

Available online at www.sciencedirect.com





Physica A 387 (2008) 2809-2822

www.elsevier.com/locate/physa

# New tools for characterizing swarming systems: A comparison of minimal models

Cristián Huepe<sup>a,\*</sup>, Maximino Aldana<sup>b</sup>

<sup>a</sup> 614 N Paulina Street, Chicago IL 60622-6062, USA <sup>b</sup> Instituto de Ciencias Físicas, Universidad Nacional Autónoma de México, Apartado Postal 48-3, Cuernavaca, Morelos 62251, Mexico

> Received 18 September 2007; received in revised form 21 December 2007 Available online 20 January 2008

### Abstract

We compare three simple models that reproduce qualitatively the emergent swarming behavior of bird flocks, fish schools, and other groups of self-propelled agents by using a new set of diagnosis tools related to the agents' spatial distribution. Two of these correspond in fact to different implementations of the same model, which had been previously confused in the literature. All models appear to undergo a very similar order-to-disorder phase transition as the noise level is increased if we only compare the standard order parameter, which measures the degree of agent alignment. When considering our novel quantities, however, their properties are clearly distinguished, unveiling previously unreported qualitative characteristics that help determine which model best captures the main features of realistic swarms. Additionally, we analyze the agent clustering in space, finding that the distribution of cluster sizes is typically exponential at high noise, and approaches a power-law as the noise level is reduced. This trend is sometimes reversed at noise levels close to the phase transition, suggesting a non-trivial critical behavior that could be verified experimentally. Finally, we study a bi-stable regime that develops under certain conditions in large systems. By computing the probability distributions of our new quantities, we distinguish the properties of each of the coexisting metastable states. Our study suggests new experimental analyses that could be carried out to characterize real biological swarms.

PACS: 64.60.-i; 05.65.+b; 05.70.Ln; 89.75.-k

Keywords: Swarms; Self-propelled particle; Swarming models; Collective motion; Phase transitions; Non-equilibrium systems

## 1. Introduction

For more than two decades, there has been a continuing interest in finding simple models that can describe the collective motion of groups of self-propelled agents [1–8]. In nature, these systems typically correspond to swarms of living organisms, such as bird flocks, fish schools, or herds of quadrupeds [9–13]. While their specific biological details can be very different and quite intricate, in all of these examples the agents are able to achieve a surprising amount of coordination in their collective displacements, in spite of having no apparent leader or long-range communication

\* Corresponding author. Tel.: +1 312261 5566.

E-mail address: cristian@northwestern.edu (C. Huepe).

URL: http://www.esam.northwestern.edu/~cristian/ (C. Huepe).

mechanism. At the origin of this swarming behavior is the ability that groups of agents have to, under certain circumstances, align and start heading in a common direction. Various models have been introduced to study this phenomenon [1,4,7,8,14–16]. Among the simplest is the one originally described by Vicsek et al. in Ref. [4]. It was in this model that numerical computations first showed the emergence of a dynamical phase transition from a disordered state (in which agents move in random directions) to an ordered one (where they head in approximately the same direction) as the noise level is decreased or the mean density is increased. More recently, other simple models with qualitatively similar agent dynamics have been introduced [7,14–16]. Typically, they all describe groups of agents moving with an imposed non-zero speed and that tend to align with their short-range neighbors. They also focus on similar order parameters to measure the amount of self-organization in the system, namely: the degree of alignment of the agents' headings. When noise is introduced, all these models display a transition from a disordered to an ordered state, analogous to the one observed in the original model by Vicsek et al. The similarity of these results presents a challenge, since we are left with no easy way to discriminate which minimal description is best at capturing the basic swarming mechanisms. Indeed, while a transition to a state where agents move in a common direction is an essential part of any credible swarming description, this appears to be the generic result of all algorithms that mimic a reasonable swarm-like behavior.

In this work we introduce a set of global quantities that allow us to better characterize the collective states resulting from different swarming dynamics. We compare three commonly used simple swarming models and find that, while the standard order parameters (that measure alignment) behave equivalently in all cases, our newly introduced quantities provide a more detailed description that can clearly distinguish their properties. These quantities can also be measured experimentally, allowing further comparison to real swarms. It is important to point out, however, that there are currently very few experiments that can measure the relative position of all agents in a swarm, which is required by our averaged quantities. Fortunately, the recent implementation of the experiments described in Refs. [30,32–34], together with a growing interest in swarming systems by communities in Biology, Physics, and Engineering, makes us expect more sets of swarming data to be obtained in the near future. It is unclear at this point, though, how these data sets will connect to our numerical calculations. The data sets recently obtained in Ref. [30], for example, seem to indicate that the interactions are not based on which agents are within a certain range (as assumed by our algorithms), but instead on a fixed number of neighboring agents that can be processed by the Starling's nervous system. For this situation, most of the results obtained here are therefore not applicable. However, Ref. [30] corresponds to experiments in a very specific system and there is no reason to expect the many other examples of swarming biological agents to behave equivalently. In summary, as more experiments become available, the analysis tools that we introduce will have to be applied in a case by case basis. In some situations, the quantities that we study will only be accessible through indirect measurements or require additional adjustable parameters. In other situations, the experiments may unveil dynamic rules that are very different from the ones assumed in our simulations, rendering our tools ineffective or meaningless. Given the variety of swarms found in Nature and our limited understanding of their mechanisms, however, we expect that our work can provide a first approach for the analysis of many of them. In some cases it will have to be adapted to the possibilities and limitations of each particular experiment, but this link to specific systems is beyond the scope of the current paper. The paper is organized as follows. In Section 2 we describe the models and order parameters considered. Here we attempt to clarify a long-standing confusion by distinguishing between two different ways of implementing the model by Vicsek et al. Section 3 presents our main results, comparing various systems by using the standard and new order parameters. Section 4 uses the cluster size distribution to further investigate the global states of the system. In Section 5, we apply our analysis techniques to shed some light on a current controversy regarding the order of the phase transition in the Vicsek model. Finally, Section 6 is our conclusion.

# 2. Definitions

# 2.1. Models

We will consider in this paper three similar models that were designed to be *minimal*, i.e., to capture the essence of a swarming behavior while including only its basic components. They are all comprised of N agents defined as point particles that (for simplicity) are forced to move at the same constant speed s in a two-dimensional periodic square box of side L. At every time-step, each agent interacts with all agents within a neighborhood of radius R, computes its new direction of motion, adds some noise, and then advances synchronously with all other agents to a new position.

It is only in the implementation details that the models differ. We will provide below a thorough description of each algorithm to avoid previous ambiguities that have been detected in the literature.

Perhaps the first minimal swarming model that displayed a phase transition was introduced by Vicsek et al. in Ref. [4]. In this model, the next state of the system is computed at every time-step *t*, given the positions  $\{\vec{x}_i\}_{i=1,...,N}$  and the directions of motion specified by the angles  $\{\theta_i\}_{i=1,...,N}$ , by using the relations [17,18]:

$$\theta_i(t + \Delta t) = \operatorname{Ang}\left\{\sum_{Z_i(t)} \vec{\operatorname{Vel}}\left[\theta_j(t)\right]\right\} + \eta \,\xi_i(t)$$

$$\vec{x}_i(t + \Delta t) = \vec{x}_i(t) + \vec{\operatorname{Vel}}\left[\theta_i(t)\right] \,\Delta t.$$
(1)

Here, the functions Ang and Vel simply map the two-dimensional particle velocity vector to its corresponding heading angle and vice versa. The set  $Z_i(t)$  contains all *j*-indexes of agents within an interaction range *R* of agent *i* at time *t*. A delta-correlated white noise is introduced through the random variable  $\xi_i(t)$ , which is homogeneously distributed in the  $[-\pi, \pi]$  interval. This is exactly the algorithm originally introduced in Ref. [4] and we will thus refer to it as the *Original Vicsek Algorithm* (OVA).

Unfortunately, under certain computational implementations, the OVA may appear as somewhat unnatural, since it can also be cast as:

$$\vec{x}_i(t+\Delta t) = \vec{x}_i(t) + \vec{\text{Vel}} \left[ \text{Ang} \left\{ \sum_{Z_i(t-\Delta t)} \vec{\text{Vel}} \left[ \theta_j(t-\Delta t) \right] \right\} + \eta \, \xi_i(t-\Delta t) \right] \Delta t,$$
(2)

which implies that two previous time-steps are required to compute the state of the system at  $t + \Delta t$ . Given this, several authors [14,19–21] have implemented a slightly different version of the OVA, which is equivalent to changing  $(t - \Delta t) \rightarrow t$  in Eq. (2). The algorithm thus becomes:

$$\theta_i(t + \Delta t) = \operatorname{Ang}\left\{\sum_{Z_i(t)} \vec{\operatorname{Vel}}\left[\theta_j(t)\right]\right\} + \eta \,\xi_i(t)$$

$$\vec{x}_i(t + \Delta t) = \vec{x}_i(t) + \vec{\operatorname{Vel}}\left[\theta_i(t + \Delta t)\right] \,\Delta t.$$
(3)

When implemented in a computer program, this corresponds to first finding the new directions of motion for all particles, and then using these to update their positions. In what follows, we will refer to this algorithm as the *Standard Vicsek Algorithm* (SVA). Note that in the context of Hamiltonian systems the OVA would correspond to the Euler method for solving the equations of motion numerically, whereas the SVA would correspond to the leapfrog method. It is known that these two methods can lead to very different numerical solutions for the same system of differential equations. In particular, for conservative systems it can be shown that the Euler method does not conserve the volume in phase space (i.e. it is a non-symplectic integrator), whereas the leapfrog method does conserve the phase space volume (it is a symplectic integrator) [22–25]. Because of this, the Euler method is considered to be a bad algorithm for Hamiltonian systems and is never used to integrate the equations of motion numerically. Swarming models, however, are neither Hamiltonian nor conservative and it is therefore not clear which updating algorithm, the OVA or the SVA, should be used. While the difference between these two algorithms may appear to be minor, we will show in the following sections that their results for self-propelled systems can differ strongly, as it occurs for symplectic and non-symplectic integrators in Hamiltonian systems. Note that even in the limit of  $\Delta t \rightarrow 0$  the results given by these two algorithms may be different, as it happens in the Hamiltonian case.

The third algorithm that we consider is the one introduced by Grégoire and Chaté in Ref. [14], where the authors analyze results obtained with both this new algorithm and the SVA described above. Here, the noise is included in a slightly different way, by adding a small vector in a random direction to the sum in Eq. (3). The model becomes:

$$\theta_i(t + \Delta t) = \operatorname{Ang}\left\{\sum_{Z_i(t)} \vec{\operatorname{Vel}}\left[\theta_j(t)\right] + \eta \, n_i(t) \vec{\operatorname{Vel}}\left[\xi_i(t)\right]\right\}$$

$$\vec{x}_i(t + \Delta t) = \vec{x}_i(t) + \vec{\operatorname{Vel}}\left[\theta_i(t + \Delta t)\right] \Delta t,$$
(4)

where  $n_i(t)$  is the number of neighbors of particle *i* at time *t*, which is equal to the number of elements in the  $Z_i(t)$  set. We will refer to this last algorithm as the *Grégoire-Chaté Algorithm* (GCA). It is important to note that this GCA algorithm has been implemented in the literature in two different ways. In one version the agents always interact if they are within a certain distance of each other, but in another version they must *also* be in neighboring Voronoi cells. In the current work, we will concentrate in this first version of the GCA algorithm and not use any Voronoi tessellation.

## 2.2. Order parameters

In the study of swarming models, the most commonly used order parameter is defined to provide a measure of the degree of alignment of the agents' velocities. It is equivalent to the normalized magnitude of the total system momentum (or also, in the context of ferromagnetism, to its magnetization) and is defined by:

$$\psi(t) = \frac{1}{Ns} \left| \sum_{i=1}^{N} \vec{\text{Vel}}[\theta_i(t)] \right|.$$
(5)

An ordered system where all agents are approximately heading in the same direction therefore has  $\langle \psi \rangle \approx 1$  and a disordered one, where they move in random directions, has  $\langle \psi \rangle \approx 0$ . Using this definition, it was first shown in Ref. [4] that the OVA system undergoes a phase transition from an ordered to a disordered state as the amount of noise  $\eta$  is increased in the system. The transition was originally thought to be of second order, but this result has been recently disputed by Grégoire and Chaté in Ref. [14], where they claim that it will always be of first order for large enough systems. Unfortunately, the difference between a continuous and a discontinuous phase transition can only be observed here for very large systems, where it is hard to carry out extensive numerical computations. While we will not attempt to settle this dispute here, we believe that the additional global quantities that we introduce can help better understand the system's dynamics close to the phase transition, which may be quite different from the one in standard equilibrium phase transitions. We also confirm in Section 5 that it appears numerically as indeed a first-order transition for at least one of the parameter sets presented in Ref. [14] by observing a discontinuous jump in various global quantities characterizing the system.

It is well known that the particle motion plays an essential role in achieving an ordered state in the presence of noise [4,19,26]. This motion produces a distribution of particles in space that is non-trivial and strongly inhomogeneous, with zones of high density where particles aggregate into clusters and large empty regions. Given that only agents within an interaction range R of each other can communicate, it is natural to define a density-related global parameter to describe the system as

$$\tilde{\rho}(t) = \langle C_n[Z_i(t)] \rangle_i \,. \tag{6}$$

Here  $C_n[\cdot]$  provides the number of elements in the set (cardinality) and  $\langle \cdot \rangle_i$  averages over all *N* particles. With this definition,  $\tilde{\rho}$  is equal to the average number of neighbors with which each agent interacts at a given moment in time (including itself) and also corresponds to the local density of interacting particles. It can be easily computed not only in numerical simulations, but also in actual biological systems provided that we can define sensing criteria that determine when agents interact. Note that in more realistic situation these may be not only distance related, but could depend on non-isotropic sensing capabilities, line-of-sight interactions, screening, etc. Unfortunately, this analysis requires measuring the position of every agent within a swarm, and there are currently very few experiments that do this. However, a growing interest in the dynamics of swarms within communities in biology, physics, applied math, and engineering; together with the development of new tracking and detection techniques for biological setups allows us to expect new experiments to be carried out in the near future. An additional challenge is posed here by the need to use an interaction rule for computing the local density of interacting agents, since for biological systems the interactions are mostly ill understood. Here again additional work that will depend on each experiment is needed.

Another order parameter that we will utilize in our analysis is the mean distance to the nearest neighbor

$$\tilde{\delta} = \left\langle \min\left( \left| \vec{x}_j - \vec{x}_i \right|_{j \in Z_i} \right) \right\rangle_i.$$
<sup>(7)</sup>

This quantity has been previously used in the swarming literature [27] but, to the best of our knowledge, not for characterizing the minimal models considered here. Like  $\tilde{\rho}$ , it is also related to the local density of agents.



Fig. 1. Snapshots of the numerical simulations of the three different swarming algorithms considered in this paper. The two top- and bottomleft panels correspond to moderately ordered states of the OVA, SVA, and GCA models. The figure shows that the OVA produces slightly more condensed clusters than the SVA for the same amount of noise. The bottom-right panel displays a high-noise SVA simulation for comparison. All snapshots were obtained using N = 2048 agents in a box of side L = 128 and a particle speed s = 0.1. Each vector shows the position and heading of an agent. Its displayed length is equal to three times the interaction range R.

The final global analysis tool that we will consider in this paper is the *distribution of cluster sizes*. A cluster is defined here as a set of agents that are connected (i.e. within the interaction range of each other) either directly or through other agents at a given moment in time. Its size is the number of agents it contains. We will further discuss the properties of this distribution when we use it for our analysis in Section 4.

## 3. Order parameter analysis

We begin by computing the standard order parameter  $\psi$  defined in Eq. (5) as a function of the amount of noise for the OVA, the SVA, and the GCA. The length- and time-scales of the computation are fixed by defining the interaction range as R = 1 and using  $\Delta t = 1$ . All our data was obtained by integrating the dynamics for over 10<sup>7</sup> time-steps and averaging every 10<sup>4</sup> steps after an initial relaxation time.

Fig. 1 shows three typical snapshots of moderately ordered states computed using the OVA, SVA, and GCA models. The bottom-right snapshot shows an SVA simulation in its disordered state for comparison. The top panels on Fig. 2 present the results of numerical computations performed with N = 2048 and N = 8192 particles in a periodic square



Fig. 2. Average degree of agent alignment  $\psi$  (top), defined in Eq. (5), and average number of interacting neighbors per agent  $\tilde{\rho}$  (bottom), defined in Eq. (6), as a function of the amount of noise  $\eta$  for three different minimal swarming models and two system sizes. All simulations were set up with the same mean density  $\rho = 1/8$ , either N = 2048 or N = 8192 agents, and the same agent speed s = 0.1. Left panels: OVA and SVA detailed in Eq. (1) and (3). Right panels: GCA detailed in Eq. (4). On the top  $\psi$  vs.  $\eta$  graphs all models appear to behave similarly, with an ordered phase at low  $\eta$  and a disordered one at high  $\eta$ . On the bottom graphs, however, the  $\tilde{\rho}$  vs.  $\eta$  curves clearly display different properties for each model.

box of side L = 128 and L = 256, respectively. Both cases correspond to the same mean density  $\rho = 1/8$  and use a particle speed equal to s = 0.1. These parameters, as well as those to be considered later in the paper, correspond to regimes previously studied in the literature [14,28]. The panels display the usual phase transition for all six cases, which appears as being of second order for this system size. The main observation here is that the curves only display small differences: The SVA would seem to have a smaller finite-size effect than the SVA [29] and the scale of  $\eta$  is not the same for the GCA, which is simply a consequence of the fact that the noise is introduced differently in Eq. (4). No other qualitative differences are apparent when we only study this quantity.

When we consider  $\tilde{\rho}$ , however, the three systems become quite distinct. The bottom panels on Fig. 2 show this order parameter as a function of  $\eta$  for the same runs displayed on the top panels. In spite of the subtle differences between the OVA and the SVA, the former produces much higher local densities. In fact, for the relevant  $\eta$  values close to the phase transition, the OVA generates a dynamics where each agent interacts with over thirty neighbors in average, which is well above what is expected for realistic systems [30]. Note that all tree algorithms produce a too high number of interacting agents  $\tilde{\rho}$  as  $\eta$  is decreased to zero. This is due to the fact that the models contain no repulsive term and it is only a sort of "thermal pressure" (here, a combination of the agents' noise and speed) that keeps particles apart. This is a known shortcoming of these minimal models but it can be disregarded if in the parameter region of interest the local density does not become exceedingly unrealistic. In the case of the OVA regimes considered here, however, our computation of  $\tilde{\rho}$  shows that unrealistically high densities occur precisely in the parameter regions of interest; there where the transition is observed. The higher  $\tilde{\rho}$  values observed in the OVA can be explained by analyzing the details of each algorithm. In the OVA, particles are moved in a direction that is computed based on the neighborhood state at  $t - \Delta t$ , while the SVA uses the neighborhood state at t. The alignment of particles belonging to a locally ordered group at a given moment in time will therefore persists for two time-steps in the OVA, making it more difficult for agents to escape the group and thus producing higher local densities.



Fig. 3. TOP: Average degree of agent alignment  $\psi$  (top), defined in Eq. (5), and average number of interacting neighbors per agent  $\tilde{\rho}$ , defined in Eq. (6), as a function of noise  $\eta$  for three different minimal swarming models and two system sizes. As in Fig. 2, all simulations were set up with the same mean density  $\rho = 1/8$  and either N = 2048 or N = 8192 agents, but here the agent speed was set to s = 0.5. Left panels: OVA and SVA detailed in Eqs. (1) and (3). Right panels: GCA detailed in Eq. (4). On the top  $\psi$  vs.  $\eta$  graphs all models appear to behave similarly, with an ordered phase at low  $\eta$  and a disordered one at high  $\eta$ . As in Fig. 2, however, the  $\tilde{\rho}$  vs.  $\eta$  curves on the bottom graphs show different properties for each model.

Fig. 3 presents  $\psi$  vs.  $\eta$  and  $\tilde{\rho}$  vs.  $\eta$  curves analogous to the ones in Fig. 2, but for a higher particle speed s = 0.5. The same qualitative features are found but the  $\tilde{\rho}$  values are significantly lower in these cases. This is due to a higher particle mobility that tends to separate agents more for a given noise level. Note that the differences that we observe on Figs. 2 and 3 between the simulations produced by each algorithm can only be seen on the  $\tilde{\rho}$  vs.  $\eta$  curves and not on the  $\psi$  vs.  $\eta$  curves, which are very similar for all cases. We found an interesting additional feature on the  $\tilde{\rho}$  vs.  $\eta$  curves: they present an interval where the second derivative changes sign, becoming negative at  $\eta$  values right under the phase transition. This "bump" is present in all three algorithms, but is more prominent for larger N and s values. It indicates a non-trivial behavior which could be measured in real biological swarms. We will further discuss this effect in the following sections.

The third-order parameter that we will analyze is the mean distance to the nearest neighbor  $\delta$ . This quantity has the advantage that it can be easily and unambiguously computed given only the agent positions. In contrast to  $\tilde{\rho}$ , it does not depend on R, and can therefore be extracted from experimental data without any further assumption regarding the biological interactions among individuals, which are often poorly understood. Fig. 4 displays  $\delta$  vs.  $\eta$  curves for the same s = 0.5 runs presented on Fig. 3. The qualitative behavior of  $\delta$  is opposite to that of  $\tilde{\rho}$ : As the local density rises, the distance to the nearest neighbor decreases and vice versa. More precisely, in all numerical simulations tested, the approximate relation  $\tilde{\rho} \propto \delta^{-2}$  was verified (data not shown). This is what would result from randomly positioning  $\tilde{\rho}$  neighbors within a distance R of each agent. Our analysis thus suggests that  $\delta$  may provide an effective way of gauging the density of interactions within a swarm, since it can be readily computed using experimental data such as the one collected in Refs. [30,34].



Fig. 4. Average distance between nearest interacting neighbors  $\tilde{\delta}$ , defined in Eq. (7), as a function of the amount of noise  $\eta$  for three different minimal swarming models and two system sizes. The simulations are the same that were used on Fig. 3. All  $\tilde{\delta}$  curves follow a behavior that is opposite to the one of their corresponding  $\tilde{\rho}$  curves on Fig. 3 (bottom panels).



Fig. 5. Average degree of agent alignment  $\psi$  (left panel) and number of interacting neighbors per agent  $\tilde{\rho}$  (right panel) as a function of the amount of noise  $\eta$  for a GCA model with mean density  $\rho = 2$ . Each simulation contains N = 2048 agents moving at a speed s = 0.5. In this regime, the order-to-disorder phase transition appears as clearly discontinuous and  $\tilde{\rho}$  develops a prominent peak just below the critical noise.

One of the interesting properties of the GCA is that it can display a sharp discontinuous phase transition in certain regimes, even for relatively small system sizes [14,21]. In order to study these cases, we computed  $\psi$ ,  $\tilde{\rho}$ , and  $\tilde{\delta}$  as a function of  $\eta$  for the same parameters used in Ref. [14]: N = 2048, s = 0.5, and a mean density of  $\rho = 2$ , significantly higher than the one that we have considered up to now. The left-side panel of Fig. 5, which is equivalent to the one published in Ref. [14], shows that  $\psi(\eta)$  evidences a clearly discontinuous phase transition in this regime. The right-side panel displays a novel feature:  $\tilde{\rho}$  increases sharply just under the critical noise value. It presents the same "bump" observed on the bottom panels of Figs. 2 and 3, but here more pronounced and delimited at the high  $\eta$  end by a discontinuous jump in  $\tilde{\rho}$ . An analogous but opposite behavior, as in the previous case on Fig. 4, is followed by the  $\delta$  vs.  $\eta$  curve (data not shown). This is further evidence of the existence of an underlying generic mechanism, which remains to be understood, that increases the local agent density close to the phase transition in these minimal swarming models.

Finally, it is important to point out that in all the cases that we studied and even at very high-noise levels, the average local density  $\tilde{\rho}$  generated by the dynamics was much higher than the mean density  $\rho$  that would result from an homogeneous agent distribution. This confirms that clustering is a prominent feature in these swarming models. We should also point out that all the plots in this section were presented as a function of the amount of noise  $\eta$ , which may not be an easy parameter to control in biological experiments. However, given that swarming data is scarce and that we do not know a priori what future experiments may be able to measure and control directly or indirectly, we have presented our results in this way, which is standard in the context of numerical swarm analyses and traditional

phase transitions in physics. Additional work will be needed to find out how the features described here in terms of  $\eta$  will translate to specific signatures in a given future experiment.

## 4. Cluster size distribution analysis

In this section, we will further analyze the OVA, SVA, and GCA swarming simulations by computing their distribution of cluster sizes. In previous work, we had shown that the SVA typically produces clusters of all sizes that are continuously forming and breaking apart, and that the distribution of cluster sizes resembles a power-law in its ordered phase [20]. We found no characteristic cluster size but that instead the probability of having larger clusters decays smoothly with size. We will extend here this analysis to consider the cluster distribution of all three models in a larger range of parameters.

We define a cluster as a group of agents that are interacting either directly or *via* other agents at a given moment in time. Two agents are therefore in the same cluster if they are within the interaction range R of each other or if they are connected by a path of interacting agents. The *size* of a cluster is then given by the number n of agents it contains. Note that the shape of a cluster is not constrained here by any condition on its contour, since our swarming dynamics contain no effect resembling surface tension.

Figs. 6 and 7 display the cumulative distributions of cluster sizes for the same runs presented in Section 3. Like before, the curves were computed by averaging data from  $10^3$  snapshots or more, taken every  $10^4$  time-steps and after reaching a statistical steady state [31]. The cumulative nature of the distributions means that each point in a C(n) curve gives the probability of having a cluster containing *n* agents or more. We present two panels per row, one for s = 0.1 and one for s = 0.5, each one displaying various curves for different values of  $\eta$ .

Our first observation is that all three systems follow a similar qualitative behavior that can be summarized as follows. At high-noise values, the system typically has an exponentially decaying cluster-size distribution. As  $\eta$  is decreased, the fraction of large clusters increases and C(n) approaches a power-law distribution with a cutoff at high *n* values. It is unclear from our data if this power-law approaches any well-defined asymptotic exponent as  $\eta$  goes to zero. In fact, when  $\eta$  is identical to zero all agents must eventually coalesce into one giant cluster in any finite system, so this exponent must vanish. This makes the analysis of the  $\eta \to 0$  limit much harder, since finite-size effects become increasingly important as the probability of having all N agents in the same cluster grows. Larger simulations would therefore be required to explore this limit, which is computationally very demanding since the relaxation dynamics will slow down as the noise is decreased. With the current data we can only state that for  $\rho = 1/8$  we measure exponents at low-noise values that are typically between  $n^{-1/2}$  and  $n^{-3/4}$ , which we include as reference lines on the figures. This is consistent with our previous results in Ref. [20] for a case with  $\rho = 0.4$ , but corresponds to shallower asymptotic slopes than the ones obtained for the high density case ( $\rho = 2$ ) presented on Fig. 7. Note that we are describing here the cluster size distribution curves loosely as being close to an exponential or to a power-law. We do not intend to imply that the curves are ideally fitted by these distributions. Instead, we are simply referring to the large *n* behavior of the distribution curves as presenting a sharp cutoff (similar to the one in an exponential distribution) or a long tail (like the one typically found in a corresponding power-law distribution), and focusing on what appears to be the asymptotic behavior of these curves. An interesting anomaly occurs in the cluster size distributions displayed on the center and bottom panels of Fig. 6 and on Fig. 7. As the noise level is lowered, C(n) first approaches a power-law as expected, but then curves back to an exponential-like distribution at intermediate noise levels before converging again towards a power-law as  $\eta$  approaches zero. This behavior seems to be related to the  $\eta$  intervals where the second derivatives of the  $\tilde{\rho}(\eta)$  and  $\tilde{\delta}(\eta)$  curves change sign, as pointed out in Section 3 (see the bottom panels of Figs. 2–5). The curving towards an exponential distribution at intermediate noise values appears to occur only at the lower  $\eta$ end of this "bump" in the  $\tilde{\rho}(\eta)$  curves, where their slope is either positive (Fig. 3 – bottom – and Fig. 5) or the least negative (Fig. 2 – bottom). The relation between  $\tilde{\rho}$  and C(n) is non-trivial, however, as it changes in each case. Indeed, for the SVA and the GCA simulations with  $\rho = 1/8$ , the C(n) curves that are closest to a power-law at intermediate noise levels occur near the local maximum of  $\tilde{\rho}$ . This can be seen by looking at the bottom panels of Fig. 3 and the right-side panels of Fig. 6 at  $\eta \approx 0.13$  and  $\eta \approx 0.3$ , respectively. In contrast, the most power-law-like C(n) curve reached by the GCA computations with  $\rho = 2$  at intermediate noise levels is obtained at  $\eta = 0.62$ , above the value at which  $\tilde{\rho}$  jumps to its local maximum as the noise is decreased to  $\eta = 0.615$  (see Figs. 5 and 7). Note that the former case appears to be more natural since  $\tilde{\rho}$  grows together with the fraction of large clusters. The latter case is



Fig. 6. Cumulative distribution of cluster sizes (in number of agents *n*) of OVA (top), SVA (center), and GCA (bottom) simulations at various noise levels. Each curve corresponds to a run that was used to compute one point on Fig. 2 or Fig. 3. All simulations were carried out with N = 8192 agents moving at speeds s = 0.1 (left panels) or s = 0.5 (right panels) and mean density  $\rho = 1/8$ . In the OVA case the fraction of large clusters increases as the noise level is decreased, with C(n) approaching a power-law distribution. In the SVA and GCA cases most curves also approach a power-law as the noise level is decreased, but this trend is reversed in the  $\eta = 0.13$  to  $\eta = 0.07$  and  $\eta = 0.2$  to  $\eta = 0.4$  intervals, respectively.

somewhat anomalous since the fraction of large clusters drop as  $\tilde{\rho}$  jumps to its peak when the noise level is reduced from  $\eta = 0.62$  to  $\eta = 0.615$ . The system here becomes a more tightly connected set of smaller clusters.

As a final note, let us point out that in most cases the transition towards a power-law distribution occurs very gradually in the C(n) curves as the noise level is decreased, and that it is only apparent when we focus on relatively large *n* values. This implies that the cluster-size distribution analysis is, unfortunately, not a good tool for detecting the phase transition or identifying its order. It does, however, produce additional information on the agents' spatial distribution and can be used to better understand the critical behavior and the main characteristics of each phase. In



Fig. 7. Cumulative distribution of cluster sizes (in number of agents *n*) of GCA simulations at various noise levels. Here the density is  $\rho = 2$ , higher than in Fig. 6. Each curve corresponds to a run used to compute one point on Fig. 5, with N = 2048 agents moving at a speed s = 0.5. The left panel displays noise values from  $\eta = 0.45$  to  $\eta = 0.615$  and the right one from  $\eta = 0.615$  to  $\eta = 0.66$ . As in the center and bottom rows of Fig. 6, while most curves slowly approach a power-law for decreasing noise levels, the reverse trend is also observed here in the interval from  $\eta = 0.615$  to  $\eta = 0.62$ .

the next section we will use this and other tools to clearly distinguish the two coexisting phases that are present on the bi-stable regime first described in Ref. [14].

## 5. Bi-stability analysis of a large system

In Refs. [14,29] Grégoire and Chaté claim that all three minimal models (the OVA, SVA, and GCA) present a firstorder phase transition in the thermodynamic limit. They argue that this transition may only *appear* to be of second order in some finite-size systems (such as in the top panels of Figs. 2 and 3) when the size L of the simulation box is smaller than a certain crossover length-scale  $L^*(\rho, s)$ . The  $L > L^*$  condition would be required to develop density bands that become metastable to a long-wavelength instability below the critical noise value, giving rise to a firstorder transition with a hysteresis region where the ordered and disordered phases coexist. As part of their evidence, a bimodal distribution is shown in Ref. [14] for the order parameter  $\psi$  in a large (N = 32768) SVA simulation close to the phase transition. However, Nagy, Daruka, and Vicsek claim in Refs. [4,28] that the OVA and SVA transitions are always of second order for small enough particle velocities. They argue that for larger s values one can indeed observe a discontinuous transition, but that this is only due to the finite-size effects of the periodic boundary conditions, which help generate the density bands and give rise to an artificial symmetry breaking that produces the ordered phase.

In this section, we will apply our analysis tools to explore more in detail the collective behavior at the phase transition in the same regime studied in Refs. [14,28]. While the order of the phase transition in the thermodynamic limit is still a matter of controversy that we will not address here, we hope that our approach will help elucidate the mechanisms behind the transition and ultimately contribute to clarify its nature.

Figs. 8 and 9 were computed for an SVA with  $\rho = 1/8$ , N = 32768, and s = 0.5, where Grégoire and Chaté had reported a bimodal  $P(\psi)$  distribution in Ref. [14]. We used a noise value of  $\eta = 0.1915$ , placing the system very close to the phase transition. The statistics was accumulated by averaging every 10<sup>4</sup> time-steps after an initial relaxation time and integrating the dynamics for 10<sup>7</sup> steps. On Fig. 8 we confirm that  $\psi$  presents indeed a bimodal distribution, which is apparent on the top right panel. In addition, the figure displays the PDF of  $\tilde{\rho}$  (bottom-right panel) and the combined two-dimensional PDF of  $\psi$  and  $\tilde{\rho}$  (left-side panel) [31]. It distinctly shows that in this regime there are two well-defined coexisting solutions: one with a higher degree of alignment and higher local density, and one with less alignment order and lower local density. As the noise level is increased, the local maximum of the disordered state becomes more prominent while the other one shrinks, vanishing at the critical noise level. This confirms that the transition is discontinuous, at least in this numerical example. Additionally, the typical correlation function between  $\psi$  and  $\tilde{\rho}$  (which corresponds here to the steepest descent path between the local maximu in the two-dimensional PDF) follows an approximate  $\psi \propto \sqrt{\tilde{\rho} - \tilde{\rho_0}}$  relation indicated by a solid curve on the left panel. Such relation seemed to remain the same for various noise values close to the phase transition (data not shown), with the distribution exploring



Fig. 8. Probability distribution of the degree of agent alignment  $\psi$  and number of interacting neighbors per agent  $\tilde{\rho}$  for an SVA simulation with  $\rho = 1/8$ , N = 32768, s = 0.5, and  $\eta = 0.1915$  (close to the critical noise value). The left panel presents a density plot displaying a bimodal distribution on the two-dimensional ( $\tilde{\rho}, \psi$ ) space. The overlaid solid curve corresponds to a  $\psi \propto \sqrt{\tilde{\rho} - \tilde{\rho}_0}$  relation. The right panels present the same data separately as  $P(\psi)$  and  $P(\tilde{\rho})$  distributions.



Fig. 9. Cumulative distribution of cluster sizes (in number of agents *n*) for the same simulation used in Fig. 8, with  $\rho = 1/8$ , N = 32768, s = 0.5, and  $\eta = 0.1915$ . The solid curve corresponds to the distribution measured when the system is in the ordered regime (when  $\psi > 0.15$ ) and the dashed one, when it is in the disordered regime (when  $\psi \le 0.15$ ). It is apparent that each regime of the bimodal distribution on Fig. 8 has a different distribution of cluster sizes.

different regions of the correlation as we varied  $\eta$ . This suggests that the typical relation between  $\psi$  and  $\tilde{\rho}$  may be a robust feature that could be probed experimentally. Note that such analysis only requires measuring the position and orientation of all agents and does not depend on any parameters, adjustable quantities, or additional biological details.

In order to further characterize the two metastable solutions in the bimodal distribution, we computed their characteristic cluster sizes separately using the following procedure. For every snapshot considered in the statistics, we determined if the system was in the ordered or the disordered phase based on the value of  $\psi$ . If  $\psi \le 0.15$  (which corresponds to the local minimum of  $P(\psi)$  that separates both regimes) the snapshot's cluster size distribution was added to the disordered regime statistics and if  $\psi > 0.15$ , it was added to the ordered regime statistics. Fig. 9 displays the resulting two curves on the same graph. It is apparent that the two coexisting regimes are also distinguished by their distribution of cluster sizes. In the ordered regime the distribution is closer to a power-law while in the disordered one it is more exponential-like, decaying faster as *n* grows. The dynamics in this case therefore explores intermittently two distinct metastable regimes: an ordered one with a higher degree of alignment, higher mean local density, and higher fraction of large clusters; and a disordered one with a lower degree of alignment, lower mean local density, and lower fraction of large clusters.

## 6. Conclusion

We have examined various global statistical quantities that characterize in new ways the collective behavior of swarming systems. In contrast to previous analyses, we concentrate not only on the agents' orientation but also on their spatial distribution. We focused in this paper on three well-known minimal models that display similar features when analyzed with standard tools. By computing the new quantities, we were able to better characterize and clearly distinguish the dynamics of the three cases. Our study, however, does not need to be restricted to minimal numerical models, but is susceptible to be applied to more realistic algorithms or to experimental systems.

Our investigations led us to unveil previously unnoticed properties of the minimal models considered. We found that the lack of a repulsive effect produces regimes where each agent interacts with an unreasonably high number of neighbors at every moment in time. This can strongly hinder the applicability of these models if it occurs within the region of interest. The original model studied by Vicsek et al. Ref. [4] (which we referred to as OVA), for example, presents certain (otherwise realistic) parameter regions where the system can only organize to start swarming when each agent is interacting on average with close to 50 neighbors. This is an order of magnitude above what is expected for most biological swarms. We conclude that the OVA is not a practical option for simulating swarms within this regime since the high number of neighbors is not only unrealistic, but also slows down substantially the numerical computations. This suggests that the quantities considered here should be monitored within any simulation to verify that they behave properly.

Another property revealed by our analysis concerns the curves that trace the average number of interacting neighbors per agent (or local density of interacting agents) and the mean distance to the nearest neighbor as a function of the amount of noise. We discovered that, in all three minimal models, there is an interval of noise values just below the phase transition where the sign of their curvatures change, producing a higher number of neighbors and smaller distances to the nearest neighbor than otherwise expected. This feature appears to be connected to a change in the relation between the distribution of cluster sizes and the amount of noise. In the usual situation, we found that this distribution is closer to a power-law for low-noise levels and that it approaches an exponential as the noise is increased. In the interval where the curvature changes, however, this trend is reversed. Our analysis suggests that there may be a universal behavior that characterizes this family of non-equilibrium phase transitions close to the critical noise.

Perhaps more important than the specific conclusions reached for the models studied in this paper is the fact that the quantities we introduced can be computed in other more realistic systems. Indeed, various experiments have been recently set up to measure the position and motion of each organism in biological swarms [30,32–34]. While these experiments differ strongly (some deal with large swarms in three dimensions, such as Ref. [30], while others only trace the trajectories of a few individuals in two dimensions), they all coincide in the facts that they explore individual and collective interactions by studying the details of each agent's motion and in that they were published within the last four years. This recent focus on the detailed behavior of biological swarms prompts us to expect a greater number of swarming experiments to be carried out in the near future in a larger variety of biological systems. The question then will become how to summarize the accumulated data into meaningful statistical quantities that provide a good measure of the global state of the system. In this context, we have shown that our analysis tools can be used successfully to characterize such systems. In particular, all the quantities that we introduced can be computed in experiments, only requiring the agent positions in time and, in some cases, an estimation of the interaction range. In situations where the control parameters cannot be easily specified, we have also shown that the correlation between the measured quantities that we have explored numerically in this paper.

#### Acknowledgements

The work of C. Huepe was supported by the National Science Foundation under Grant No. DMS-0507745. M. Aldana acknowledges CONACyT and PAPIIT-UNAM for partial support under grants P47836-F and IN112407-3, respectively.

#### References

[1] C.W. Reynolds, Flocks, herds, and schools: A distributed behavioral model, Comput. Graph. 21 (1987) 25.

- [2] K. Warburton, J. Lazarus, Tendency-distance models of social cohesion in animal groups, J. Theoret. Biol. 150 (1991) 473.
- [3] D. Grunbaum, A. Okubo, Modeling social animal aggregations, in: Frontiers in Theoretical Biology, in: Lecture Notes in Biomathematics, vol. 100, Springer-Verlag, New York, 1994, pp. 296–325.
- [4] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, O. Shochet, Novel type of phase transition in a system of self-driven particles, Phys. Rev. Lett. 75 (1995) 1226.
- [5] N. Shimoyama, K. Sugawara, T. Mizuguchi, Y. Hayakawa, M. Sano, Collective motion in a system of motile elements, Phys. Rev. Lett. 76 (1996) 3870–3873.
- [6] E.V. Albano, Self-organized collective displacements of self-driven individuals, Phys. Rev. Lett. 77 (1996) 2129–2132.
- [7] H. Levine, W.-J. Rappel, I. Cohen, Self-organization in systems of self-propelled particles, Phys. Rev. E 63 (2000) 017101.
- [8] I.D. Couzin, J. Krause, R. James, G.D. Ruxton, N.R. Franks, Collective memory and spatial sorting in animal groups, J. Theoret. Biol. 218 (2002) 1.
- [9] S. Gueron, S.A. Levin, D.I. Rubenstein, The dynamics of herds: From individuals to aggregations, J. Theoret. Biol. 182 (1996) 85-98.
- [10] G. Flierl, D. Grunbaum, S. Levin, D. Olson, From individuals to aggregations: The interplay between behavior and physics, J. Theoret. Biol. 196 (1999) 397.
- [11] J.K. Parrish, L. Edelstein-Keshet, Complexity, pattern, and evolutionary trade-offs in animal aggregation, Science 284 (1999) 99.
- [12] J.K. Parrish, W.H. Hamner (Eds.), Animal Groups in Three Dimensions, Cambridge University Press, Cambridge, MA, 1997 (and references therein).
- [13] I.D. Couzin, J. Krause, N.R. Franks, S.A. Levin, Effective leadership and decision-making in animal groups on the move, Nature 433 (2005) 513.
- [14] G. Grégoire, H. Chaté, Onset of collective and cohesive motion, Phys. Rev. Lett. 92 (2004) 025702.
- [15] F. Cucker, S. Smale, Emergent behavior in flocks, IEEE Trans. Automat. Control 52 (2007) 852.
- [16] U. Erdmann, W. Ebeling, A.S. Mikhailov, Noise-induced transition from translational to rotational motion of swarms, Phys. Rev. E 71 (2005) 051904.
- [17] A. Czirók, H. Eugene Stanley, T. Vicsek, Spontaneous ordered motion of self-propelled particles, J. Phys. A 30 (1997) 1375.
- [18] A. Czirók, T. Vicsek, Collective behavior of interacting self-propelled particles, Physica A 281 (2000) 17.
- [19] M. Aldana, C. Huepe, Phase transitions in self-driven many-particle systems and related non-equilibrium models: A network approach, J. Stat. Phys. 112 (2003) 135–153.
- [20] C. Huepe, M. Aldana, Intermittency and clustering in a system of self-driven particles, Phys. Rev. Lett. 92 (2004) 168701.
- [21] M. Aldana, V. Dossetti, C. Huepe, V.M. Kenkre, H. Larralde, Phase transitions in systems of self-propelled agents and related network models, Phys. Rev. Lett. 98 (2007) 095702.
- [22] B. Leimkuhler, S. Reich, Simulating Hamiltonian dynamics, in: Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, 2004.
- [23] R.I. McLachlan, G. Reinout, W. Quispel, Geometric integrators, J. Phys. A 39 (2006) 5251–5285.
- [24] T.J. Bridges, S. Reich, Numerical methods for Hamiltonian PDEs, J. Phys. A 39 (2006) 5287–5320.
- [25] B.A. Shadwick, J.C. Bowman, P.J. Morrison, Exactly conservative integrators, SIAM J. Appl. Math. 59 (3) (1999) 1112–1133.
- [26] J. Toner, Y. Tu, Long-range order in a two-dimensional dynamical XY model: How birds fly together, Phys. Rev. Lett. 75 (1995) 4326–4329.
- [27] A. Huth, C. Wissel, The simulation of fish schools in comparison with experimental-data, Ecol. Model. 75 (1994) 135–146.
- [28] M. Nagy, I. Daruka, T. Vicsek, New aspects of the continuous phase transition in the scalar noise model (SNM) of collective motion, Physica A 373 (2006) 445.
- [29] H. Chaté, Private communication, 2007.
- [30] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, V. Zdravkovic, Interaction Ruling Animal Collective Behaviour Depends on Topological Rather than Metric Distance: Evidence from a Field Study, arXiv:0709.1916v1 [q-bio.PE] (2007).
- [31] Wolfram Research, Inc. Mathematica 6.0 Wolfram Research, Inc., Champaign, IL 2005.
- [32] S.V. Viscido, J.K. Parrish, D. Grunbaum, Individual behavior and emergent properties of fish schools: A comparison of observation and theory, Marine Ecology - Progress Series 273 (2004) 239–249.
- [33] J.H. Tien, S.A. Levin, D.I. Rubenstein, Dynamics of fish shoals: Identifying key decision rules, Evol. Ecol. Res. 6 (4) (2004) 555–565.
- [34] J. Buhl, D.J.T. Sumpter, I.D. Couzin, J.J. Hale, E. Despland, E.R. Miller, S.J. Simpson, From disorder to order in marching locusts, Science 312 (2006) 1402.