section with cluster ‘mass’, i.e. the number of particles in the cluster, and the scaling of
cluster perimeter with cluster mass. The scaling of the cluster cross-section and of the cluster
perimeter are characterized by exponents $\alpha$ and $\beta$, respectively. Furthermore, we present a
comprehensive analysis of this kinetic cluster model, including a finite size (FS) study of it.

The FS study of the kinetic model reveals that: (i) the transition to the collective clustering
phase—characterized by a non-monotonic CSD—is generic to all SPPs with either a polar or a
nematic alignment mechanism, and (ii) there exists a critical asymptotic CSD exponent $\gamma$ which
is a function of $\alpha$ and $\beta$ only. By studying the physically meaningful parameter space $\alpha$–$\beta$,
we found that $\gamma$ always falls in the range $0.8 < \gamma < 1.5$. Note that in the absence of cluster–cluster
correlations, this transition marks the onset of collective motion as defined in [6]. Although the
simplified clustering theory is, strictly speaking, only valid in the absence of global order, we
show through simulations—using a large number of particles—that a comparable picture holds
when global order is observed.

This paper is organized as follows. In section 2 we present our simple kinetic clustering
model under the assumption that there are no cluster–cluster correlations, and show that for any
finite system there are two clustering phases. We perform a system size analysis of the kinetic
model in section 3, where we report on the scaling properties of the obtained CSD, which we
found to depend on the scaling of the cluster cross-section and cluster perimeter. In section 4 we
study cluster formation in an SPP model that is known to exhibit cluster–cluster correlations,
and contrast the obtained results with our cluster–cluster uncorrelated clustering theory. The
implications and limitations of the proposed kinetic approach are discussed in section 5.

2. A kinetic model of clustering

We look for a description of the clustering process in terms of $\langle n_1(t) \rangle$, $\langle n_2(t) \rangle$, . . . , $\langle n_N(t) \rangle$,
where $\langle n_i(t) \rangle$ represents the average value at time $t$ of clusters formed by $i$ particles and $N$ the
number of particles in the system. To ease the notation, $\langle n_1(t) \rangle$ is hereafter referred to as $n_1(t)$.

Our strategy consists in deriving a coagulation Smoluchowski equation with fragmentation
for these $n_i(t)$ objects. For an introduction to this kind of coagulation equation, see [27]. In
SPP systems, the only conserved quantity is the number of particles $N$. Neither the overall
orientational order nor the number of clusters is conserved. Thus, our generalized coagulation
Smoluchowski equation should conserve the number of particles. We start by simplifying the
clustering dynamics. Because we are dealing with self-propelled particles, we assume that
clusters are the result of either an explicit or an effective alignment mechanism such that
particles inside a cluster move coherently in the same direction. This implies that clusters
move at speeds comparable with that of the individual particles, independently of the size
of the cluster. Below we also discuss what can be expected if this condition is relaxed. In

New Journal of Physics 15 (2013) 065009 (http://www.njp.org/)
contrast, clusters of passive particles, driven by thermal fluctuations, are such that their mobility decreases as a function of their size. We stress that the assumption of a size-independent cluster speed evidences the non-equilibrium nature of our simple cluster theory. We further assume that (binary) collisions among clusters may result in cluster–cluster fusion, and neglect the possibility of cluster fragmentation induced by cluster–cluster collisions. This assumption is justified as long as the dynamics is overdamped, and is in accordance with what is observed in experiments with gliding bacteria [6] and simulations with self-propelled rods [17]. We simplify the cluster fragmentation dynamics by assuming an evaporation-like process by which clusters shrink in size by losing one by one those particles that are on the cluster boundary. In summary, we are assuming an irreversible clustering process in which a cluster of mass \( j \) can undergo the following ‘reactions’:

\[
C_j + C_k \xrightarrow{A_{j,k}} C_{j+k}, \quad C_j \xrightarrow{B_j} C_{j-1} + C_1. \tag{1}
\]

Note that because the reaction \( C_{j+k} \rightarrow C_j + C_k \) does not occur, the process is, in this sense, irreversible. Now, we look for a description of the process in terms of the (average) number \( n_j(t) \) of clusters with \( j \) particles (i.e. the number of \( C_j \)) at time \( t \), whose time evolution takes the form

\[
\dot{n}_1 = 2B_2n_2 + \sum_{k=3}^{N} B_k n_k - \sum_{k=1}^{N-1} A_{k,1} n_k n_1,
\]

\[
\dot{n}_j = B_{j+1}n_{j+1} - B_j n_j - \sum_{k=1}^{N-j} A_{k,j} n_k n_j + \frac{1}{2} \sum_{k=1}^{j-1} A_{k,j-k} n_k n_{j-k}, \quad \text{for } j = 2, \ldots, N-1,
\]

\[
\dot{n}_N = -B_N n_N + \frac{1}{2} \sum_{k=1}^{N-1} A_{k,N-k} n_k n_{N-k},
\]

where the dot denotes the time derivative, \( B_j \) represents the rate at which a cluster of mass \( j \) loses particles, defined as

\[
B_j = \frac{D}{d^2} j^\beta, \tag{3}
\]

and \( A_{j,k} \) is the collision rate between clusters of mass \( j \) and \( k \), defined by

\[
A_{j,k} = \frac{q_s v_c \sigma_0}{L^2} \left( j^\alpha + k^\alpha \right), \tag{4}
\]

where \( L^2 \) is the area of the two-dimensional (2D) space where particles move and \( v_c \) is the cluster speed, which we assume to be \( v_c \sim v_0 \) with \( v_0 \) the speed of individual particles. Note that equations (2) are such that \( \sum m n_m = 0 \) and thus the number of particles \( N = \sum m n_m(t) \) is conserved. In equation (3), \( d^2/D \) is the typical time a particle located on the cluster boundary needs to detach from a cluster, with \( d \) the maximum distance two particles can be apart to still be considered as connected and \( D \) the diffusion coefficient with respect to the center of mass of the cluster. We stress that in equation (2) we have assumed that clusters are uncorrelated. Several statistical features of the model can be given explicitly if one assumes that the SPPs follow a dynamics where their translation is determined by a constant speed and the direction of
motion is subject to alignment interactions with neighboring SPPs and an angular noise with amplitude $\eta$. An example of such dynamics is given, in two dimensions, by the following equations of motion:

$$
\theta_j^{t+1} = \arg \left[ \sum_k f \left( e^{i\delta_{k,j}}, e^{i\theta_j^t} \right) \right] + \eta_j^t,
$$

$$
\mathbf{x}_j^{t+1} = \mathbf{x}_j^t + v_0 e^{i\theta_j^t},
$$

where $j$ is the index of the particle, $\theta_j^t$ defines (in two dimensions) the particle moving direction, $\mathbf{x}_j^t$ denotes the position at time $t$, the sum is taken over all particles within a unit distance of $j$, which includes $j$, and $\eta_j^t$ is a uniformly distributed random variable such that $-\eta/2 < \eta_j^t < \eta/2$, $\langle \eta_j^t \rangle = 0$ and $\langle \eta_j^t \eta_{k}^t \rangle = (\eta^2/12)\delta_{j,k}$. The dynamics given by equation (5) is frequently used in simulation studies, e.g. in [10–16, 28, 29] (for a review see [1]). More specifically, $f(e^{i\delta_{k,j}}, e^{i\theta_j^t}) = e^{i\theta_j^t}$ defines a polar alignment rule as in the VM [10]. $f(e^{i\delta_{k,j}}, e^{i\theta_j^t}) = \text{sign}(\cos(\theta_j^t - \theta_k^t))e^{i\theta_j^t}$ defines SPP with a nematic alignment rule as used [14, 15]. Assuming that the SPP obey equation (5), the diffusion around the center of mass of a polar cluster of SPP can be approximated, as detailed in [16], by $D = v_0^2 \left[ 1 - (\sin(\eta/2)/(\eta/2))^2 \right] \Delta t$, where $v_0$ is the speed of individual particles and $\Delta t$ sets the discrete time step for typical SPP dynamics such as the one given by equation (5). By making a Taylor expansion we find that $D \sim v_0^2/12\eta^2 \Delta t$. Note that for non-interacting random walkers subject to the same kind of noise, $D \propto v_0^2/\eta^2$. In addition, we have to consider that the splitting rate $B_j$ has to be proportional to the number of particles on the cluster boundary. In equation (3), we have assumed that the perimeter $l$ of a cluster scales with its mass $j$ as $l \propto j^\beta$, with $\beta$ constrained to $1/2 \leq \beta \leq 1$. Let us provide a physical context to this assumption, always assuming that we are in a 2D space. Surface tension would tend to minimize cluster perimeters, and so clusters would be round and $\beta = 1/2$, as observed in liquid–vapor drops [30]. On the other hand, for random deposition of particles as in classical percolation, $\beta \approx 1$. If clusters are ‘chains’ of particles, then also $\beta = 1$. In short, we expect $0.5 \leq \beta < 1$, with $\beta = 0.5$ representing the most compact way we can arrange $j$ particles to minimize the number of particles at the cluster interface, and $\beta = 1$ representing the opposite extreme, i.e. a situation where all particles are at the cluster boundary.

The collision rate $A_{j,k}$ is derived in analogy to the collision rate in kinetic gas theory [31], which means that we assume that clusters move ballistically in between collisions, which leads to a collision rate proportional to the sum of the scattering cross-section of the involved clusters, the relative cluster speed and the cluster density. Due to the latter assumption, $A_{j,k} \propto L^{-2}$. The relative cluster speed can be assumed to be not too different from that of the individual clusters, $v_c$. In SPP systems, we often find that the individual cluster speed is close to that of individual particles, i.e. $v_c \sim v_0$. If $v_c$ exhibits a dependence on cluster size of the form $v_c \propto j^{-\omega}$, the exponent $\omega$ can be absorbed into $\alpha$. For simplicity, here we assume that $v_c \sim v_0$. For instance, the speed of SPP clusters can be approximated by $v_c = v_0 \sin(\eta/2)/(\eta/2) \propto v_0 (1 - \eta^2)$ [16]. The scattering cross-section of a 2D cluster is the (average) projection of the cluster on a given axis. Thus, the scattering cross-section has to be always less than or equal to the cluster perimeter. If the scattering cross-section of a cluster scales as $\propto \sigma_0 j^\alpha$, the previous observation implies that $\alpha \leq \beta$ is the only meaningful physical scenario. Below, we will see that there are two qualitatively very different physical scenarios, $\alpha = \beta$ and $\alpha < \beta$. Finally, $q_s$ represents the probability that a cluster–cluster collision resulted in a successful fusion of clusters. Here, we
assumes it is a constant but certainly it can depend on various (intensive) variables of the actual system. In particular, it depends on the alignment symmetry. A very rough assumption would be to assume that if $q_s = q_0$ for the ferromagnetic alignment, $q_s = q_0/2$ for the nematic one. Equations (2) are scaled and transformed into dimensionless form by dividing equations (3) and (4) by $D/d^2$, which leads to the following dimensionless parameter:

$$P = \frac{q_s \sigma_0 d^2 v_c}{L^2 D}.$$  

Thus, equation (3) reduces to $B_j = j^\beta$, and equation (4) to $A_{j,k} = P (j^\alpha + k^\alpha)$. The parameter $P$ controls the relative weight of fragmentation with respect to coagulation. For the 2D SPP model type considered here, we obtain $P \propto (L \eta)^{-2}$. Note that while in a three-dimensional space we can expect $A_{j,k}$ and $B_j$ to follow a similar scaling as the one proposed above, one-dimensional flocking models are essentially different, exhibiting a homogeneous density phase and a condensed phase [32]. Finally, computation of the critical point $P$ can be done, alternatively, by directly studying the stability of the individual phase and ignoring the actual shape of the CSD, as recently proposed in [33].

Except for a few special cases, for instance $\alpha = \beta = 0$, equation (2) cannot be solved analytically. Nevertheless, direct numerical integration of equation (2), for instance through a fourth-order Runge–Kutta method, can be easily implemented. Note that the numerical implementation has to ensure particle number conservation at every integration step, which often requires a renormalization checkpoint in the integration loop. We integrate the dimensionless version of equations (2), with the initial condition $n_j(t = 0) = N \delta_{j,1}$, and find that the (weighted) CSD, defined as

$$p(m, t) = \frac{m n_m(t)}{N},$$

reaches a steady state, i.e. $p(m, t \to \infty) = p(m; P)$. This also implies that the number of clusters in the system, $M(t) = N \sum m^{-1} p_m(t)$, also reaches a steady-state value as $t \to \infty$. In the literature the term CSD is frequently used to refer to $\tilde{p}(m, t) = n_m(t)/M(t)$. The disadvantage of this definition is that its normalization constant, $M(t)$, varies with time. While $\tilde{p}(m)$ refers to the probability of finding a cluster of size $m$, $p(m)$ indicates the probability of a randomly selected particle to be in a cluster of size $m$. Given $\alpha$ and $\beta$, the CSD $p(m)$ depends on $N$ and the value of $P$. Figure 3 summarizes the clustering behavior with $P$ for a given $N$. Note that there exists a critical $P_c(N)$ that separates two different types of clustering behavior. For small values of $P$, the distribution $p(m)$ is monotonically decreasing, dominated by an exponential tail. As $P \to P_c$, $p(m)$ approaches a power law, with a system size cutoff. For $P > P_c$, $p(m)$ is non-monotonic and exhibits a peak at large cluster sizes. This dramatic change of behavior at $P_c$ unveils a phase transition, which is evidenced by figures 3(c) and (d) that show $\langle m \rangle = \sum m \ p(m)$ and $\sigma^2 = \sum (m - \langle m \rangle)^2 \ p(m)$, respectively. Below, we will see that this transition is a genuine phase transition—in the thermodynamical sense—only for $\alpha = \beta$.

3. Scaling properties

To study the asymptotic behavior of the CSD, equation (7), we need to solve the set of $N$ coupled equations given by equation (2). These equations depend on the dimensionless parameter $P$, the
Figure 3. For a given system size $N$, the steady-state CSD $p(m)$ is monotonically decreasing for $P < P_c$ (a), while for $P > P_c$ the distribution exhibits a peak at large cluster sizes (b). At $P = P_c$ the distribution $p(m)$ is a power law. The transition between these two types of behavior, (a) and (b), can be represented in a more phase-transition-like form, by first looking at two moments of the distribution $p(m)$, (c) and (d), respectively. The other parameters are: $N = 512$ and $\alpha = \beta = 1/2$.

Kernel exponents $\alpha$ and $\beta$, as well as $N$, i.e. the involved number of equations in equation (2). We conduct a FS study of equation (2) to extrapolate the behavior of $p(m)$ for $N \to \infty$. Let us recall that the thermodynamical limit is given by $N \to \infty$ as well as $L \to \infty$, with the constraint that $\rho = N/L^2$ is a constant. In the previous version we have learned that $P \propto L^{-2}$. In this section, we show that the critical point $P_c(N; \alpha, \beta)$—defined in the previous section—depends on $N$. For every $N$, we numerically estimate $P_c(N)$, using a bisection method, as the point at which the CSD $p(m)$ is no longer monotonically decreasing. We find that at the critical point $P_c(N)$, $p(m)$ scales as

$$p(m; P_c) \propto m^{-\gamma},$$

where $\gamma$ is a critical exponent that depends on $\alpha$ and $\beta$. This finding results from the evident collapse of the various system size curves at the critical point shown in figure 4(a). The figure indicates that for a given $N$, $p(m; P_c)$ follows the scaling given by equation (8) up to a certain cluster size $m^*(N)$ above which the FS of the system becomes evident. The exponent $\gamma$ results from extrapolating the behavior for $N \to \infty$ and it is numerically obtained by minimizing the error defined as $W(\gamma; N) = \sum_m m^\gamma p(m; P_c)$. This means that to obtain $\gamma$ as a function of $\alpha$ and $\beta$ (figure 5), for every pair $(\alpha, \beta)$ a FS study, such as the one described above, has been realized. This implies first the estimation of $P_c$ for every system size $N$ before computing the actual value of $\gamma$. Below the critical point, the following scaling is obeyed by $p(m)$:

$$p(m; N, P, \alpha, \beta) = K p(m; bN, b^{-\xi} P, \alpha, \beta),$$

where $K$ is a constant. This is confirmed by the obtained collapse of the various system size curves where $P$ was rescaled as indicated in equation (9), see figure 4(b). Below we explain that $\xi$ is connected to the scaling of $P_c$ with $N$. For $\alpha = \beta$, we find that $K = 1$, while for
Figure 4. FS study of the steady state CSD $p(m; P)$. The critical point $P_c(N)$ is itself a function of the system $N$, equation (10). The CSD at the critical point, i.e. $p(m; P = P_c(N))$, approaches a power law—see equation (8)—as $N \to \infty$ as indicated by the collapse of the various system size curves shown in (a). Below the critical point $P_c(N)$, the FS scaling of $p(m; P)$ is given by equation (9) as illustrated by the collapse of the curves corresponding to $P < 2.50 \times 10^{-3}$, $1.25 \times 10^{-3}$ and $6.25 \times 10^{-4}$. Above the critical point $P_c(N)$, the scaling given by equation (9) leads to a collapse for small $m$ values, while the position of the peak and its width are observed to scale with $N$, see the curves for $P = 4 \times 10^{-2}$, $2 \times 10^{-2}$ and $1 \times 10^{-1}$ and the definition of $Q(N)$ in the text. Panels (a) and (b) correspond to $\alpha = \beta = 1/2$. For the value of $\gamma$ in panel (a), see figure 5.

$\alpha \neq \beta$, $K$ depends on $b$, $N$ and $\beta$. All this means that if, for instance, we double the system size and reduce $P$ by $2^{-\xi}$, we fall on the same distribution $p(m)$ for $\alpha = \beta$, while for $\alpha < \beta$ there is a multiplicative constant $K$. For $P < P_c$, we find that the CSD is well fitted by $p(m) \sim m^{-\gamma} \exp(-m/\tilde{m}(P))$.

For $P > P_c$, we observe that there is a shift with $N$ of the peak that emerges at large cluster sizes (figure 3(b)). The scaling given by equation (9) works up to a given size $m_m(N)$ which is given by the minimum of $p(m)$. We are interested in knowing the behavior of the peak with the system size $N$. To answer these questions, we study the scaling of $Q(N) = \sum_{m_m(N)} p(m)$. We find that for $\alpha = \beta$, $Q(N \to \infty) \to 1$. This indicates the presence of a collective phase in which most of the particles are part of large clusters. Note that this does not mean the formation of a single giant cluster; the width of the peak does not shrink to zero. For $\alpha < \beta$, $Q$ does not converge to 1 as $N \to \infty$. This is because for $\alpha < \beta$, in the limit of $N \to \infty$ the transition does not occur. Nevertheless, for any finite $N$, we always observe the transition to the collective phase. This observation can be understood by studying the system size behavior of $P_c(N)$. We
find that
\[ P_c(N) \propto N^{-\xi} \]  
with \( \xi = 1 \) for \( \alpha = \beta \) and \( \xi < 1 \) for \( \alpha < \beta \). For a fixed value of \( \alpha, \xi \) becomes a function of \( \beta \) as illustrated in the inset of figure 5(b) for \( \alpha = 1/2 \). In the thermodynamical limit, with \( N \to \infty \) and \( L \to \infty \), while \( \rho = N/L^2 \) is constant, we can estimate the critical point. Combining equations (6) and (10) we obtain, for a given set of parameters \( D, v_0, \sigma_0 \) and \( d \), the critical density \( \rho_c \) above which a peak at large cluster sizes emerges:

\[ \rho_c = N^{1-\xi} \frac{D}{v_c \sigma_0 d^2}. \]  

From equation (11) it is clear that when \( \xi < 1 \), \( \rho_c \) diverges with the system size \( N \). Only for \( \xi = 1 \) is \( \rho_c \) a finite quantity in the thermodynamical limit. Thus, for \( \alpha = \beta \) there is a critical density \( \rho_c \) above which the collective transition occurs. In contrast, for \( \alpha < \beta \), \( \rho_c \to \infty \) and any finite density \( \rho_0 \) corresponds to the individual phase. Equation (11) predicts the critical noise intensity \( \eta_c \) below which one expects the collective clustering phase to occur as \( \eta_c \propto \sqrt{\rho N^{\xi-1}} \), from which it follows that for \( \xi = 1 \), the critical noise scales as \( \eta_c \propto \sqrt{\rho} \). Interestingly, a similar scaling was reported for the onset of collective motion in simulations with the VM [10, 26, 29].

4. Clustering in the presence of global orientational order

The simple set of Smoluchowski-type equations (2) has been derived assuming that clusters move in an uncorrelated fashion, as observed for instance in bacterial experiments [6]. However,
Figure 6. Simulation snapshots of the Vicsek model (VM) for various noise intensity values. Particles belonging to the same cluster are labeled with the same color. As can be seen in figure 7, for $\eta = 1.0$ the system is disordered, while for $\eta < 1$ there is a preferred direction of motion. Although the snapshots for $\eta = 1.0$ and $\eta = 0.9$ seem comparable, their CSD is remarkably different: an exponential distribution for the former and a power law for the latter. Traveling bands in the VM are composed of many correlated clusters ($\eta = 0.7$). Finally, for very low $\eta$ values, most particles form part of very large clusters ($\eta = 0.05$).

several SPP models exhibit long-range orientational order and macroscopic structures such as bands, which indicates that there are strong correlations among cluster velocities. Thus, \textit{a priori} we cannot expect equation (2) to describe quantitatively the cluster dynamics of such systems deep in phases with strong cluster–cluster correlations. Interestingly, it has already been observed that the VM in its (global) ordered phase displays CSDs that can be power law distributed, with $\gamma$ roughly in the range $0.8 < \gamma < 1.3$, according to [11, 12, 26]. Here, we carried out simulations in the VM, i.e. equations (5) with a choice of $f = e^{i\theta_k}$. Our simulation results confirm that indeed the VM exhibits power law distributed CSDs with exponents in this range, and show that this occurs close to the well-known disorder–order transition of the VM. Furthermore, we found that the cluster statistics in the VM, below and close to the critical point—i.e. when cluster–cluster correlations are weak or absent—resembles that obtained with the kinetic clustering model. As expected, far away from the critical point and well in the ordered phase, cluster–cluster correlations induce effects that cannot be accounted for by the kinetic clustering model. In short, the simulation data strongly suggest that close to the disorder–order critical point, the VM exhibits a transition to a collective clustering phase like the one described with the kinetic model.

The order–disorder transition in the VM is observed when the noise intensity $\eta$ is decreased below a critical value $\eta^*$. The orientational order is characterized by $\phi = |\sum_j \exp(i\theta_k^j)|/N$, which defines the usual polarization order parameter using complex number notation [10]. In the disorder phase, the clusters cover the space homogeneously (figure 6) and the CSD is dominated by an exponential tail, see figure 7. If $\eta$ is decreased below $\eta^*$, velocity–velocity correlations among particles become important, $\phi$ increases and clusters grow significantly in size. Below the critical point $\eta^*$, the CSD is a power law with an exponent in the range $[0.8, 1.3]$, which results also in a change in behavior of $\langle m \rangle / N$, as shown in figure 7. The power law distributed CSD indicates that the system displays arbitrarily large clusters. Indirectly, this also means that particle–particle correlations are long-ranged, i.e. arbitrarily large like the cluster sizes. Note
that this does not necessarily imply the existence of cluster–cluster correlations. Nevertheless, in the VM at low values of the noise intensity, cluster–cluster correlations become evident. Arguably, due to these cluster–cluster correlations, we observe a clear deviation from what the theory predicts for \( \eta \ll \eta^* \). Moreover, in the VM, \( \gamma \) is a function of \( \eta \). We observe that close to \( \eta^* \), \( \gamma \approx 1.3 \), and that when \( \eta \) is decreased, \( \gamma \) approaches 1. It is close to \( \eta^* \) that we observe the emergence of a traveling band. Note that this band is not a connected component, but a cloud of highly correlated moving clusters. We find that the CSD is power law distributed in the band regime. For \( \eta \ll \eta^* \) the system is highly ordered and clusters move roughly in the same direction, which means that cluster collisions are less frequent. On the other hand, low values of \( \eta \) also imply that spreading of a cluster around its center of mass is also small. In summary, at low noise values the cluster dynamics slows down significantly in the VM, which results in a non-monotonic response of \( \langle m \rangle / N \) with \( \eta \) (figure 7). Interestingly, we observe that at extremely low \( \eta \) values the CSD decreases even more slowly than \( 1/m \) (see figure 6).

5. Concluding remarks

We have argued that SPP systems exhibit for a finite system size two phases: an individual phase, where the CSD is dominated by an exponential tail, and a collective (clustering) phase characterized by a non-monotonic CSD. Assuming that the moving directions of clusters are uncorrelated, equation (2) justifies the existence of these two phases. At the transition
point, the CSD is a power law characterized by an exponent $\gamma$ that depends, according to the proposed kinetic clustering model, on the scaling with cluster size of the cluster cross-section—characterized by an exponent $\alpha$—and the cluster perimeter—characterized by an exponent $\beta$. A systematic study of the $\alpha$–$\beta$ parameter space revealed that $\gamma$ always falls in the range between 0.8 and 1.5, which is consistent with experimental observations [5, 6] and simulations [17, 18]. However, we learn that in the thermodynamical limit, the above-mentioned transition occurs only for $\alpha = \beta$, while otherwise, the system remains in the individual phase. It is worth pointing out that for the special case $\alpha = \beta$ the kinetic model predicts a critical density value that is independent of the system size. Hence, the properties for the transition to non-equilibrium clustering found in the kinetic model for intermediate values of $N$ hold true in the thermodynamic limit.

In addition to the study of the kinetic clustering theory given by equation (2), we characterized the cluster statistics of the VM [10]. Despite the fact that cluster–cluster correlations and global order emerge in the simulations of the VM that are not treated in equation (2), we found that a transition to a collective clustering phase is also present in the VM. Moreover, $\gamma$ falls in the expected range if simulations are performed with ferromagnetic alignment, as shown here and in [11, 12, 26], as well as with nematic alignment [16]. Notably, we found that the CSD is power law distributed for a range of noise intensity values $\eta$ close to the order–disorder critical point, with $\gamma$ a function of $\eta$. In summary, the existence of two clustering phases is found as well in SPP systems in the presence or absence of global orientational order.

Our kinetic approach is complementary to the ongoing effort at deriving coarse-grained nonlinear field equations for the description of active matter [34–42] (for a review, see [2, 43]) that are better suited to capture large-scale structures or the emergence of global orientational order in large systems. In contrast, the observed clustering phenomena are typically studied in the absence of long-range order, on a small scale, and for intermediate system sizes. It is also worth noting that the non-equilibrium clustering phase leads to (apparent) giant number fluctuations (GNF) [6], which often have similar properties as the ones predicted by Toner and Tu [34] in the phase with long-range orientational order. The kinetic clustering theory indicates that the CSD can be power law distributed, in the thermodynamical limit, in the absence of global orientational order. Since a power law distributed CSD necessarily leads to apparent GNF, the theory supports recent observations regarding the presence of GNF in the absence of orientational order, see [6, 45]. Moreover, in section 4 we showed that in large-scale simulations, the CSD exhibits fat tails in the presence of global order. This certainly leads again to GNF. While the presence of fat tails in the CSD leads to apparent GNF, the contrary is not necessarily true. A solid connection between GNF and fat tails in the CSD remains to be established.

Finally, we expect a similar collective (clustering) transition and cluster dynamics to occur in all SPP systems where the particle speed is not strongly affected by the local density. The presence of a density-dependent speed, as in [44–47], can dramatically change the above-given clustering picture, since in these models, large clusters, when they are not strongly aligned, are prone to slowing down significantly. As result of this effect, the SPPs can form—when the alignment mechanism is either weak or inexistent—a single large cluster which coexists with a background gas of particles in the individual phase, as found in [44–47]. The cluster dynamics of these systems is likely to be related to an equilibrium-like phase separation, as suggested in [44, 46].
Acknowledgments

We thank H Chaté, F Ginelli and J Toner for useful discussions. MB acknowledges financial support from DFG through GRK 1558.

References

[1] Vicsek T and Zafeiris A 2012 Phys. Rep. 517 71
[2] Marchetti M C, Joanny J F, Ramaswamy S, Liverpool T B, Prost J, Rao M and Simha R A 2012 arXiv:1207.2929
[3] Schaller V, Weber C, Semmrich C, Frey E and Bausch A 2010 Nature 467 73
[4] Köhler S, Schaller V and Bausch A 2011 Nature Mater. 10 462
[5] Zhang H, Be’er A, Florin E-L and Swinney H 2010 Proc. Natl Acad. Sci. USA 107 13526
[6] Peruani F, Starruss J, Jakovljevic V, Sogaard-Andersen L, Deutsch A and Bär M 2012 Phys. Rev. Lett. 108 098102
[7] Starruss J, Peruani F, Jakovljevic V, Sogaard-Andersen L, Deutsch A and Bär M 2012 Interface Focus 2 774
[8] Theurkauff I, Cotton-Bizzone C, Palacci J, Ybert C and Bocquet L 2012 Phys. Rev. Lett. 108 268303
[9] Palacci J, Sacanna S, Steinberg A, Pine D and Chaikin P 2013 Science 339 936
[10] Vicsek T, Czirok E A, Jacob E B, Cohen I and Shochet O 1995 Phys. Rev. Lett. 75 1226
[11] Huepe C and Aldana M 2004 Phys. Rev. Lett. 92 168701
[12] Huepe C and Aldana M 2008 Physica A 387 2809
[13] Romensky M and Lobaskin V 2013 arXiv:1301.6294v1
[14] Peruani F, Deutsch A and Bär M 2008 Eur. Phys. J. Spec. Top. 157 111
[15] Ginelli F, Peruani F, Bär M and Chaté H 2010 Phys. Rev. Lett. 104 184502
[16] Peruani F, Schimansky-Geier L and Bär M 2010 Phys. Rev. E 84 030904
[17] Yang Y, Marceau V and Gompper G 2010 Phys. Rev. E 82 031904
[18] Weber C, Schaller V, Bausch A and Frey E 2012 Phys. Rev. E 86 030901
[19] Grossman D, Aranson I and Ben-Jacob E 2008 New J. Phys. 10 023036
[20] Deseigne J, Dauchot O and Chaté H 2010 Phys. Rev. Lett. 105 098001
[21] Llopis I and Pagonabarraga I 2006 Europhys. Lett. 75 999
[22] Yang Y, Elgeti J and Gompper G 2008 Phys. Rev. E 78 061903
[23] Wensink H H, Dunkel J, Heidenreich S, Drescher K, Goldstein R E, Löwen H and Yeomans J M 2012 Proc. Natl Acad. Sci. USA 109 14308
[24] Wensink H H and Löwen H 2012 J. Phys.: Condens. Matter 24 464130
[25] Chaté H, Ginelli F, Grégoire G and Raynaud F 2008 Phys. Rev. E 77 046113
[26] Krapivsky P L, Redner S and Ben-Naim E 2010 A Kinetic View of Statistical Physics (New York: Cambridge University Press)
[27] Baglietto G and Albano E 2009 Phys. Rev. E 80 050103
[28] Peruani F, Ginelli F, Bär M and Chaté H 2011 J. Phys.: Conf. Ser. 297 012014
[29] Stauffer D 1979 Phys. Rep. 54 1
[30] Reif F 1965 Fundamentals of Statistical and Thermal Physics (Singapore: McGraw-Hill)
[31] J.O’Loan O and Evans M R 1999 J. Phys. A: Math. Gen. 32 99
[32] Weber C, Thueroff F and Frey E 2013 New J. Phys. 15 045014
[33] Toner J and Tu Y 1998 Phys. Rev. E 58 4828
[34] Bertin E, Droz M and Gregoire G 2006 Phys. Rev. E 74 022101
[35] Mishra S, Baskaran A and Marchetti M 2010 Phys. Rev. E 81 061916
[36] Ihle T 2011 Phys. Rev. E 83 030901
[37] Grossmann R, Schimansky-Geier L and Romanczuk P 2012 New J. Phys. 14 073033

New Journal of Physics 15 (2013) 065009 (http://www.njp.org/)
[39] Dunkel J, Heidenreich S, Bär M and Goldstein R 2012 arXiv:1208.4464
[40] Gopinath A, Hagan M, Marchetti M and Baskaran A 2012 Phys. Rev. E 85 061903
[41] Peshkov A, Aranson I, Bertin E, Chaté H and Ginelli F 2012 Phys. Rev. Lett. 109 268701
[42] Dunkel J, Heidenreich S, Bär M and Goldstein R 2013 New J. Phys. 15 045016
[43] Ramaswamy S 2010 Annu. Rev. Condens. Matter Phys. 1 323
[44] Peruani F, Klauss T, Deutsch A and Voss-Boehme A 2011 Phys. Rev. Lett. 106 128101
[45] Fily Y and Marchetti M C 2012 Phys. Rev. Lett. 108 235702
[46] Farrell F D C, Marchetti M C, Marenduzzo D and Tailleur J 2012 Phys. Rev. Lett. 108 248101
[47] McCandlish S R, Baskaran A and Hagan M F 2012 Soft Matter 8 2527