Universitat
de les Illes Balears

Active cluster crystals with Vicsek-like alignment interaction

Javier Aguilar-Sánchez

Master's Thesis

Master's degree in Physics of Complex Systems
at the
Universitat De Les Illes Balears

2017 - 2018

September 2018

Thesis Supervisor : Cristóbal López

ACKNOWLEDGMENTS

I would like to thank my supervisor, Cristóbal López, for his dedication to this project besides his professional and personal advice. I also acknowledge my relatives from Mallorca, especially to my auntie Silvia and her unconditional support through the entire year. I am grateful for the help received from the researchers of IFISC Emilio Hernández, for his corrections and discussion about active matter and Tomás Sintes and Pere Colet for their comments on the numerical methods. I have to show my recognition to Yaouen Fily[1] for helping me to fully understand his work [1]. The last, but not the least, I honor my dear colleagues from the master; for they reinforced my idea that cooperation is much more efficient than competition.

"Non-equilibrium statistical mechanics largely remains messy, elusive and non-rigurous ... and this is precisely why it is such an interesting field" [2]

[1] Wilkes Honors College, Florida Atlantic University.

CONTENTS

Acknowledgments 1

Abstract 3

I. Introduction 3

II. Background 4
Passive matter: Brownian Motion 4
Active matter 6
Vicsek Model 7
Active cluster crystals 9

III. The model 13
Low diffusion regime: effect on empty clusters 14
Phase diagrams 15
Large diffusion regime 18

IV. Conclusions 19

V. Appendix 21
A. Diffusion equation: dimension argument 21
B. Numerical methods for solving SDE 22
C. Equivalence between Active Ornstein-Uhlenbeck process and Active Brownian particle. 23
D. Directional Statistics 24
Angular averages 24
E. URL 25
F. Dictionary of abbreviations and summary of stochastic processes 26

References 27

ABSTRACT

Our aim is to mix together the Vicsek model, that settled the basis of the understanding of the movement of swarming animals, with a model that performs active crystal clusters. To include interactions that depend on the orientation is a simple way of generalizing the usual isotropic forces associated to the crystal clusters formation. Moreover, the two models lead to the non-trivial emergence of collective phenomena, besides stressing the relevance of the non-equilibrium nature intrinsically linked with the active movement of the particles. We study how the pattern formation affects the well-known results from the Vicsek model and under which conditions it is possible to induce a coherent motion on the crystal structure. Through the process of building the model, we enter in fundamental questions about active matter, in particular, we seek for a microscopic argument that can tell us why active matter is out of equilibrium. Our method is based on the numerical integration of the equations of motion, we will either study single realizations and develop a statistical study.

I. INTRODUCTION

The emergence of complexity at the macroscopic level from a microscopic description is one of the key topics of the physics of complex systems. Non-trivial pattern formation, synchronization and collective movement merge spontaneously from a reduced number of (usually short ranged) interactions at the individual level [3] [4] [5] [6]. This situation holds the understanding of nature introduced by Statistical Mechanics and is perfectly summarized on the complex physicists' mantra "More is different" [7]. Moreover, there is a really stimulating landscape of different strategies to deal with these problems that range from well-known equilibrium statistical mechanics machinery to agent-based models and field descriptions of the non-equilibrium framework passing through sophisticated simulations for solving either non-equilibrium and equilibrium questions.

To illustrate what we have said, we refer to [3], which acts as the seminal work for this entire thesis as well as a good example of collective phenomena. Lets think about a box filled with random walkers just equipped with purely repulsive long ranged interactions. Our intuition on the subject could tell us that diffusion will dominate the phenomenology and therefore any stationary state should be statistically homogeneous (as the usual picture of a gas filling a box). However, since the system is bounded and the interactions are long ranged, particles will aggregate into clusters which will conform an hexagonal regular pattern. Hence, particles will be embedded in a "fight" between a repulsive individual interaction versus an effective attractive collective force. In other words, purely repulsive interactions which could represent a competing strategy for the space give raise to a cooperative behavior at the macroscopic level.

The general picture in physics is that particles are driven by external influences, such as some potential or the stochastic collisions with the environment. Nevertheless, active particles which are able to extract energy from the environment in order to sustain persistent motion autonomously [8] [9] [10] [11] have obtained a lot of attention lately. Firstly, there is a huge number of physical and biological examples of this kind of motion, spanning an enormous range of length scales (Escherichia coli, swarming animals, self-propelled artificial particles, colloidal systems, ...) [1]. Secondly, it is intrinsically fascinating since the active matter can only be studied from an out of equilibrium perspective and is linked with interesting phenomena as pattern formation, order-disorder transitions, swarming-like movement and anomalous fluctuations [1] .

The starting point for this thesis was reproducing the work [4]. One more time, non-trivial order and cluster formation arise from stochastic particles with purely repulsive interactions; the key difference with [3] is that particles have also an active nature. The aim of this thesis is to equip a system that displays active clusters like in [4] within the most famous interaction between the directions of the self-propelled movement of the particles, that is, the Vicsek model [5]. Many interesting questions

can be addressed to this combination. The first one is quite obvious, How the repulsive interactions change the well-known phenomenology of the Vicsek model? But there are more: In which situations the alignment interaction breaks the periodic crystal structure? Is it possible to induce a coherent synchronized movement for the pattern? Are the empty clusters presents in [4] robust?

The body of the work is the direct numerical simulation of the system, that is, each particle will be described by an stochastic differential equation (SDE) and we will use some numerical methods to integrate them. We will basically gain insights about the general behavior of the system with the direct observation of trajectories, then we will identify relevant quantities that serve as order and control parameters to proceed with the exhaustive variation of the later.[2]

In section II, since all the terms in our equations of motion have been largely discussed separately, we see it convenient to give but a brief summary about the construction of the model. Through this section we will not only reproduce and summarize well-known results but we will enter also in a discussion about up to date problems related with the theoretical basis of active matter. In particular, we introduce several active matter models and study the conditions under which they are statistically equivalent; then, we develop a microscopic argument that make us understand the non-equilibrium nature of active matter. The rest of the thesis is organized as follows: In section III, the complete model is presented and the main results are given. General conclusions, comments and proposals for future lines of research on the topic will conclude the main part of the thesis (Section IV). Finally, some calculations and comments complete the text and makes it self-contained in the Appendix section V. Despite not being really frequent in this type of texts, we found it useful to include the full URL directions for all the movies attached in the main text, a dictionary of abbreviations and a quick summary of the stochastic processes defined through all the thesis also in the Appendix.

II. BACKGROUND

Passive matter: Brownian Motion

The Brownian motion is the erratic movement of mesoscopic particles suspended in a fluid (or gas) generated by a myriad of collisions with fast moving microscopic particles (molecules or atoms). The effect of the microscopic particles is modeled by an stochastic term in the equations of motion. Since the time scales in which we are interested are usually large in comparison with the characteristic time between collisions, it is suitable to use a white noise (Gaussian[3] fully-uncorrelated stochastic process) as the stochastic contribution [13]. Having into account friction, we are able to write down the full Newton's second-law equation for a Brownian particle.

$$\begin{gathered} m\ddot{\vec{r}}_i = -\gamma\dot{\vec{r}}_i + \sqrt{2D^*}\vec{\xi}_i(t)\,, \\ \langle \xi_i^\alpha(t)\rangle = 0, \\ \left\langle \xi_i^\alpha(t)\,\xi_j^\beta(t')\right\rangle = \delta_{ij}\delta_{\alpha\beta}\delta(t-t')\,. \end{gathered} \tag{1}$$

Where the Greek characters label the components of Euclidean d-dimensional vectors and the latin characters label the different particles.

Whereas the above Langevin equation (1) gives the complete dynamical description of the system, the *overdamped* limit is usually assumed [9]. That is, the friction is dominant so we can neglect

[2] Being rigorous, the control and order parameters are terms characteristics from the theory of phase transitions [12]. Despite avoiding the discussion of the existence of a true thermodynamic phase transition in this work, we talk about control parameters in the sense that varying them we provoke relevant behavioral changes in the system and order parameters in the sense that they inform in a compact way about the state of the system.

[3] It is a suitable situation for letting the central limit theorem to work since the Brownian motion is the result of a *great number* successive collisions.

the inertia term. This situation will obviously simplify our analysis since we will deal with a first order SDE. Moreover, it is a suitable approximation to describe a wide range of physical systems (e.g. microswimmers in a low Reynolds-number regime [9] [10]). Therefore, our equation of movement reads:

$$\dot{\vec{r}}_i = \sqrt{2D}\vec{\xi}_i\,(t)\,. \tag{2}$$

We could integrate the eq. (2) directly, giving as a result that $\vec{r}_i\,(t)$ is a Wiener process (or random walk) [13]. Note that the noise in eq. (2) (and for all SDE in the rest of this work) appears in an additive way, so there is no need to choose an interpretation for the stochastic integrals (Îto, Stratonovich, etc.) [13]. However, it has sense to briefly study the Fokker-Planck equation associated to eq. (2) [14].

$$\partial_t P(\vec{r_j};t) = D\nabla^2 P(\vec{r_j};t). \tag{3}$$

Where $P(\vec{r_j};t)d\vec{r_j}$ is the probability at time t that the particle j is located in the infinitesimal volume $d\vec{r_j}$.[4] There are different ways of solving the eq. (3) (We point out one of these strategies in V A), giving as a result for the probability distribution function (pdf):

$$P(\vec{r_j};t) = \frac{1}{(4\pi Dt)^{\frac{d}{2}}} e^{-\frac{\vec{r}^2}{4Dt}}. \tag{4}$$

Since the standard deviation of the Gaussian distribution grows with time, each particle will be more delocalized as time goes on; so that the $t \to \infty$ limit of equation (4) is an uniform distribution (figure 1c). Indeed, a pure Brownian motion is the microscopical mechanism that leads to the usual macroscopic Fick's diffusion $\left\langle \vec{r}^2(t) \right\rangle = \int_V d\vec{r} P(\vec{r};t)\vec{r}^2(t) = 2dDt$ (figure 1b) [5] in which any initial condition for the density of particles will eventually reach an homogeneous state. In this sense, as the noise will tend to disorder the system, it is usual to think on the strength of the noise as kind of "temperature"-like variable. Moreover, if thermodynamic equilibrium is assumed, the fluctuation-dissipation theorem gives the explicit relationship between the strength of the noise and the temperature $D = \frac{k_B T}{\gamma}$ [14] [17] . In the later situation, the Brownian particles are in thermal equilibrium with its environment so the energy given by the thermal reservoir equals the energy dissipated into the reservoir and there is no entropy production.

---

[4] We could describe the mean density of particles at each point by nothing more than multiplying equation (3) by the homogeneous density $\rho = \frac{N}{V}$. The resulting equation, $\partial_t \rho(\vec{r_j};t) = D\nabla^2 \rho(\vec{r}_i,t)$ is exactly the equation for the first moment (or the deterministic approximation) for the microscopic density $\hat{\rho}(x,t) = \sum_i \delta(x - x_j(t))$ that we could easily derive as the Dean-Kawasaki equation for the Overdamped Brownian motion [3] [15].

[5] Exploring smaller time-scales would require the use of time-correlated stochastic processes that could lead to different kinds of *anomalous diffusion* (e.g. an Orstein-Uhlenbeck process would lead to ballistic diffusion $\left\langle x^2(t) \right\rangle \propto t^2$ for small time-scales [16]).

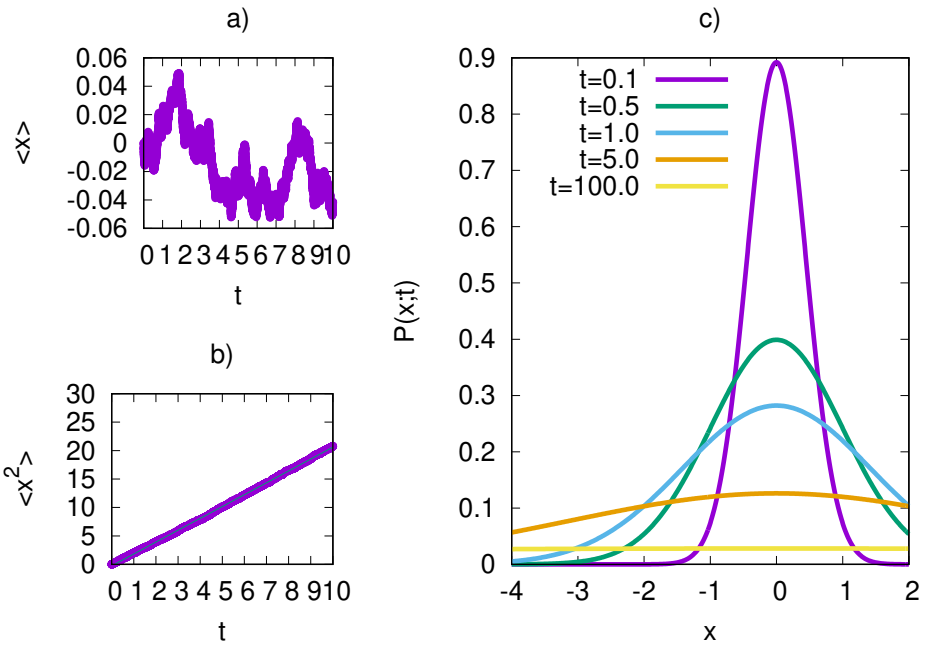


Figure 1: 1D Brownian particle described by equation (2) for $D = 1$ and $x(t = 0) = 0$. a) Mean position as time goes on, isotropy implies that particles will fluctuate around the origin. b) Second moment fitted with a linear regression $\langle x^2(t)\rangle = 2.0847(4)t$, in agreement with Fick's law. c) Pdf (4) for different times.

Active matter

By contrast with the usual driven (passive) particles, active particles can consume internal energy in order to execute movements independently from external influences. There is a wide number of recent reviews which introduce the topic and summarize the state of the art [9] [8] [10] [11] . Most of the published papers have as a starting point the following reasoning: The extraction of energy linked with the self-propelled movement is not balanced with the dissipation of energy to the thermal reservoir, hence there is a net flux of energy that drives the system far from equilibrium [6]. The above thermodynamic argument should convince ourselves that self-propulsion implies an out of equilibrium situation as any unbalanced thermodynamic flux is linked with an entropy production.

Nevertheless, thermodynamic equilibrium is subtle. We could ask ourselves if the self-propulsion mechanism is that different from a Brownian motion that, as we have already seen, can be understood from an equilibrium framework if the dissipation-fluctuation theorem does hold. Moreover, it has been shown that in some active systems it is difficult to find a signature of non equilibrium at large scales since there is an absence of particle flows and entropy production [19] [20] [21].

We would like to give a microscopic argument that could show why active matter can only be described within a non-equilibrium framework[7]. As there is a lack of a well-established formalism for non-equilibrium statistical mechanics, it is an usual strategy to prove general statements on particular

[6] Where we use the term "far from equilibrium" to refer that there is no relaxation towards an equilibrium state. This implies that stationary states will not be characterized by the Boltzmann distribution of the ensemble theory. Moreover, we reject by this expression the possibility of being close to equilibrium, so the entropy production is related to the thermodynamic fluxes through linear relations (Onsager's phenomenological linear theory [18]).

[7] This demonstration is quite trivial for other paradigmatic non-equilibrium models. As for those with absorbing phase transitions, in which the rates from the master equation description show how detailed balance is broken.[12]

models and then try to generalize the results. Following this approach we introduce a really simple active matter model, which is basically an Ornstein-Uhlenbeck (OU) process, and we elucidate exactly why is it a non-equilibrium process:

$$\begin{cases} \dot{\vec{r}} = \vec{\xi}'(t), \\ \left\langle \vec{\xi}'(t) \right\rangle = \vec{0}, \\ \left\langle \xi'_\alpha(t)\, \xi'_\beta(t') \right\rangle = 2\delta_{\alpha\beta} \left(D\delta(t-t') + \frac{U_0^2}{4} e^{-2v_r|t-t'|} \right). \end{cases} \tag{5}$$

Eq. (5) describes a particle performing a thermal Brownian motion ($D = \frac{k_B T}{\gamma}$) plus an OU process with characteristic time $\frac{1}{2v_r}$. This formulation has been largely used [19] [20] [22]. It is of special interest since is more tractable mathematically than others . It is a trivial result that in the limit $|t-t'| \to \infty$, the OU process reduces to a Brownian motion, therefore the system is described by a diffusion equation with diffusive constant $D' = 2D + \frac{U_0^2}{2}$. Actually the strength of the OU process $\frac{U_0^2}{2}$ tell us "how far from equilibrium" is the process (5), as the fluctuation-dissipation theorem does not hold by this quantity. Moreover, it explains why the long time dynamics are difficult to differentiate from those of an equilibrium diffusion process.

There are many other ways of introducing the self-propelled motion. The *fixed norm models*, in which the self-propelled velocity has a constant modulus U_o and a direction of stochastic nature, were inferred from experimental observations (e.g. movement of bacteria *Escherichia coli* [21] and swarming animals [5]). Depending on this stochastic nature, we can diferenciate the *Run-and-tumle particles* (RTP), which are affected by sudden and rapid changes on the direction, and the *Active Brownian Particles* (ABP), which are characterized by an smoother rotational diffusion [2].

Nonetheless, the different models for non-interacting active particles seem to be totally equivalent for the long time dynamics. These equivalence has been proven in different works for the RTPs and ABPs [21] [2], whereas the equivalence between the ABTs and an OU process is mentioned in [1] and explicitly proven in the appendix of this thesis (V C). Therefore, As the fixed norm models reduces to an OU process, they are basically diffusive process for the long time and the simple discussion given before tell us why they are out of equilibrium too. Furthermore, as it has been pointed out recently [1] [21] , we could be in front of a new universality class in which the phenomenology depends on the symmetry and dimensions of the system rather than on specific details of the model.

In principle, interacting active particles do not posses this nice universality property. The interacting forces will appear as drift terms in the diffusion equation, which will depend on the type of active model used [21]. Moreover, the correlations between the interacting forces and the rest of relevant parameters may also affect the fluctuation-dissipation theorem [23]. Therefore, we can not do a general discussion, as we did above, for interacting active particles.

Vicsek Model

The Boids [6] and Vicsek [5] [24] models settled the basis for the microscopic understanding of the collective motion of swarming animals. Basically particles within a distance less than R_V (Vicsek length) steer to point their self-propelled velocities in the same direction, while this behavior is partially hinder by a noise term. It is indeed impressive that the model capture a really complex phenomenology (collective motion, synchronization, order-disorder phase transitions) with such a reduced set of rules. The equation of motion for each ABP reads:

$$\begin{array}{c} \dot{\vec{r}}_i = U_o \hat{n}_i \left(\theta_i\right). \\ \dot{\theta}_i \left(t\right) = \sqrt{2D_r}\xi_i^r(t) + \left\langle \theta_i \left(t\right)\right\rangle_{R_V}. \\ \left\langle \xi_i^r \left(t\right)\right\rangle = 0. \\ \left\langle \xi_i^r \left(t\right) \xi_j^r \left(t'\right)\right\rangle = \delta_{ij}\delta\left(t-t'\right). \end{array} \tag{6}$$

Where:

- $\hat{n}_i$ is the direction of the self-propelled 2D velocity, which is fully characterized by the angle θ_i, $\hat{n}_i \left(\theta_i\right) = \left(\cos\left(\theta_i\right), \sin\left(\theta_i\right)\right)$. U_o is the constant self-propelled speed.

- Since $\{\theta_i\}$ are random angular variables, the stochastic processes $\xi_i^r(t)$ should be distributed according to the wrapped normal distribution (or its approximation, the von Mises distribution) [25]. However, since we are not working with complex functions which could have multi-valuate problems, and following the guidance of [4]; $\xi_i^r(t)$ is distributed according to an usual Gaussian distribution (see V D).

- $\left\langle \theta_i \left(t\right)\right\rangle_{R_V}$ is the averaged direction of the self-propelled velocity of the particles within a distance lower than R_V. Since $\{\theta_i\}$ are random angular variables, the local averages $\left\langle \theta_i \left(t\right)\right\rangle_{R_V}$ have to be computed having into account some subtleties that we summarize in V D.

We could think on the 2D Vicsek model as a kind of XY spin model, where the role of the temperature is substituted by the strength of the noise D_r and the magnetization would be the "degree of alignment" defined as the mean of the modulus of the particle's speed $\vec{v}_j(t) \equiv \dot{\vec{r}}_j(t)$ normalized to U_o:

$$< v >= \frac{1}{NU_0} \sum_i \left|\vec{v}_i\right|. \tag{7}$$

We have done nothing more than defining the control (D_r) and order ($< v >$) parameters for the system. By varying the control parameter and plotting the order parameter we could obtain the classical *phase diagram* for a second order phase transition (figure 2). Actually, due to the locality of the interactions, order cannot arise in very diluted systems; so we should have into account the density of particles $\frac{N}{L}$ as a second control parameter. Therefore this systems displays a *critical line* rather than a single critical point. We remark that the relevance of the density for the order-disorder transitions is a general characteristic of the active matter. We have simulated the 2D Vicsek model for a previous work, it is possible to access the code (Code1) and some videos of the simulations (movie1).

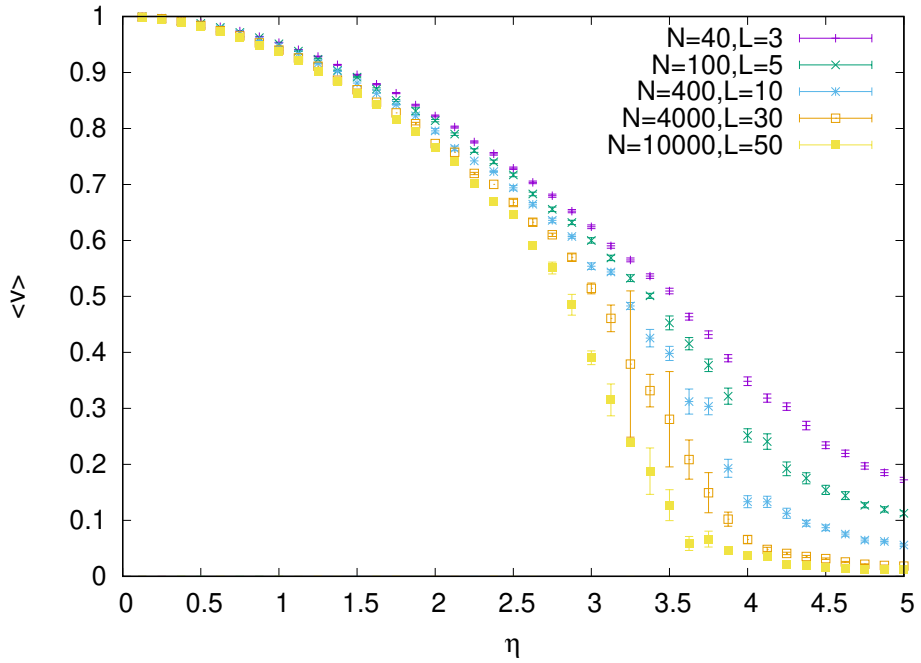


Figure 2: Phase diagram for the 2D Vicsek model for different sizes L and densities. The transition gets sharper as the size of the system grows, in fact, the true phase transition is only obtained for $L \to \infty$ (finite size effects).

Active cluster crystals

Whereas the non-interacting Brownian motion is the microscopical mechanism that leads to a macroscopic diffusion, interacting Brownian particles can have very different behaviors depending on the specific interaction. In particular, it has been shown that confined Brownian particles with repulsive soft-core[8] interactions can give rise to non-homogeneous crystal-like structures [3]. Thus we introduce central conservative forces that carry out the interaction between particles. In particular, the force will be the gradient of a sum of two-body central potentials $\vec{F}(\vec{r}_i) = -\sum_{j\neq i} \nabla_i V(|\vec{r}_i - \vec{r}_j|)$. Following the works [3] and [4], we will study a class of soft-core repulsive potentials, the generalized exponential model of exponent α (GEM-α): $V(|\vec{r}_i - \vec{r}_j|) = \epsilon \exp\left[-\left(\frac{|\vec{r}_i - \vec{r}_j|}{R}\right)^{\alpha}\right]$, being $\epsilon > 0$ and R the constant "strength" and characteristic length of the interaction respectively (see figure 3).

[8] By soft core we mean that the potential does not diverge at $|\vec{r}_i - \vec{r}_j| = 0$, so particles can overlap.

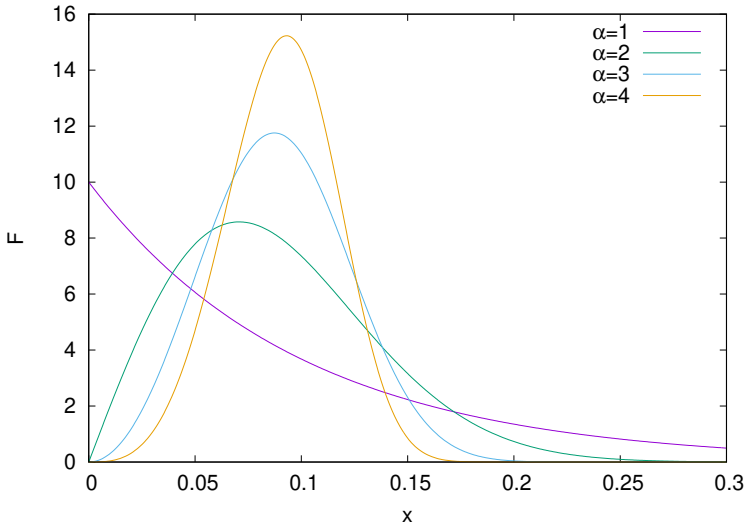


Figure 3: GEM-α force exerted by a particle located at the origin to a second particle located at a distance x from the first one, $F(x) = -\frac{d}{dx}\left[\epsilon \exp\left[-\left(\frac{|x|}{R}\right)^{\alpha}\right]\right]$ with $\epsilon > 0$ (repulsive force), for different values of α.

The work [4] adds an active nature to the interacting Brownian particles introduced above. The resulting equation of motion reads:

$$\begin{aligned} \dot{\vec{r}}_i &= U_o \hat{n}_i(\theta_i) + \vec{F}(\vec{r}_i) + \sqrt{2D}\vec{\xi}_i(t), \\ \dot{\theta}_i(t) &= \sqrt{2D_r}\xi_i^r(t). \end{aligned} \tag{8}$$

Summarizing, eq. (8) describes N ABPs with repulsive interactions submitted to a thermal Brownian motion. We remark that the self-propelled directions are not coupled as in the Vicsek model ([5], sec. II), so we do not expect any rotational symmetry breaking for this system. Our aim is to develop a numerical study of eq. (8) in two dimensions. In order to enclosure the system in a LxL box without losing the translational symmetry of an infinite system, we will implement the usual periodic boundary conditions (PBC). Since eq. (8) are additive stochastic differential equations, it is suitable to integrate them within a Euler-Muruyama algorithm (see sec. V B and [13] for details) , as it has be done in previous works [3] [4]. The naive implementation of the code is of order $\mathcal{O}\left(N^2\right)$, since we have to compute the distance having into account the PBC for each pair of particles (i, j). Nevertheless, we remark that we reduce the number of operations to $\frac{N^2-N}{2}$ just having into account the third Newton's law, that is, it is only needed to compute the force for each pair $(i, j > i)$ since $\vec{F}(i \to j) = -\vec{F}(j \to i)$ and $\vec{F}(i \to i) = 0$.[9]

Despite having a complex parameter space, we can clearly differentiate three phases. When the Brownian motion or self propulsion terms dominates over the repulsive force, particles will eminently diffuse, generating homogeneous states from any initial condition. However, for crowded environments ($\rho = \frac{N}{L}$ large) and sufficient low values of U_o and D (parameters which tune the diffusion), the particles will be organized into equidistant clusters creating an hexagonal pattern (figure 4b). We shall

[9] Due to the long range nature of our interactions, we are not able to use more sophisticated algorithms as the Link-Cell method [24], that allows to simulate the dynamics of short range interacting particles (e.g. The Vicsek model[5]) as an $\mathcal{O}(N)$ problem (see implementation at Code1). In the context of molecular dynamics, the Ewald summation techniques [26] are $\mathcal{O}(N \log N)$ methods that are used to deal with long range interactions with "infinite" systems (the system is not infinite, but there is a hierarchy of copies of the LxL cell). We can not use these kind of Fourier algorithms, since our particles will only interact with each other using the closer distance, we implicitly introduce a cut off for the interaction range.

remark the non-triviality of this result, individual repulsive interactions pile up generating an effective attractive force that sustains the clusters. The probability of particle exchange between clusters will depend on U_o and D.

Whereas there are some examples of cluster formation caused by the self propelled motion [9], the clusters that we study here are the result of having long-range repulsive interaction between enclosed particles. In fact, we obtain the same phenomena described in the paragraph above for the passive version of the model [3].

The third possibility is obtained when diffusion is really low but the density is high enough so there is pattern formation. The distribution of particles inside each cluster becomes non-homogeneous (particles will stay close to the edges) (figure 4a) (movie2). It is really interesting that this effect is not present for passive particles [3], hence it is linked with the non-equilibrium nature of active particles and we expect that it will disappear for small persistence length (small U_o or large D_r). Therefore, this exemplifies that the self-propelled motion can cause effects that differ from a Brownian-like diffusion (even for long times) when there is interaction between particles.

This is the general picture for the GEM-$\alpha \geqslant 2$ potentials. However, for $\alpha < 2$, the late time state is always statistically homogeneous (there is no cluster formation) [3]. From this result, we deduce that a "peaked interaction" (figure 3) is needed to generate the confinement collective behavior.

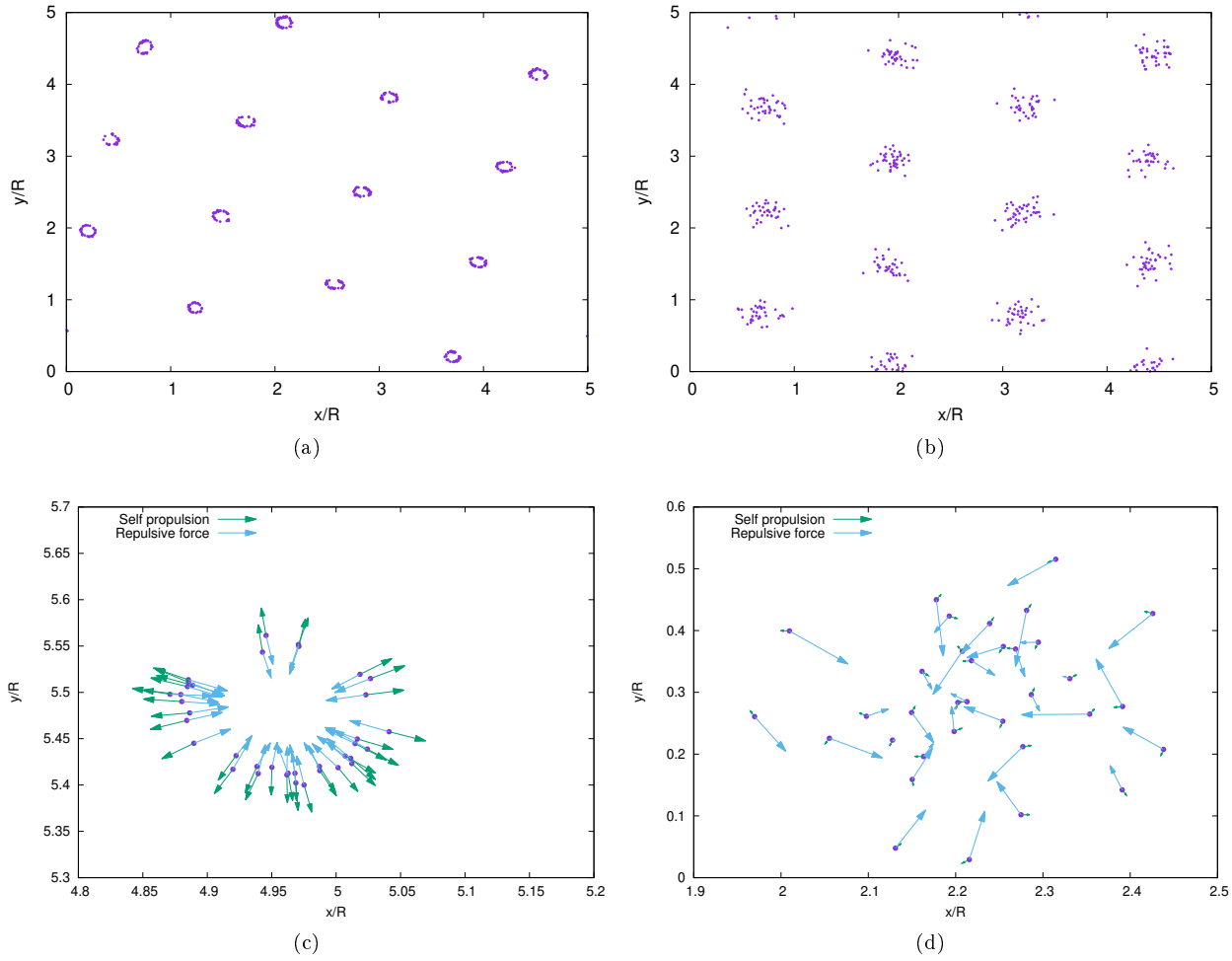
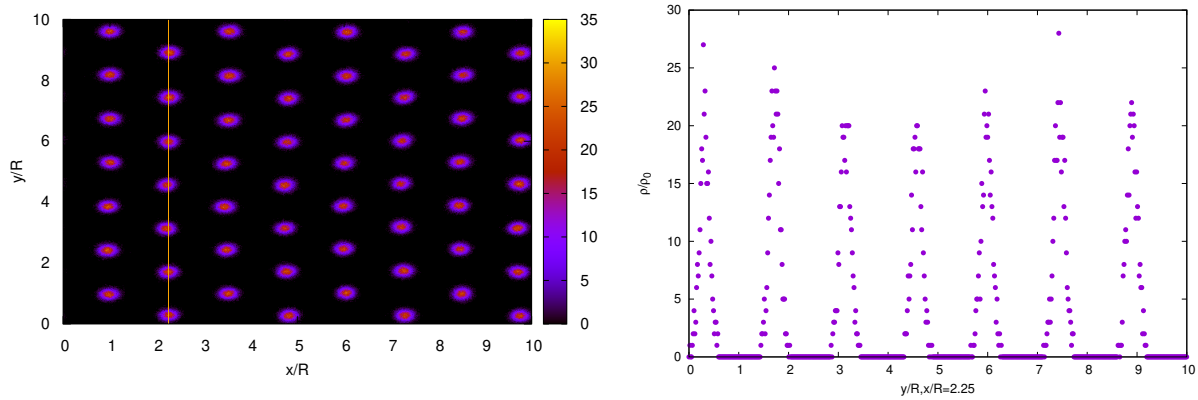


Figure 4: (a) and (b): Snapshots of stationary state obtained from a numerical simulation of eq. (8). (c): zoom of one of the empty clusters of (a). (d): zoom of one of the filled clusters of (b). *Parameters:* $L = 1$, $\epsilon = 0.0333$,$R = 0.1$, $\rho = 2000$, $D_r = 0.1$, $\alpha = 3$. (a) $D = 3 \cdot 10^{-2}$ and $U_o = 1$. (b) $D = 10^{-4}$ and $U_o = 3$.

It is of special interest the spontaneous synchronization that arise between the collective repulsive force and the individual self propulsion. Note that these two terms are not coupled in equation (8). Moreover their nature is very different: whereas the repulsive force is deterministic , depends on the distance between the particles and merge from the sum of all the interactions; the self-propulsion has a stochastic nature and is independent of the state of the particles. Nevertheless, these two terms are constantly balanced (see figure 4c and movie2). It is remarkable that these non-trivial correlations between the position of the particle and the self-propulsion arise in various active models [22], stressing one more time the relevance of the interactions in active systems.

A more quantitative study of the structure of the clusters is supported by the figure 5. We firstly wait a thermalization time to avoid the inclusion of transient states into the statistics. Then, we discretize the space into 0.002×0.002 squared cells and compute the local density in these squares. The coarse-grained density is averaged on time so we can have an idea of what the mean average density field of the system looks like. By studying a section of the 3D plot of figure 5a, we can estimate the mean distance between clusters (a) and the mean cluster size. The mean distance between clusters is quite robust

whereas the cluster size is highly dependent on D. For $D \to 0$, the clusters will collapse into a point; as D increases the cluster size will grow until the clusters disappear and the system is homogeneous.



(a) (b)

Figure 5: Particles inside cluster diffuse within a characteristic radius for the same parameters as in figure 4b. (a): Heat map of coarse grained time-averaged density. (b): Values of the coarse grained time-averaged density in the section signaled at figure 5a. From it, we compute the mean distance between clusters $\frac{a}{R} = 1.37(2)$ and the mean cluster size.

III. THE MODEL

We now study a combination of all the contributions previously commented in sec. II. The novelty of this model comes from adding the Vicsek-like alignment interaction to the active crystals described above, in sec. II [4]. Thus, we generalize the model [4] in a simple but fundamental way. As interactions will depend not only in the distance between particles but on the orientation of the self-propelled motion too, we expect to observe the usual rotational symmetry breaking of the spin models. The equations of motion will be similar to eq. (8), except for the presence of a Vicsek-like interaction (6):

$$\begin{aligned} \dot{\vec{r}}_i &= U_o \hat{n}_i\left(\theta_i\right) + \vec{F}\left(\vec{r}_i\right) + \sqrt{2D}\vec{\xi}_i\left(t\right). \\ \dot{\theta}_i\left(t\right) &= \sqrt{2D_r}\xi_i^r(t) + \left\langle \theta_i\left(t\right)\right\rangle_{R_V}. \end{aligned} \tag{9}$$

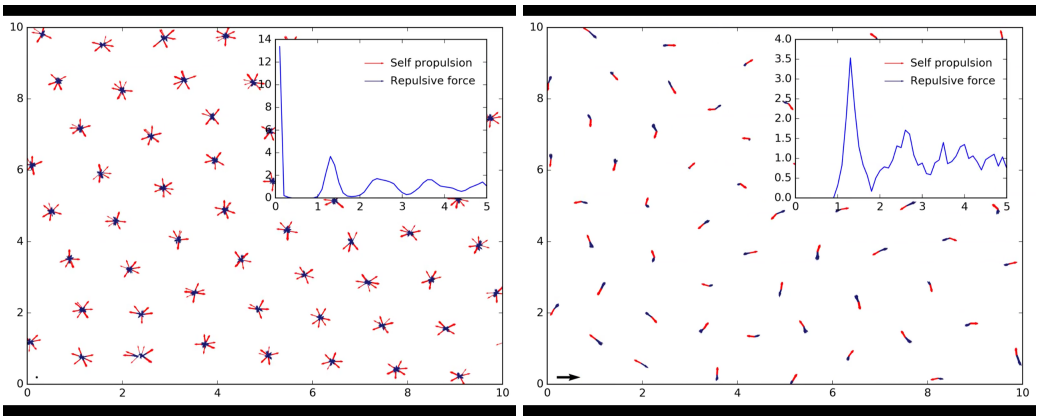
Therefore, the self-propelled directions will be locally coupled. At the same time, there is a long-range (but weak) GEM-α repulsive interaction depending on the distance between particles. Hence, there are three characteristic distances for the system: R (repulsive interaction distance), R_v (alignment or Vicsek distance) and L (system size). Whereas these distances are parameters that define the model, the mean cluster size and the distance between clusters (a) are emergent distances that characterize the stationary state when there is cluster formation.

We will draw on dimensional arguments as well as fixing parameters to reduce the parametric complexity of the system. In particular, we focus on values of D_r such that a pure Vicsek model would be in the ordered state (see figure 2) so the differences with the active crystals without alignment interactions will be stressed.

The FORTRAN code that was built for integrating eq. (9) with PBC in a $L \times L$ square box is shared on (Code2) and has free license.

Low diffusion regime: effect on empty clusters

We focus our study in the low diffusion regime ($D \to 0$) where the active clusters without Vicsek interaction ([4], sec. II) showed the "empty cluster" structure (see figures 4a and 4c). Our aim is to give a phenomenological description of the behavioral dependence of the system with R_V (which model the locality of the alignment interaction).



(a) $\frac{R_V}{R}$=0.01, $\langle v \rangle = 0.02(2)$.(movie4).

(b) $\frac{R_V}{R}$=0.5, $\langle v \rangle = 0.06(2)$.(movie5).

(c) $\frac{R_V}{R}$=1, $\langle v \rangle = 0.72(2)$.(movie6).

(d) $\frac{R_V}{R}$=2, $\langle v \rangle = 0.99(2)$.(movie3).

Figure 6: *Main plots:* Typical snapshots of $N = 2000$ particles evolving according to eq. (9) for different values of R_V. *Inset:* radial distribution of particles. *a)* Metastable empty-clusters state that can be observed for low enough values of R_V. Radial distribution in clusters ensures $< v > \to 0$. Mean radius of the clusters and mean distance between particles can be extracted from inset (first and second peaks of the radial distribution respectively). *b)* Angular diffusion has destroyed the empty clusters, local symmetry breaking mechanism ensures $< v > \to 0$. *c)* Clusters partially interact by Vicsek-alignment interaction, thus $< v >$ increase. *d)* $R_V > a$, clusters always interact by alignment interaction so $< v > \to 1$. *Parameters:* $L = 1$, $\epsilon = 0.0333$, $\rho = 2000$, $D_r = 0.1$, $\alpha = 3$, $D = 10^{-4}$, $R = 0.1$ and $U_o = 3$.

Our background in stochastic processes tell us that the study of single realizations is an useful tool, complementary to the systematic statistical analysis of the system [13]. We observe that the alignment interaction creates a local ($R_V \sim 0$) or universal ($R_V \sim L$) collective preferred direction for the self-propelled movement. The main observation is that this spontaneous symmetry breaking will destroy the empty clusters (figure 6). Basically, when two particles are aligned, the collective force push them

into the same point; once they collapse, they will not repel each other (figure 3) and the Brownian motion is not strong enough to avoid this stable state. These alingment-and-collapse effect is nearly instantaneous when R_V is of the order of magnitude of the cluster size (movie3). When R_V is smaller than the typical radius of the clusters, there is a metastable state of empty clusters (or "cruxes") that will be eventually destroyed by the angular noise.

Since we were concerned about the precision of the stability analysis of the empty crystals, we decided to use a more efficient integration algorithm (which allows us to explore larger times or to increase the precision without an increase of computational time). In particular, we change the numerical integration method for a Heun algorithm (see sec. V B for a brief presentation), in principle it does not generate any advantage since its convergence is of order $\mathcal{O}\left(h^{\frac{3}{2}}\right)$ as the Euler-Maruyama method, however, the improvement comes from the integration of the deterministic part which will have a $\mathcal{O}\left(h^3\right)$ convergence. We decided that it was a suitable option for this particular problem since one of the stochastic contributions (the translational Brownian motion) is not relevant.

The quantity $\langle v \rangle$ defined for first time in 7 has to be redefined in order to maintain its meaning as "degree of alignment" of the self-propelled directions as:

$$< v >= \frac{1}{N} \sum_i \left| \hat{n}_i \left(\theta_i \right) \right| \tag{10}$$

It is needed to look for some quantity that can carry information about the spatial order and the quality of the pattern, The radial distribution function $g(r)$ is the density of probability for finding a particle at a distance r of another given particle (see details about meaning and normalization at [4] , Wiki and inside the scripts of Code2). The defects of the pattern are related with peaks of $g(r)$, the distribution is more peaked when there is a good pattern. We also can measure the mean distance between clusters and the mean size of the clusters from the position of the peaks (insets of figure 6).

Regarding the pattern formation, as $R_V \to 0$, we should recover the quasi-static hexagonal structure modeled by equation (8). Due to the radial distribution of the self-propelled velocities in the empty clusters of the transient state (see figure 6a, movie4), $\langle v \rangle$ is close to zero. Whereas if R_V is of the order of magnitude of the system (L), the hexagonal pattern will be moving within a constant speed as a rigid body; all the clusters would have collapsed within the same direction so $\langle v \rangle = 1$ (see figure 6d and movie3).

The interesting situation lays between these two behaviors, where all the clusters have already collapse, but local consensus will not spread through the entire system. This will generate deformations so the hexagonal symmetry is continuously broken by the self-propelled motion and restored by the collective attractive force (see figure 6b, figure 6c). When varying continuously the parameter R_V, there is a transition on $\langle v \rangle$ that somehow maps the structural transition on the system. Therefore, the "Vicsek's order parameter" maintains its informative relevance for this model.

Phase diagrams

Inspired by the previous analysis, we seek for a more systematic discussion of the variation of $\langle v \rangle$ with the rest of parameters. The self-propelled direction could resemble one more time the order parameter of a paramagnetic-ferromagnetic phase transition [5]. Despite obtaining phase diagrams that reinforce this idea (figures 7 and 8), it also could not be a true thermodynamic phase transition [3]. To probe the existence of such phase transition we should keep on with the usual tests like looking for the power laws that relate control and order parameters, the power law divergence of the correlation lengths and susceptibility and the finite size scaling properties [12] [27]. However, this probe is out of the scope of this work. Hence, as in previous papers ([3] and [4]), when referring to states such as the ones displayed

at figure 6, we do not imply the existence of a true phase transition but simply highlight that the distribution for the particle positions and velocities are very different depending on the values of some relevant parameters.

In order to explore the complex parametric space, it has sense to define a set of dimensionless parameters with well defined physical interpretations. To do so, we partially follow the prescription given by [4]:

$$\begin{gathered} L = 1, \rho = \frac{N}{L} = 2000, \alpha = 3, R = 0.1 \rightarrow \text{Fixed.} \\ D_r = 0.1 \rightarrow \text{Fixed (choice of time units).} \\ \widetilde{U} = \frac{U}{\sqrt{D D_r}} \rightarrow \text{Compares strength of self propulsion and translational diffusion.} \\ \widetilde{\epsilon} = \frac{\epsilon D_r}{U^2} \rightarrow \text{Compares strength of interaction forces with self-propulsion.} \\ \widetilde{l} = \frac{U}{D_r R} \rightarrow \text{Compares persistence and interaction lengths.} \end{gathered} \tag{11}$$

Apart from this, all lengths will be normalized by the characteristic length of the repulsive interaction R as before.

We start by re-doing the analysis of the previous section III in a systematic way. Keeping constant the translational diffusion parameter (D) and all the other parameters as in 11, after waiting a proper thermalization time (inplot of figure 7), we compute the time average of $\langle v \rangle$. Repeating this process for different values of R_v and D we obtain the phase diagram showed in the figure 7. We could define something like a critical value for R_V that separates an aligned (or ordered) phase and an isotropic (or disordered) phase, this critical value is of the order of magnitude of the average distance between clusters, a. As discussed before, all the particles in a cluster will end up having the same self-propelled direction. However, when $R_v < a$, the clusters cannot interact by the Vicsek interaction and therefore they will not have the same self-propelled direction; since the spontaneous symmetry breaking respects the rotational symmetry (the preferred direction is "chosen" randomly for each cluster) we expect $\langle v \rangle$ to to be close zero. The contrary occurs for $R_v > a$, when a preferred direction will spread through the entire system.

We can explain that $\langle v \rangle$ is not exactly zero at the ordered phase in figure 7. First of all, the number of clusters in the system is of the order of fifty, we expect that when the number of clusters approach infinity ($N \rightarrow \infty$ and $L \rightarrow \infty$), the sum of the preferred directions will be closer to zero. Secondly, we saw in the previous section III that, when preferred directions are chosen locally, hexagonal symmetry is sometimes broken with defects so the distance between particles is lower than a and they could be able to interact via the Vicsek interaction. If the last statement is true, the alignment should be enhanced by increasing the strength of the self-propelled movement. This is exactly what we can see in figure 8, were we have obtained another order-disorder transition by varying $\widetilde{l}$ and keeping constant the rest parameters of 11 and the Vicsek radius ($R_V < a$). When the persistence length[10] is high enough (the exact value will depend on $\widetilde{\epsilon}$), the cluster's mobility is such that they can overcome the repulsive forces commissioned to maintain the hexagonal symmetry.

It is also important to note the role of D, producing a drift in the critical position of R_V. When translational diffusion is enhanced, particles explore the space easily, the jumps of particles between clusters is more likely. This is somehow analogous to letting the clusters interact by the Vicsek interaction.

---

[10] The persistence length $\frac{U_0}{D_r}$ is the mean distance that an active particle travels due to its self-propelled movement in an straight line (like the "Run" of a RTP).

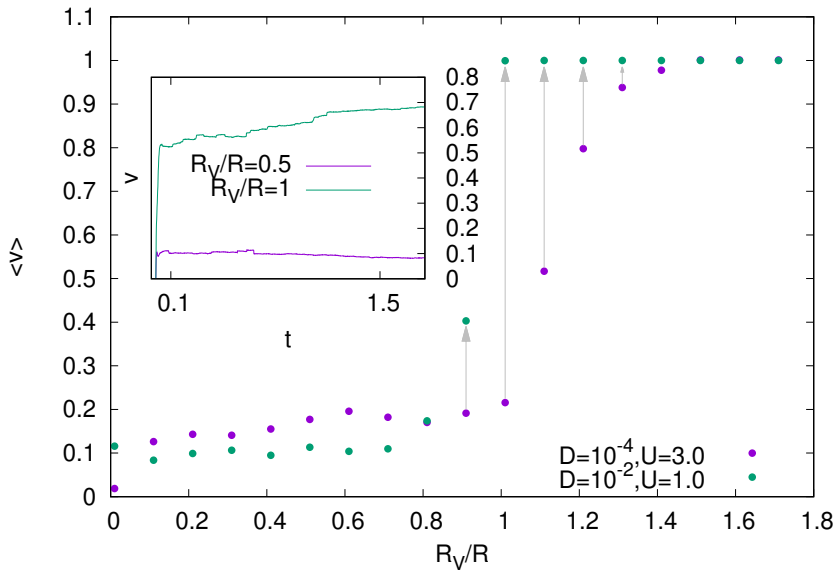


Figure 7: *Main plot:* Degree of alignment $< v >$ for different values for the alignment radius (R_V) and the translational diffusive constant (D). *Inset:* Thermalization depends on the value for R_V we fix the thermalization time as 2 (in $\frac{1}{h}$ units, being h the integration step, see V B).

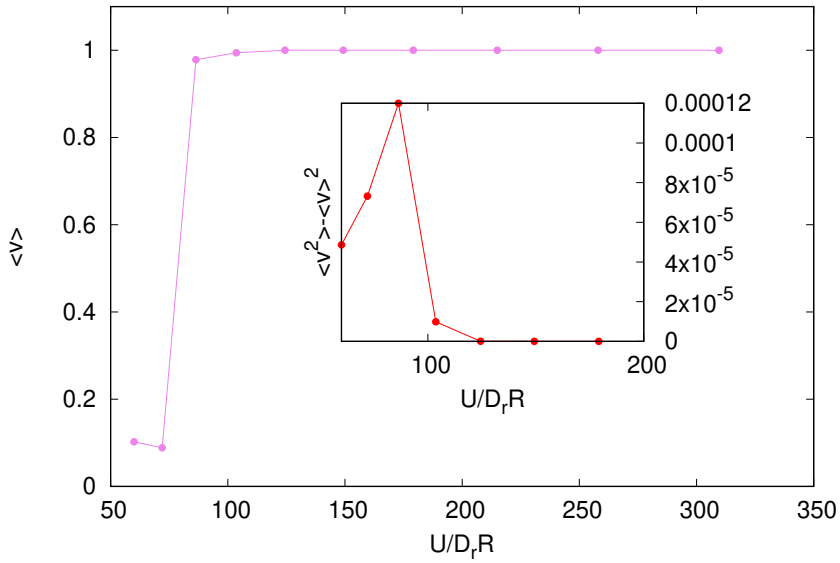


Figure 8: *Main plot:* "Degree of alignment" $< v >$ as we vary the ratio $\widetilde{l}$ between the characteristic lengths of the problem (equation (11)) for $\frac{R_V}{R} = 0.5$. The rest of dimensionless parameters were constants when varying $\widetilde{l}$. *Inset:* Variance, it would be a susceptibility in an equilibrium framework. it gives us some insights about the possibility of having a true phase transition (susceptibility should diverge on the critical point for $N \to \infty$ and $L \to \infty$).

Despite not being visible, all plots (figures 7 and 8) have errorbars that inform about the statistical errors. We had into account either the variance (inplot in 8) and the errors introduced by the temporal correlations between measures (see [13] for a discussion about the computation of statistical errors). Despite having the typical susceptibility blow up (inplot in 8), it is strange for us that the errors are so low (even close to the transition). We know that the divergence of the susceptibility in a thermodynamic phase transition is only obtained in the thermodynamic limit ($N \to \infty$ and $L \to \infty$), but the situation remains unusual, moreover when large fluctuations are also a common behavior of active matter [1].

Large diffusion regime

We now address the question of what happens if we increase the relevance of the Brownian motion, lets remember that the empty clusters would disappear from the pure active crystal cluster model without alignment interaction (section II). The main difference is that the Brownian motion avoids the particle collapse (see figure 9 and movie7). As mobility gets enhanced, the relevance of the repulsive interactions will get lost and the degrees of freedom related with the self-propelled motion will behave as in the ordered phase of a pure Vicsek model (section II) (figure 10).

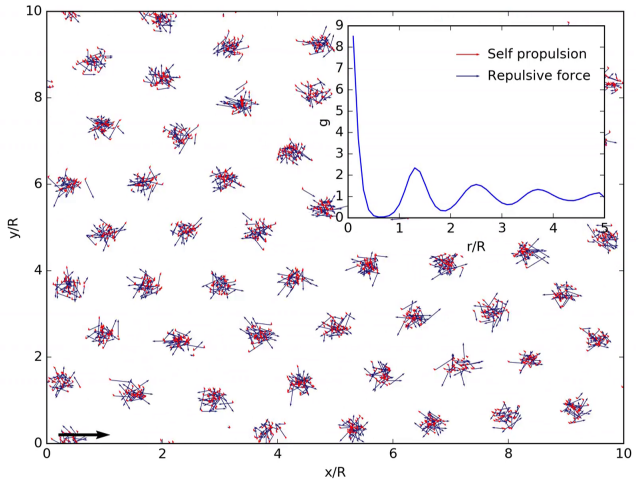


Figure 9: Snapshot from movie7. Brownian motion avoids that particles collapse into a point. The self-propelled motion of all particles is aligned ($< v >= 0.98(2)$) so the pattern will move as a rigid body. *Parameters* $D = 10^{-2}$, $\frac{R_V}{R} = 0.9$, $U_0 = 1.0$ (the rest of them are fixed as in (11)).

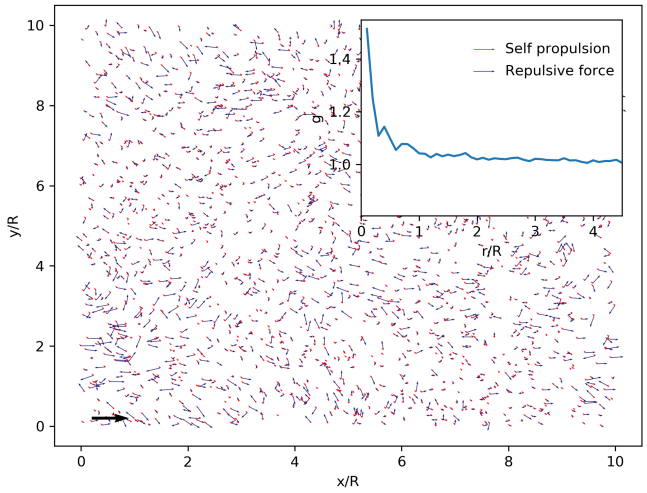


Figure 10: Brownian motion generates statistically homogeneous state (as we can see from the radial distribution function of the inset). Alignment gets enhanced by the free mobility ($< v >= 0.95(2)$). *Parameters* $D = 1$, $\frac{R_V}{R} = 0.7$, $U_0 = 1.0$ (the rest of them are fixed as in (11)).

IV. CONCLUSIONS

Through our interest in developing a self-contained work, we got into fundamental and up-to-date active matter problems. All the different models of non-interacting active matter have long time dynamics that can be described through a Fick's diffusive equation like eq. (3). The summary of our argument is that all non-interacting active models can be somehow mapped into an OU process (see example in V C) that can be understood as a Brownian motion at large scales, which is nothing more than the microscopic mechanism that leads to a a Fick's diffusion (see II). Since this diffusion equation will have a diffusion constant with is not in agreed with the fluctuation-dissipation theorem, such a system will perform far from equilibrium.

However, different active models will do lead to different results when dealing with interacting particles. These differences appears explicitly in the field description of the models, where the diffusion equation is drifted by terms that depend on the interactions [15] [21]. Moreover, there are effects that just arise from the conjunction of deterministic interactions and an active nature; the appearance of "empty clusters" in [4] (sec. II) that can not be observed in the passive version of the model [3] is a good example of this. The relevance of the interactions gets one more time stressed since non-trivial correlations between the positions and active parameters can arise due to the presence of interactions (see figure 4c). This all makes us think in a possible application of this knowledge: it could be possible to discern the active model that is more suitable to describe real active particles by applying a potential (or ,alternatively, that it could be possible to identify the interactions presence in a real set of well-identified active particles).

The novelty in this thesis was the study of how a Vicsek alignment interaction affects the behavior of active clusters. We found out that the degree of alignment $\langle v \rangle$, which is an order parameter for the Vicsek model analogous to the magnetization on a spin model, results very useful in the behavioral description of the system since it does not only inform on the alignment of the self-propelled components of the velocity but on the "quality" of the hexagonal structure.

We focus our attention in the low diffusion range, because in this regime the relevance of the active nature of the system is stressed through the presence of the empty clusters. We kind of regard the empty clusters as a micelle-like structure, since this kind of structures are of great importance in

biology and chemistry, it would have been of great interest to induce a coherent displacement on this "micelles" thanks to the Vicsek interaction. However, it is a negative result of this thesis that the Vicsek interaction will eventually destroy all the empty clusters diffusion as a low Brownian motion implies that particles that get aligned will collapse into the same point. It would be a question interesting to address in future works if the desired behavior would be observed in hard core active particles.

When varying the relevant parameters of the system (equation (11)) to observe the effect on $\langle v \rangle$, we observe the same kind of phase diagrams obtained from a pure Vicsek-model. The system pass from a isotropic situation ($\langle v \rangle = 0$) to another one in which a preferred self-propelled direction have been chosen ($\langle v \rangle = 1$) through a kind of abrupt symmetry breaking mechanism. Whereas the extrema values of $\langle v \rangle$ are associated with a proper hexagonal symmetry for the system, in the intermediate values, some structural defects will arise.

The main result is that to get a preferred direction that spreads through the entire system ($< v > \rightarrow 1$), we need the same requirements for the strength of the noise and density of particles of the pure Vicsek model and one of the following situations:

- The radius for the Vicsek interaction (R_V) is greater than the mean distance between clusters (a). That is, $R_V > a$ at every time.
- The self-propelled motion is such that hexagonal structure gets deformed (locally in time and space) so the distance between clusters can be lower than R_V. That is, $R_V > a$ locally.

In the large diffusion regime, spatial disorder (homogeneous state) lives together with order in the self-propelled directions. Basically the Brownian mobility makes it easier for the particles to interact by the Vicsek interactions, in fact, we can observe the drift in the phase diagrams induced by the Brownian movement in the figure 7 . The behavioral picture for $D \rightarrow \infty$ is a Brownian motion with a constant drift term U_o in the velocity.

V. APPENDIX

A. Diffusion equation: dimension argument

Eq. (3) can be solved by using well-known strategies (e.g. separation of variables, doing the Fourier transform). However, we find it really easy to solve the problem using dimensional considerations. We first identify the dimensions of the key magnitudes of the problem [11].

$$[D] = \frac{L^2}{T} \qquad [P] = \frac{1}{L^d}. \tag{12}$$

Where d are the spatial dimensions. Therefore we can write the probability distribution function as:

$$P(\vec{r};t) = \frac{1}{(Dt)^{\frac{d}{2}}} f(\vec{r};t). \tag{13}$$

Where $f(\vec{r};t)$ is a *dimensionless* probability distribution function. The key point of the argument is that there is just a fundamental dimensionless quantity that can be constructed since the problem has just a parameter, D. This quantity is $\vec{\xi} = \frac{\vec{r}}{\sqrt{Dt}}$, hence the dependence of the dimensionless ansatz 13 is restricted to $f\left(\vec{\xi}\right)$. Therefore, we can write eq. (3) in a more convenient way:

$$\partial_t P(\vec{r};t) = -\frac{1}{2D^{\frac{d}{2}} t^{\frac{d}{2}-1}} \left(d + \vec{\xi}\cdot\nabla_\xi\right) f\left(\vec{\xi}\right). \tag{14}$$

$$D\nabla^2 P(\vec{r};t) = -\frac{D}{D^{\frac{d}{2}} t^{\frac{d}{2}-1}} \nabla_\xi^2 f\left(\vec{\xi}\right). \tag{15}$$

So eq. (3) reads:

$$\nabla_\xi \cdot \left[2D\nabla_\xi f\left(\vec{\xi}\right) + f\left(\vec{\xi}\right)\vec{\xi}\right] = 0. \tag{16}$$

Assuming that there is no flux of probability, that is, fixing the integration constant equal to zero:

$$2D\nabla_\xi f\left(\vec{\xi}\right) + f\left(\vec{\xi}\right)\vec{\xi} = 0. \tag{17}$$

Basically we have reduced the problem from a partial differential equation to an ODE. Actually, the eq. (17) can be seen as a system of d identical ODEs (one for each component of $\vec{\xi}$). Since the problem is isotropic, we can factorize the function $f\left(\vec{\xi}\right) = \prod f_\alpha\left(\xi_\alpha\right)$. The ODE $2D\frac{d}{d\xi_\alpha} f_\alpha\left(\vec{\xi}\right) + f_\alpha\left(\xi_\alpha\right)\xi_\alpha = 0$ has trivial solution $f_\alpha\left(\xi_\alpha\right) = \frac{1}{(4\pi)^{\frac{1}{2}}} e^{-\frac{\xi_\alpha^2}{4}}$, where the prefactor was obtained imposing normalization. Finally, from equation (13):

$$P(\vec{r};t) = \frac{1}{(4\pi Dt)^{\frac{d}{2}}} e^{-\frac{\vec{r}^2}{4Dt}}. \tag{18}$$

---

[11] Where $[y]$ means "dimensions of the object y" and L and T are space and time units respectively.

B. Numerical methods for solving SDE

In this section, we briefly summarize the fundaments of the numerical methods for integrating SDE used in this work. A more detailed introduction to these algorithms can be found at [13]. We focus on the numerical solutions of additive SDEs with white noise of the form:

$$\begin{array}{c} \dot{x}(t) = f(x(t),t) + \sqrt{D}\eta\left(t\right). \\ \langle \eta\left(t\right)\rangle = 0 \qquad \langle \eta\left(t\right)\eta\left(t'\right)\rangle = \delta(t-t'). \end{array} \tag{19}$$

Being $f(x(t),t)$ a deterministic term.

- Euler-Muruyama algorithm:

  We start by integrating equation (19) between t and $t+h$:

$$x(t+h) = x(t) + \int_t^{t+h} ds\, f(x(s),s) + \sqrt{D}\int_t^{t+h} ds\, \eta(s). \tag{20}$$

  The integral of the deterministic part can be approximated by $hf(x(t),t)$ since we consider $f(x(t),t)$ a well-behaved function and h small enough (this is indeed the key point of the Euler method for solving ODEs). The stochastic term can be integrated easily if we remember that the white noise is the derivative of a Wiener process $W(t)$ [13]. Therefore:

$$x(t+h) = x(t) + hf(x(t),t) + \sqrt{D}\left[W(t+h) - W(t)\right]. \tag{21}$$

  The difference of a Wiener process in different times $[W(t+h) - W(t)]$ is a Gaussian stochastic process (sum of Gaussians is still a Gaussian). When time is discretized in steps of width h, the evaluation of this stochastic process is analogous to the generation of a random gaussian variable $\sqrt{h}u_t$ with $\langle u_t u_{t'}\rangle = \delta_{tt'}$ and $\langle u_t\rangle = 0$ (since a Gaussian random variable is fully characterized by its mean and correlations). Therefore:

$$x(t+h) = x(t) + hf(x(t),t) + \sqrt{hD}u_t + \mathcal{O}\left(h^{\frac{3}{2}}\right). \tag{22}$$

- Heun method:

  Here we have to apply the same ideas than before, we just change the way of integrating the deterministic part from an Euler method to Runge-Kutta of order 2 prescription:

$$\begin{array}{c} k \equiv hf(x(t),t). \\ x(t+h) = x(t) + \frac{h}{2}\left(f(x(t),t) + f(x(t)+k,t+h)\right) + \sqrt{hD}u_t + \mathcal{O}\left(h^{\frac{3}{2}}\right). \end{array} \tag{23}$$

  Comparing the efficiency of both algorithms, there is a trade-off between a more precise algorithm and the need of two evaluations of $f(x(t),t)$ at each step.

C. Equivalence between Active Ornstein-Uhlenbeck process and Active Brownian particle.

Our aim is to demonstrate that an active particle whose 2D movement is described by the equations:

$$\begin{cases} \dot{\vec{r}} = U_0 \hat{v}\left(\theta\left(t\right)\right) + \vec{\eta}^{\,T}\left(t\right). \\ \dot{\theta} = \eta\left(t\right). \\ \left\langle \eta\left(t\right)\right\rangle = 0 \qquad \left\langle \vec{\eta}^{\,T}\left(t\right)\right\rangle = \vec{0}. \\ \left\langle \eta\left(t\right)\eta\left(t'\right)\right\rangle = 2v_r\delta(t-t') \qquad \left\langle \eta_\alpha^T\left(t\right)\eta_\beta^T\left(t'\right)\right\rangle = 2D\delta_{\alpha\beta}\delta(t-t'). \end{cases} \tag{24}$$

Is equivalent to a pure translatiotal Ornstein-Uhlenbeck process.

$$\begin{cases} \dot{\vec{r}} = \vec{\xi}'\left(t\right). \\ \left\langle \vec{\xi}'\left(t\right)\right\rangle = \vec{0}. \\ \left\langle \xi'_\alpha\left(t\right)\xi'_\beta\left(t'\right)\right\rangle = 2\delta_{\alpha\beta}\left(D\delta(t-t') + \frac{U_0^2}{4}e^{-2v_r|t-t'|}\right). \end{cases} \tag{25}$$

Where U_0 represents the constant modulus of the self-propelled speed, the Greek letters label components of Euclidean vectors and all the stochastic processes are Gaussians (and therefore characterized by its mean and variance).

Basically, the demonstration is based in the formal integration of the angular part of equation (24) and the statistical characterization of the stochastic process $\hat{v}\left(\theta\left(t\right)\right) \equiv \cos\left(\theta\left(t\right)\right)\hat{i} + \sin\left(\theta\left(t\right)\right)\hat{j}$.

Formally, integrating equation (24):

$$\theta\left(t\right) = \int_{t_0}^{t} \eta\left(s\right)\, ds + \theta\left(t_0\right). \tag{26}$$

Where $\theta\left(t_0\right)$ is a random variable with an a priory well-known pdf (Gaussian, uniform, delta,...). We assume that $\theta\left(t_0\right)$ is a Gaussian random variable uncorrelated with the stochastic process $\eta\left(s\right)$. Therefore, $\theta\left(t\right)$ is also a Gaussian stochastic process (the sum of normal distributed random variables is also a normal distributed random variable). It will be very useful to compute the first moment and the temporal correlations of $\theta\left(t\right)$.

$$\begin{cases} \left\langle \theta\left(t\right)\right\rangle = \left\langle \theta\left(t_0\right)\right\rangle + \int_{t_0}^{t} \cancelto{0}{\left\langle \eta\left(s\right)\right\rangle} ds \equiv 0. \\ \left\langle \theta\left(t\right)\theta\left(t'\right)\right\rangle = \underbrace{\left\langle \theta\left(t_0\right)^2\right\rangle}_{\equiv\sigma^2} + 2\int_{t_0}^{t} \cancelto{0}{\left\langle \eta\left(s\right)\right\rangle}\left\langle \theta\left(t_0\right)\right\rangle\, ds + \int_{t_0}^{t}\int_{t_0}^{t'} \left\langle \eta\left(s\right)\eta\left(s'\right)\right\rangle\, dsds' \equiv \sigma^2 + 2\nu_r \min\left\{t-t_0, t'-t_0\right\}. \end{cases} \tag{27}$$

Where $\langle\rangle$ in this context means averaging over different realizations of the noise $\eta(t)$ and the random variable $\theta(t)$. Without loosing generality, we fix $t_0 = 0$ so $\left\langle \theta\left(t\right)\theta\left(t'\right)\right\rangle = \sigma^2 + 2\nu_r \min\left\{t, t'\right\}$. It is useful to write down the explicit dependence of $\vec{v}$ with $\theta\left(t\right)$ using exponentials.

$$\hat{v}\left(\theta\left(t\right)\right) = \frac{e^{i\theta(t)} + e^{-i\theta(t)}}{2}\hat{i} + \frac{e^{i\theta(t)} - e^{-i\theta(t)}}{2}\hat{j}. \tag{28}$$

Using the results of expression 27, we compute all the possibilities for $\left\langle \nu_\alpha\left(\theta\left(t\right)\right)\nu_\beta\left(\theta\left(t'\right)\right)\right\rangle$.

▷

$$\langle \cos(\theta(t)) \cos(\theta(t')) \rangle = \frac{\left\langle e^{i(\theta(t)+\theta(t'))} \right\rangle + \left\langle e^{-i(\theta(t)+\theta(t'))} \right\rangle + \left\langle e^{i(\theta(t)-\theta(t'))} \right\rangle + \left\langle e^{i(-\theta(t)+\theta(t'))} \right\rangle}{4}. \tag{29}$$

Using the well-known result for Gaussian random variables: $\langle e^x \rangle = e^{\langle x^2 \rangle}$.Therefore:

$$\langle \cos(\theta(t)) \cos(\theta(t')) \rangle = \frac{e^{-\left\langle (\theta(t)-\theta(t'))^2 \right\rangle} + e^{-\left\langle (\theta(t)+\theta(t'))^2 \right\rangle}}{2} = \frac{e^{-2v_r|t-t'|} + e^{-\left(t+t'+2\min\{t,t'\}+4\sigma^2\right)}}{2}. \tag{30}$$

▷

$$\langle \sin(\theta(t)) \cos(\theta(t')) \rangle = \frac{\left\langle e^{i(\theta(t)+\theta(t'))} \right\rangle - \left\langle e^{-i(\theta(t)+\theta(t'))} \right\rangle + \left\langle e^{i(\theta(t)-\theta(t'))} \right\rangle - \left\langle e^{i(-\theta(t)+\theta(t'))} \right\rangle}{4i} = 0. \tag{31}$$

By the same way, $\langle \sin(\theta(t')) \cos(\theta(t)) \rangle = 0$.

▷

$$\langle \sin(\theta(t)) \sin(\theta(t')) \rangle = -\frac{\left\langle e^{i(\theta(t)+\theta(t'))} \right\rangle + \left\langle e^{-i(\theta(t)+\theta(t'))} \right\rangle - \left\langle e^{i(\theta(t)-\theta(t'))} \right\rangle - \left\langle e^{i(-\theta(t)+\theta(t'))} \right\rangle}{4} = \tag{32}$$

$$= \frac{e^{-2v_r|t-t'|} - e^{-\left(t+t'+2\min\{t,t'\}+4\sigma^2\right)}}{2}.$$

Hence, we obtain the classic correlations for an Ornstein-Uhlenbeck process applying any of the limits $t, t' \to \infty$, that is, when the system is "far from the initial condition". *We remark that this limit differs from the one that we should apply to obtain a white noise from an OU process,* $|t - t'| \to \infty$.

D. Directional Statistics

Directions on the plane can be regarded as unitary vectors on $\mathcal{R}^2$ or as a complex numbers of modulus one, then directions are characterized by an angle.

$$\hat{r} = \cos(\theta)\,\hat{i} + \sin(\theta)\,\hat{j} \qquad r = e^{-i\theta} \tag{33}$$

Since the angle θ and $\theta + 2\pi$ gives the same point on the unit circle, it is a good practice to use modulo 2π arithmetics to deal with angles [25] (it avoids, for example, that certain complex functions becomes multi-valuated as the complex logarithm). Angular random numbers can be generated according to probability distribution functions which respect the chosen cut for the unit circle ($[0, 2\pi[$, $[-\pi, \pi[$, ...) as the wrapped normal distribution or the Von Mises distribution [25]. Nevertheless, since it is computationally expensive to deal with this kinds of distributions we will use the usual real probability distribution functions, the only trade off of this choice is that the angular averages have to be computed carefully.

Angular averages

Let $\{\theta_i\}_{i=1,..,N}$ be the set of N angles which are the outcomes from a random experiment. It is clear that the usual definition for the average ($< \theta >= \frac{1}{N}\sum_i \theta_i$) is meaningless, since it depends on where

the circle is cut. As an example, if we consider the outcomes $\theta_1 = 2\pi - x$ and $\theta_2 = x$, then $< \theta >= \pi$; we look for a definition of the average such that in this situation $< \theta >= 0$. We follow the prescription given by [25] and [5], which is based on making averages over the sine and the cosine:

$$< \theta >= \arctan 2 \left(\frac{\frac{1}{N} \sum_i \sin(\theta_i)}{\frac{1}{N} \sum_i \cos(\theta_i)} \right) \tag{34}$$

We remind that the arctan function cannot distinguish between the first and the third quadrant an the second and forth, to do so we need to take into account the sings of the denominator and numerator; this is done by the $\arctan 2$ function. As a last remark, computing averages of angles according to 34 maintains a pathological error: since we are basically computing the direction of the resultant of the vectors defined from 33, if the resultant is zero the averaged angle will be one more time meaningless (the origin on $\mathcal{R}^2$ has not a well defined angle in polar coordinates).

E. URL

Code1: Code and documentation for Vicsek model simulations.

https://github.com/jvrglr/Flocking

Code2: Code and documentation for the simulation of Active Clusters with Vicsek-like interactions.

https://github.com/jvrglr/Master-Thesis

Wiki: Information about radial distribution function.

https://en.wikipedia.org/wiki/Radial_distribution_function

Movie1: Transcient and stationary states for Vicsek model in ordered phase.

https://www.youtube.com/watch?v=KwlSgCWQt4k

Movie2: Empty active crystal clusters without Vicsek interactions.

https://www.youtube.com/watch?v=j4bS2HwzLqU

Movie3: Active Clusters with Vicsek interaction, $R_V \sim L$ (system size). Preferred direction spreads through the entire system.

https://www.youtube.com/watch?v=JQ3KJnAH-CI&feature=youtu.be

Movie4: Active Clusters with Vicsek interaction, $R_V \sim 0$. Transient empty clusters.

https://www.youtube.com/watch?v=WdLyVVRROZO

Movie5: Active Clusters with Vicsek interaction, $R_V < a$ (distance between clusters).

https://www.youtube.com/watch?v=DZwftbqdwXE&feature=youtu.be

Movie6: Active Clusters with Vicsek interaction, $R_V \lesssim a$ (distance between clusters).

https://www.youtube.com/watch?v=71ADtP7tsk4&feature=youtu.be

Movie7: Active Clusters with Vicsek interaction, Brownian motion avoids that the particles collapse into a point.

https://youtu.be/NeXH7kobWcE

F. Dictionary of abbreviations and summary of stochastic processes

ABP: Active Brownian Particles.
GEM-α: generalized exponential model of exponent α.
ODE: Ordinary Differential Equation.
PBC: Periodic Boundary Conditions.
PDF: Probability Distribution Function.
PDE: Partial Differential Equation.
RTP: Run-and-Tumble Particles.
SDE: Stochastic Differential Equation.

We have only dealt with Gaussian stochastic processes that are therefore characterized by its mean and correlations:

- Translational white noise: $\vec{\xi}_i(t)$

$$\begin{gathered}\langle \xi_i^\alpha(t)\rangle = 0. \\ \left\langle \xi_i^\alpha(t)\,\xi_j^\beta(t')\right\rangle = \delta_{ij}\delta_{\alpha\beta}\delta(t-t').\end{gathered}$$

- OU process + white noise: $\vec{\xi}'(t)$

$$\begin{gathered}\left\langle \vec{\xi}'(t)\right\rangle = \vec{0}. \\ \left\langle \xi'_\alpha(t)\,\xi'_\beta(t')\right\rangle = 2\delta_{\alpha\beta}\left(D\delta(t-t') + \tfrac{U_0^2}{4}e^{-2v_r|t-t'|}\right).\end{gathered}$$

- Rotational white noise: $\xi_i^r(t)$

$$\begin{gathered}\langle \xi_i(t)\rangle = 0. \\ \langle \xi_i(t)\,\xi_j(t')\rangle = \delta_{ij}\delta(t-t').\end{gathered}$$

[1] Yaouen Fily and M. Cristina Marchetti. Athermal phase separation of self-propelled particles with no alignment. *Physical review letters*, 108(23):235702, 2012.

[2] Alexandre P. Solon, ME Cates, and Julien Tailleur. Active brownian particles and run-and-tumble particles: A comparative study. *The European Physical Journal Special Topics*, 224(7):1231–1262, 2015.

[3] Jean-Baptiste Delfau, Hélène Ollivier, Cristóbal López, Bernd Blasius, and Emilio Hernández-García. Pattern formation with repulsive soft-core interactions: Discrete particle dynamics and dean-kawasaki equation. *Physical Review E*, 94(4):042120, 2016.

[4] Jean-Baptiste Delfau, Cristóbal López, and Emilio Hernández-García. Active cluster crystals. *New Journal of Physics*, 19(9):095001, 2017.

[5] Tamás Vicsek, András Czirók, Eshel Ben-Jacob, Inon Cohen, and Ofer Shochet. Novel type of phase transition in a system of self-driven particles. *Physical review letters*, 75(6):1226, 1995.

[6] Craig W. Reynolds. Flocks, herds and schools: A distributed behavioral model. *ACM SIGGRAPH computer graphics*, 21(4):25–34, 1987.

[7] Philip W. Anderson et al. More is different. *Science*, 177(4047):393–396, 1972.

[8] Pawel Romanczuk, Markus Bär, Werner Ebeling, Benjamin Lindner, and Lutz Schimansky-Geier. Active brownian particles. *The European Physical Journal Special Topics*, 202(1):1–162, 2012.

[9] Clemens Bechinger, Roberto Di Leonardo, Hartmut Löwen, Charles Reichhardt, Giorgio Volpe, and Giovanni Volpe. Active particles in complex and crowded environments. *Reviews of Modern Physics*, 88(4):045006, 2016.

[10] Jens Elgeti, Roland G. Winkler, and Gerhard Gompper. Physics of microswimmers single particle motion and collective behavior: a review. *Reports on progress in physics*, 78(5):056601, 2015.

[11] M. Cristina Marchetti, Jean-François Joanny, Sriram Ramaswamy, Tanniemola B. Liverpool, Jacques Prost, Madan Rao, and R. Aditi Simha. Hydrodynamics of soft active matter. *Reviews of Modern Physics*, 85(3):1143, 2013.

[12] M. Henkel, H. Hinrichsen, and S. Lubeck. Non-equilibrium phase transtions: Volume 1: Absorbing phase transitions, 2008.

[13] Raúl Toral and Pere Colet. *Stochastic numerical methods: an introduction for students and scientists.* John Wiley & Sons, 2014.

[14] N. Gr. van Kampen. *Stochastic processes in physics and chemistry.* 1992, 1995.

[15] David S. Dean. Langevin equation for the density of a system of interacting langevin processes. *Journal of Physics A: Mathematical and General*, 29(24):L613, 1996.

[16] Anna S. Bodrova, Aleksei V. Chechkin, Andrey G. Cherstvy, Hadiseh Safdari, Igor M. Sokolov, and Ralf Metzler. Underdamped scaled brownian motion. 2016.

[17] Rep Kubo. The fluctuation-dissipation theorem. *Reports on progress in physics*, 29(1):255, 1966.

[18] Dilip Kondepudi and Ilya Prigogine. *Modern thermodynamics: from heat engines to dissipative structures.* John Wiley & Sons, 2014.

[19] Étienne Fodor, Cesare Nardini, Michael E. Cates, Julien Tailleur, Paolo Visco, and Frédéric van Wijland. How far from equilibrium is active matter? *Physical review letters*, 117(3):038103, 2016.

[20] Cesare Nardini, Étienne Fodor, Elsen Tjhung, Frédéric Van Wijland, Julien Tailleur, and Michael E Cates. Entropy production in field theories without time-reversal symmetry: Quantifying the non-equilibrium character of active matter. *Physical Review X*, 7(2):021007, 2017.

[21] ME Cates and J. Tailleur. When are active brownian particles and run-and-tumble particles equivalent? consequences for motility-induced phase separation. *EPL (Europhysics Letters)*, 101(2):20010, 2013.

[22] Grzegorz Szamel. Self-propelled particle in an external potential: Existence of an effective temperature. *Physical Review E*, 90(1):012111, 2014.

[23] WF Van Gunsteren and HJC Berendsen. On the fluctuation-dissipation theorem for interacting brownian particles. *Molecular Physics*, 47(3):721–723, 1982.

[24] Francesco Ginelli. The physics of the vicsek model. *The European Physical Journal Special Topics*, 225(11):2099–2117, Nov 2016.

[25] Kanti V. Mardia and Peter E. Jupp. *Directional statistics*, volume 494. John Wiley & Sons, 2009.

[26] Abdulnour Y Toukmaji and John A. Board Jr. Ewald summation techniques in perspective: a survey. *Computer physics communications*, 95(2-3):73–92, 1996.

[27] Alessandro Attanasi, Andrea Cavagna, Lorenzo Del Castello, Irene Giardina, Stefania Melillo, Leonardo Parisi, Oliver Pohl, Bruno Rossaro, Edward Shen, Edmondo Silvestri, et al. Finite-size scaling as a way to probe near-criticality in natural swarms. *Physical review letters*, 113(23):238102, 2014.

[28] Davide Loi, Stefano Mossa, and Leticia F. Cugliandolo. Effective temperature of active matter. *Physical Review E*, 77(5):051111, 2008.