

Motility-Induced Phase Separation of harmonic Active Ornstein-Uhlenbeck particles fulfilling Detailed Balance.

Author: Lior Tetro

Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain.

Advisor: Demian Levis

Abstract: Active matter constitutes a class of nonequilibrium systems which has attracted a lot of attention over the past decades, and Motility-Induced Phase Separation (MIPS) lies among one of its most salient collective phenomena. Although MIPS has been studied in depth, the question of whether this active phase transition can only occur in the presence of out-of-equilibrium fluctuations remains open. In this work, we numerically show that harmonic Active Ornstein-Uhlenbeck Particles (AOUP), with an equilibrium dynamics fulfilling detailed balance, undergo MIPS. We studied analytically 2-body interactions, identifying an effective attraction increasing with activity, to show that n-body effects are needed in order to account for the nature of the transition, and in particular, of the dense phase. Finally, a recursive method to obtain a multibody expansion of the partition function for the complete system is developed.

I. INTRODUCTION

Although many-body systems in thermodynamic equilibrium have been studied and generally understood for over a century now, the generic behavior of non-equilibrium systems constitutes a fundamental open questions in physics. One of the difficulties for studying such systems is the variety of situations from which they arise, and the lack of a general theoretical framework to describe them. One of the classes that can be defined in non-equilibrium statistical physics is active matter [1]. These are many-body systems whose elementary units consume energy from their environment to perform some kind of tasks, and in particular, self-propel. Examples include both living systems (such as flocks of birds or bacteria colonies) and synthetic ones (self-diffusophobic Janus colloids, polar grains, robot swarms, etc.) [2]. Each constituent, or active particle, of the system has an irreversible dynamics, dissipating energy at the microscopic level [3].

Important insight has been gained from the investigation of simple model systems, extending the tools of statistical mechanics to these non-equilibrium soft matter, some of the most archetypical examples being Run-and-Tumble Particles (RTP) or Active Brownian Particles (ABP) [4, 5]. In both cases what is observed is that, although the modulus of the velocity of the particles is constant, its orientation follows a stochastic process. RTP change their direction of motion at a given rate, and ABP does also so by rotational diffusion. But, the question of a minimal model of active matter that manifests the main collective and individual features representative of this class, has motivated the introduction of Active Ornstein-Uhlenbeck Particles (AOUP) [6]. This type of particles is characterized by a self-propulsion force that follows an Ornstein-Uhlenbeck process. The non-equilibrium nature of the microscopic dynamics arises from the existence of memory, characterized by a persistence time.

Active matter exhibits new behavior that is not

observed in equilibrium systems, being one of the most studied one Motility-Induced Phase Separation (MIPS) [7][8]. It can be observed that self-propelled particles with purely repulsive interactions can undergo a phase separation, like a liquid-gas phase separation, in the absence of attractive forces, emerging from the mere competition between self-propulsion and excluded volume. When active particles collide, the intrinsic persistence of their dynamics delays their separation, blocking each other for a period of time, scaling with their persistence time. At large enough densities, collisions become more likely, and thus a uniform suspension can eventually become unstable, triggering a full phase-separation between a dilute active gas in coexistence with a denser phase with reduced motility (see FIG. 1).

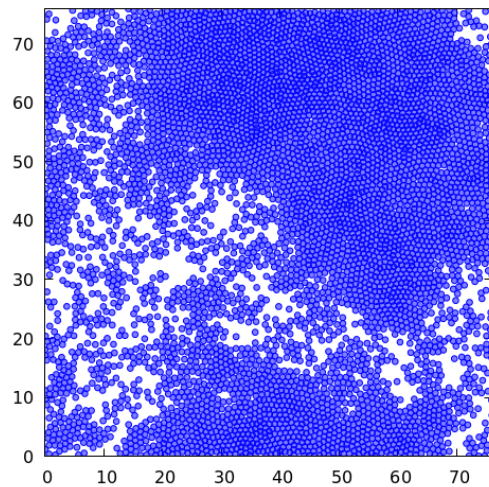


FIG. 1. Snapshot of a system of purely repulsive $N = 6000$ AOUP with $\tau = 100, \sigma = 1, k = 10, \phi = 0.8$. The system phase separates into a dense and dilute phases.

How to characterize the non-equilibrium behavior of active matter is a question that has attracted a great

deal of attention. Although the dynamics include self-propulsion, stemming from an energy injection and the single-particle level, the question of how specifically in an interacting system detailed balance is broken, is far from trivial. If we focus on AOUP, it is possible to obtain the entropy production of the system through a path integral formulation of its stochastic dynamics. The result obtained shows that a system of AOUP has an entropy production proportional to the third derivatives of the interaction potential [9]. It is also possible to address this question from the Fokker-Planck (FP) equation of the system, yielding the same result [10, 11]. Then, these results show that a system of AOUP with a potential with vanishing third derivatives has an equilibrium dynamics, and as such, one can derive its steady-state distribution function (i.e. a generalization of the Boltzmann distribution). These results do not hold for ABP, so it is a specific result for this microscopic dynamics governed by an Ornstein-Uhlenbeck process.

Normally, when system fulfills detailed balance, and so, in the regime of equilibrium statistical mechanics, the steady-state statistics is described by the Boltzmann probability distribution, where the exponential of the interaction potential appears. But in this case, as will be reproduced afterwards, the equilibrium regime is described by a probability that is not the regular Boltzmann one. What we obtain in the exponential is an effective interaction that deviates in a non-trivial manner from the original one. So, even though we normally think of an equilibrium system as one with the Boltzmann distribution, here the equilibrium behavior is demonstrated by a vanishing entropy production rate and the fulfillment of detailed balance.

Recently, the mechanism behind MIPS has been revisited by exploring in detail the role of multi-body collisions. What has been seen for ABP in three dimensions, is that the 2-body effective interactions are insufficient to engender phase separation. It is necessary to consider multibody effects that are associated with caging of particles to fully account for the phase separation observed in simulations. The idea is that as activity increases, a cage of particles is developed, creating a dense phase of "caged particles". Normally, the number of neighbors needed to see this cage is around the coordination number of liquids [12].

The importance of the multibody effects in MIPS and the existence of an equilibrium regime for a system of AOUP guide us directly to ask if in the equilibrium regime of AOUP the system undergoes MIPS if the density (ϕ) and activity (τ) are large enough.

This directly questions the fundamental nature of the transition, is it purely a nonequilibrium effect, or otherwise, is it driven by equilibrium-like n-body effects and can be seen as an effective equilibrium transition? This is the problem we address in this work. First, we will introduce AOUP and study the simple analytic results that can be obtained in the absence of interactions. Then we will introduce the general Fokker-Planck framework to

extract the detailed balance condition and the probability distribution of AOUP in the equilibrium regime. We follow with the numerical results for a system of AOUP interacting via a harmonic potential (zero third derivatives) to investigate whether MIPS occurs in this context. To investigate multibody effects, we will study analytically and numerically the 2-body problem, showing that activity gives rise to an effective pair-wise attraction. As we will show, this is not enough to explain the advent of MIPS. We thus turn into the 3-body problem and show analytically, the existence of non-trivial N-body effects that defy the simple picture arising from the 2-body calculation. Finally, as we are in an equilibrium problem, we will sketch a systematic way to obtain the partition function of the N-body problem.

II. THE MODEL: ACTIVE ORNSTEIN-UHLENBECK PARTICLES

We consider a system of N particles, located at \mathbf{r}_i , moving in a $L \times L$ plane with periodic boundary conditions (PBC). Their evolution is governed by the following overdamped Langevin equations, with a noise $\xi(t)$ following an Ornstein-Uhlenbeck process with a characteristic persistence time τ and a diffusion constant D :

$$\begin{aligned}\dot{\xi}_i^\alpha(t) &= -\frac{\xi_i^\alpha(t)}{\tau} + \frac{\sqrt{2D}}{\tau}\eta_i^\alpha(t) \\ \dot{r}_i^\alpha(t) &= -\partial_i^\alpha\Phi + \xi_i^\alpha(t)\end{aligned}\quad (1)$$

Here the latin index labels the particles and the greek index the space coordinate. Here $\eta(t)$ is a white noise with the following properties

$$\begin{aligned}\langle \eta_i^\alpha(t)\eta_j^\beta(t') \rangle &= \delta(t-t')\delta_{ij}\delta^{\alpha\beta} \\ \langle \eta_i^\alpha(t) \rangle &= 0\end{aligned}\quad (2)$$

The particles will interact via a pair-wise harmonic short-range potential with a characteristic length σ and a characteristic energy scale k .

$$\Phi(r_{ij}) = \begin{cases} \sum_i \sum_{j < i} k \left(1 - \frac{r_{ij}}{\sigma}\right)^2 & r_{ij} < \sigma \\ 0 & r_{ij} > \sigma \end{cases}\quad (3)$$

where $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ is the separation between particles i and j . (The reason why we make this choice of potential will become clear in the next section.) The Ornstein-Uhlenbeck process that models the activity through a self-propulsion of the AOUP, is decoupled from the position Langevin equation. Then, the colored noise satisfies the following self-correlation function,

$$\begin{aligned}\langle \xi_i^\alpha(t)\xi_j^\beta(t') \rangle &= \frac{D}{\tau}\delta_{ij}\delta^{\alpha\beta}e^{-|t-t'|/\tau} \\ \langle \xi_i(0)^2 \rangle &= 2D/\tau\end{aligned}$$

We have a stochastic dynamics with memory for the evolution of r_i that is described by a characteristic correlation time τ .

A. Free AOUP

If the system is non interacting $\Phi = 0$, the equations can be solved analytically. The Ornstein-Uhlenbeck process has the following evolution

$$\xi^\alpha(t) = \xi^\alpha(0)e^{-t/\tau} + \frac{\sqrt{2D}}{\tau}e^{-t/\tau} \int_0^t e^{u/\tau} \eta^\alpha(u) du$$

that in general depends on the initial condition of $\xi(t)$. If we introduce this result in the evolution of the position and solve for r we have,

$$r(t) = \xi(0)\tau(1-e^{-t/\tau})^2 + \sqrt{2D} \int_0^t \left(1 - e^{-|t-u|/\tau}\right) \eta(u) du$$

which is the trajectory of a single AOUP. These coupled equations have two free parameters $\xi(0)$ and $r(0)$. If now, we want to obtain the mean squared displacement (MSD), $\langle(r(t) - r(0))^2\rangle$, we take the average over the noise, and integrating by parts the position we have

$$\langle r(t)^2 \rangle = 4D\tau \left(\frac{t}{\tau} - (1 - e^{-t/\tau}) \right) \quad (4)$$

If we look at the behavior at short and long times, we obtain respectively a ballistic and a diffusive regime:

$$\begin{aligned} \langle r(t)^2 \rangle &= 4\frac{D}{\tau}t^2 = 4v_0^2t^2 \quad \text{for } t \rightarrow 0 \\ \langle r(t)^2 \rangle &= 4Dt \quad \quad \quad \text{for } t \rightarrow \infty \end{aligned} \quad (5)$$

where for the ballistic regime we can define the velocity $v_0 = \sqrt{D/\tau}$.

B. Numerical simulations

With the previous analytical results, it is now possible to run simulations of free AOUP and compare the analytical and numerical results, as a check of the consistency and accuracy of the code we developed for this project. To perform simulations, we discretize the equations of motion under the Euler-Mayurama scheme [13]

$$\begin{aligned} r_i^\alpha(t + \Delta t) &= r_i^\alpha(t) + \xi^\alpha \Delta t - \partial_i^\alpha \Phi \Delta t \\ \xi_i^\alpha(t + \delta t) &= \xi_i^\alpha(t) - \frac{\xi_i^\alpha}{\tau} \Delta t + \frac{\sqrt{2D}}{\tau} g_i^\alpha \sqrt{\Delta t} \end{aligned} \quad (6)$$

where g_i^α is a random number with zero mean and variance equal to one. As shown in Fig. 2 and 3, the numerical simulations reproduce the analytical expectations accurately.

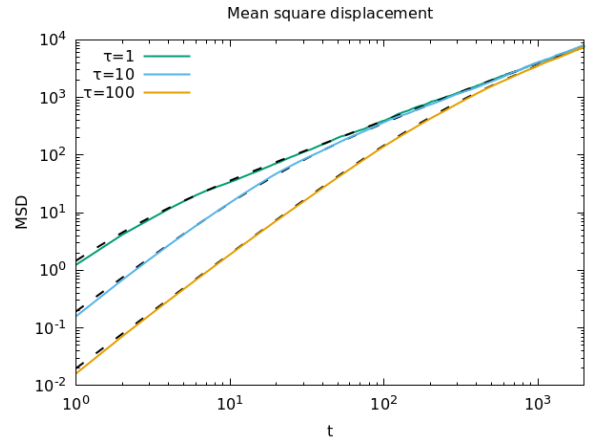


FIG. 2. MSD at a constant diffusion parameter for increasing τ . The dotted lines correspond to the exact solution of the model eq. 4. The crossover time towards a diffusive regime is at $t \sim \tau$. With $D = 1$ and a time step $\Delta t = 0.01$.

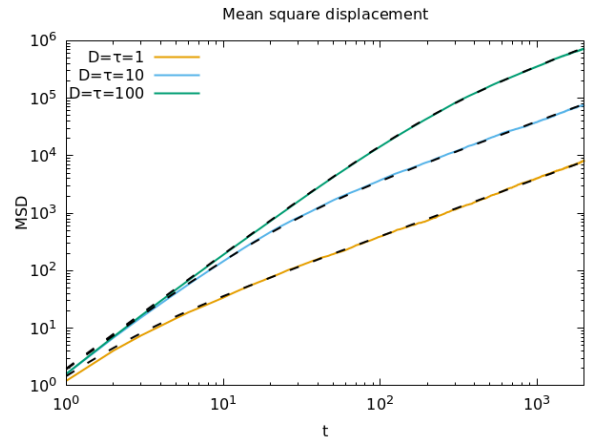


FIG. 3. MSD at a constant velocity v_0 for increasing $\tau = D$. The dotted lines correspond to the exact result. We can observe that at the ballistic regime the three cases collapse as the self-propulsion velocity is identical in all three cases.

III. GENERAL STOCHASTIC DYNAMICS

Now that we are familiarized with AOUP we need to demonstrate that a system of AOUP with a harmonic interaction is in equilibrium. We will follow the formalism introduced by Sara dal Cengio et al in [11].

A. General aspects

For a system consisting on N dynamic variables $\Gamma \doteq \{\Gamma_i\}_{i=1}^N$, we can introduce a probability distribution $\Psi(\Gamma, t)$ that describes our system at any time t . The time evolution will be governed by the operator $\Omega_0(\Gamma)$

as,

$$\partial_t \Psi(\mathbf{\Gamma}, t) = \Omega_0(\mathbf{\Gamma}) \Psi(\mathbf{\Gamma}, t) \quad (7)$$

In general, if an initial condition is specified $\Psi(\mathbf{\Gamma}, 0)$, the partial differential equation can be formally solved as

$$\Psi(\mathbf{\Gamma}, t) = e^{\Omega_0 t} \Psi_0(\mathbf{\Gamma}, 0) \quad (8)$$

For Markovian dynamics the operator takes the general Fokker-Planck (FP) form, i.e., for a constant diagonal diffusion,

$$\Omega_0(\mathbf{\Gamma}) = \sum_i \left(-\partial_i A_i(\mathbf{\Gamma}) + \sum_j \partial_i \partial_j D_i \delta_{ij} \right) \quad (9)$$

where the partial derivatives are taken under the N dynamic variables $\mathbf{\Gamma}$, A_i are the drift vectors and D_i are the diffusion constants.

B. Symmetry aspects

To obtain the detailed balance condition, we must identify the parity under time-reversal of $\mathbf{\Gamma}$. Let \mathcal{T} be the time-reversal operator, then, under its mapping $\mathcal{T} : \mathbf{\Gamma} \rightarrow \epsilon \mathbf{\Gamma} \doteq \epsilon_i \Gamma_i$. Variables with $\epsilon_i = +1$ are said to be even and $\epsilon_i = -1$ odd. For example, with 1 particle $\mathcal{T} : (r, p) \rightarrow (r, -p)$, where r is the position and p is the momentum.

From the FP equation we can define the following flux

$$J_i(\mathbf{\Gamma}, t) \doteq A_i(\mathbf{\Gamma}) \Psi(\mathbf{\Gamma}, t) - D_i \partial_i \Psi(\mathbf{\Gamma}, t) \quad (10)$$

It is now possible to decompose the drift in a reversible and irreversible part, by considering the nature under time-reversal.

$$\begin{aligned} A_i^{rev}(\mathbf{\Gamma}) &\doteq \frac{1}{2}(A_i(\mathbf{\Gamma}) - \epsilon_i A_i(\mathbf{\Gamma})) \\ A_i^{irr}(\mathbf{\Gamma}) &\doteq \frac{1}{2}(A_i(\mathbf{\Gamma}) + \epsilon_i A_i(\mathbf{\Gamma})) \end{aligned} \quad (11)$$

Finally, we can decompose the total probability flux in the following reversible and irreversible contributions,

$$\begin{aligned} J_i^{rev}(\mathbf{\Gamma}) &= A_i^{rev}(\mathbf{\Gamma}) \Psi(\mathbf{\Gamma}, t) \\ J_i^{irr}(\mathbf{\Gamma}) &= A_i^{irr}(\mathbf{\Gamma}) \Psi(\mathbf{\Gamma}, t) - D_i \partial_i \Psi(\mathbf{\Gamma}, t) \end{aligned} \quad (12)$$

Under time-reversal the irreversible current transforms as

$$J_i^{irr}(\mathbf{\Gamma}) = \epsilon_i J_i^{irr}(\mathbf{\Gamma}) \quad (13)$$

C. Detailed balance

A system is in equilibrium if any microscopic process is balanced by the reversed one at stationarity. This is the detailed balance condition (DB). We can formally write it as,

$$P(\mathbf{\Gamma}_f, t_f | \mathbf{\Gamma}_0, t_0) \Psi_0(\mathbf{\Gamma}_0) = P(\epsilon \mathbf{\Gamma}_0, t_0 | \epsilon \mathbf{\Gamma}_f, t_f) \Psi_0(\epsilon \mathbf{\Gamma}_f) \quad (14)$$

where $P(\mathbf{\Gamma}_f, t_f | \mathbf{\Gamma}_0, t_0)$ is the probability of a given stochastic trajectory in phase space, starting at $\mathbf{\Gamma}_0$ at t_0 and ending at $\mathbf{\Gamma}_f$ at t_f .

A necessary and sufficient condition for DB to hold is the absence of irreversible fluxes $J^{irr} = 0$, in steady conditions. Then the DB condition lefts unbounded the reversible steady fluxes (with physical currents having zero ensemble averages).

The absence of irreversible fluxes imposes the thermodynamic curvature of the irreversible drift to vanish. And we obtain DB in terms only of the geometrical properties of the drift and diffusion terms.

$$D_i^{-1} \partial_j A_i^{irr}(\mathbf{\Gamma}) - D_j^{-1} \partial_i A_j^{irr}(\mathbf{\Gamma}) = 0 \quad (15)$$

D. Effective equilibrium regime AOUP

If now we apply the previous formalism to AOUP it is possible to obtain a stationary solution in an equilibrium regime. If we start from eq. 1 and introduce the velocity variable as $\mathbf{p}(t) = \dot{\mathbf{r}}(t)$, we can rewrite the vectorial dynamics as,

$$\begin{aligned} \dot{\mathbf{r}}_i(t) &= \mathbf{p}_i \\ \dot{\mathbf{p}}_i(t) &= \mu_0 \left(\sum_j \mathbf{p}_j \partial_j \right) \mathbf{F}_i - \frac{\mathbf{F}_i}{\tau} - \mu_0 \frac{\mathbf{p}_i(t)}{\tau} + \frac{\sqrt{2D}}{\tau} \boldsymbol{\eta}_i(t) \end{aligned} \quad (16)$$

where for us $\mu_0 = 1$. The coupling between velocity and space derivatives in the first term of the second equation arises from the convective part of the total derivative of the velocity. From these two equations we can derive the FP generator of the time evolution problem which yields

$$\begin{aligned} \Omega_0(\mathbf{\Gamma}) &= \sum_i \left[-\mathbf{p}_i \partial_i \right. \\ &\quad \left. - \partial_{\mathbf{p}_i} \left(\mu_0 \left(\sum_j \mathbf{p}_j \partial_j \right) \mathbf{F}_i - \frac{\mathbf{p}_i}{\tau} + \mu_0 \frac{\mathbf{F}_i}{\tau} - \frac{D}{\tau^2} \partial_{\mathbf{p}_i} \right) \right] \end{aligned} \quad (17)$$

where $\partial_{\mathbf{p}_i} \doteq \partial / \partial \mathbf{p}_i$, $\partial_i \doteq \partial / \partial \mathbf{r}_i$ and $\mathbf{\Gamma} = \{\mathbf{r}_i, \mathbf{p}_i\}$. From here, it is possible to obtain explicitly the DB condition corresponding to $\mathbf{J}^{irr} = 0$ by applying the definitions.

$$J^{irr} = \left(\mu_0 \left(\sum_j \mathbf{p}_j \partial_j \right) \mathbf{F}_i - \frac{\mathbf{p}_i}{\tau} - D \partial_{\mathbf{p}_i} \right) \Psi(\mathbf{\Gamma}, t) = 0 \quad (18)$$

where $\beta \doteq \mu_0/D$. This expression can be reduced to a first order partial differential equation,

$$\frac{1}{\tau^2} \partial_{\mathbf{p}_i} \ln \Psi_o(\mathbf{\Gamma}) = \beta \left(\sum_j \mathbf{p}_j \partial_j \right) \mathbf{F}_i - \frac{\mathbf{p}_i}{D\tau} \quad (19)$$

We can solve this differential equation exactly up to a function that depends on the space coordinates $\Lambda(\{\mathbf{r}_i\})$,

$$\Psi_0 = \exp \left[-\Lambda(\{\mathbf{r}_i\}) - \frac{\beta\tau^2}{2} \left(\sum_i \mathbf{p}_i \partial_i \right)^2 \Phi - \sum_i \frac{\tau \mathbf{p}_i^2}{2D} \right] \quad (20)$$

Finally, to solve for the arbitrary function $\Lambda(\{\mathbf{r}_i\})$ and solve the problem, we impose stationarity, and obtain

$$\Psi_{eq} = N \exp \left[-\beta\Phi - \frac{\tau}{2} \sum_i \left(\frac{\mathbf{p}_i^2}{D} + \beta^2 D |\partial_i \Phi|^2 \right) - \frac{\beta\tau^2}{2} \left(\sum_i \mathbf{p}_i \partial_i \right)^2 \Phi \right] \quad (22)$$

So, for an interaction potential with vanishing third derivatives, it is possible to obtain an exact equilibrium solution of the probability distribution. It is then possible to formulate the question for the existence of MIPS and its intrinsic non-equilibrium character. It is important to notice that the probability distribution, although it is the exponential of a function, this function is not the original interaction potential. What is obtained is an effective potential that couple's positions and velocities, and it is not a result that can be directly seen from the generator of the dynamics.

IV. NUMERICAL SIMULATIONS

To search for MIPS in a system of AOUP with (3), we will take as a control parameter for the activity the persistence time such that $\tau = D$ [14][15]. The other key control parameter in the system is the density or packing fraction.

$$\phi = \frac{\pi\sigma^2 N}{4L^2}$$

We explore a broad range of $\tau = 0.1, \dots, 100$ and $\phi = 0, \dots, 0.8$. Following the results of Figures 3 and 2 we will let the system relax a time of at least 10τ in order to fully enter the diffusive regime and consider the system has reached stationarity. The velocity of the ballistic regime remains fixed $v_0 = 1$. As we will fix the density and the number of particles simulated for different values of τ and ϕ , the size of the box, L^2 , is fixed by the definition of ϕ .

that the free function must satisfy

$$\sum_i \left[\partial_i \Lambda + \frac{\beta\tau^2}{2} \left(\sum_j \mathbf{p}_j \partial_j \right)^2 \partial_i \Phi - \frac{\beta^2 D \tau}{2} \partial_i \sum_j |\partial_j \Phi|^2 - \beta \partial_i \Phi \right] \Psi_0 = 0 \quad (21)$$

In general, there is no solution to the previous differential equation. This is because there is a term proportional to the velocity squared. But, if the third derivatives of the potential vanish, the term proportional to the velocity squared is eliminated and it is possible to solve for $\Lambda(\{\mathbf{r}_i\})$, and obtain an exact equilibrium solution for a system of AOUP, given by

And to reproduce the results of the simulations, L , must be calculated for each parametrization

The computation time for the evolution of a system of interacting particles grows as N^2 when simulated by directly computing the additive pair forces. But, to explore a wide range of parameters for big systems it would cost a large amount of computation power and time. To optimize the code, we introduce the Verlet list using cells to identify the nearest neighbors and obtain a computation time of the order of N [16].

We can see Fig.4 that as we move to higher selfpropulsion for densities above $\phi \approx 0.4$ the systems tends to create clusters of higher densities in the system. But, for a system of $N = 2000$ particles it is difficult to see a big cluster with high density surrounded by a phase of lower density, as is normally seen for MIPS. For that reason, simulations with a bigger system have been performed, to see how the system behaves under finite size effects. The results obtained in these simulation Fig.5, show that for a larger system a large phase of high density is developed, and the tendency to aggregate is preserved. Also, we can see that this big cluster is stable as the system evolves.

To understand better the nature of the dense state, to see the importance of N-body effects and the relation our results have with respect to other systems undergoing MIPS, we can explore the pair correlation function(PCF) of the system in the final distribution of the simulation with $N = 6000$.

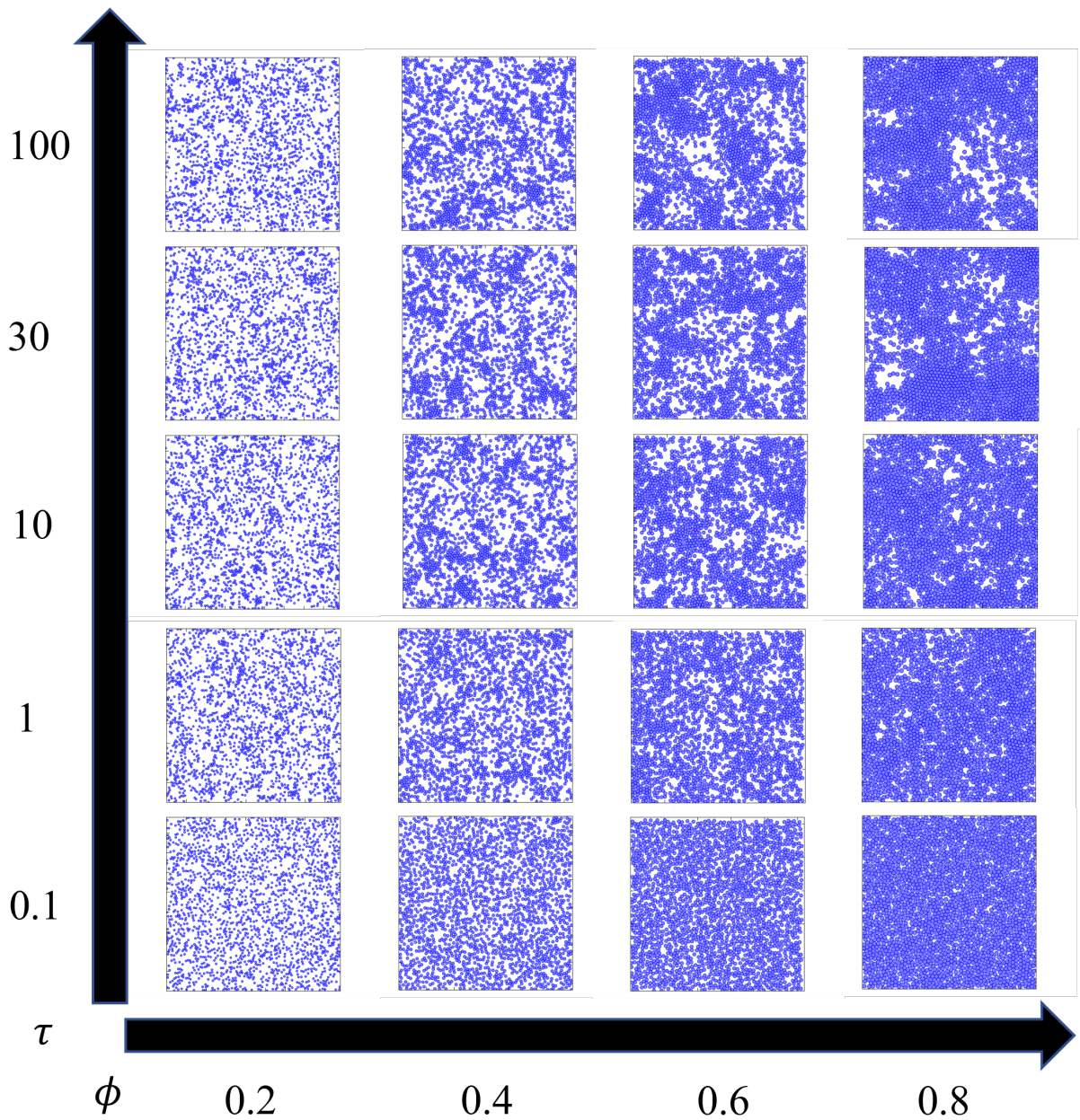


FIG. 4. Typical steady-state configuration of simulations of AOUP with parameters: $N = 2000, \sigma = 1, \text{nsteps} = 10^6, k = 10$ and values of $\tau = D$ and ϕ as indicated in the figure.

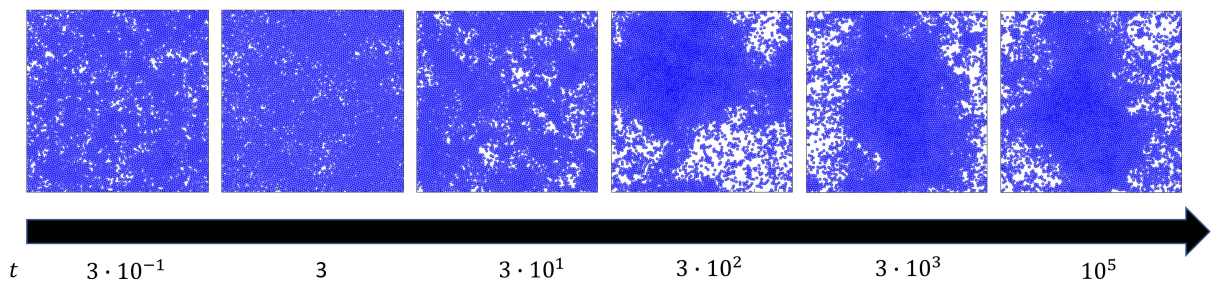


FIG. 5. Evolution of the configuration of a system of AOUP with parameters: $N = 6000, \sigma = 1, \tau = 100, \phi = 0.8, k = 10$

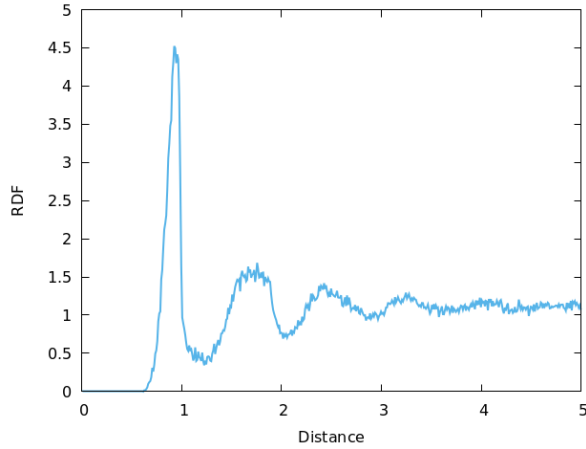


FIG. 6. Pair correlation function of the final configuration of the FIG. 5. For larger distances the value is still 1, but to better observe the peaks we have truncated the distance to 5σ .

By inspecting the PCF in FIG. 6 we observe that the first peak in the probability of observing a particle is $\sim 0.9\sigma$, and the second maximum is at $\sim 1.6\sigma$. These results manifest that in the denser phase of the system, there is overlapping between the particles. This is interesting for three reasons. First, because from purely repulsive interaction, for high enough selfpropulsion we can see a non-negligible overlap between particles. The second reason is that this characteristic can be searched when studying the 2-body problem of our system to identify if it is a 2-body effect or we need to go beyond to explain it. The third one is that this type of PCF is not

the one expected in a crystal, it resembles much more a homogeneous liquid phase.

V. MULTIBODY EFFECTS

There are different paths that can be followed to quantify the N-body effects in a system. One possibility is to measure Virial coefficients in a systematic way for the reduced 2,3...-body problem and by comparing them, isolate for example pure 3-body effects[12][17]. An alternative path, that can be used when the probability distribution is known or can be approximated, is to study the reduced problem analytically and extract the pair correlation function and effective potentials of the 2,3...-body system and compare it to simulations of those size and also to compare it with big systems. This second possibility is the one we will follow.[18].

A. 2-body problem

The first step to study the transition is to look at how a system of 2 particles behaves. This simplest case has an advantage over the rest n-body problems that we will see afterwards and is that analytically it is for some steps tractable and transparent and can be easily compared to simulations. We would want to obtain 2 results from here, an effective potential, and the pair correlation function.

The starting point is the 2-body probability distribution that can be obtained from (22)

$$\Psi_2 = \frac{1}{Z_2} \exp \left[-\beta \Phi(\mathbf{r}_1, \mathbf{r}_2) - \frac{\tau}{2} \sum_{i=1}^2 \left(\frac{\mathbf{p}_i^2}{D} + \beta^2 D |\nabla_i \Phi|^2 \right) - \frac{\beta \tau^2}{2} (\mathbf{p}_1^2 \nabla_1^2 + \mathbf{p}_2^2 \nabla_2^2 + 2\mathbf{p}_1 \mathbf{p}_2 \nabla_1 \nabla_2) \Phi(\mathbf{r}_1, \mathbf{r}_2) \right] \quad (23)$$

where the subindex references the particle. We have maintained the potential with a general dependence because it is clearer when performing the calculation. From here it would be easy following the prescription of equilibrium statistical mechanics to obtain the effective 2-body interaction, w , as

$$\Psi_2 = \frac{1}{Z_2} \exp(-\beta w(r_1, r_2, \mathbf{p}_1, \mathbf{p}_2))$$

If we fix the coordinates to be planar, we have the following generalized coordinates to describe the system $\Gamma = \{r_1, \theta_1, r_2, \theta_2, p_1^r, p_1^\theta, p_2^r, p_2^\theta\}$, where we denote the particle with subindex and when needed the coordinate with superindex. For our harmonic potential eq. 3, the terms different than zero that appear in $w(\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_1, \mathbf{p}_2)$ are $\partial_r \Phi = -(2k/\sigma)(1 - r/\sigma)$, $\partial_{r_1} \partial_{r_2} \Phi = -2k/\sigma^2$ and $\partial_r^2 \Phi = 2k/\sigma^2$. So, particularizing to our case we have

$$w(r_1, r_2, \mathbf{p}_1, \mathbf{p}_2) = \left(1 - \frac{|r_1 - r_2|}{\sigma}\right)^2 \left(k + 4 \frac{\tau k^2}{\sigma^2}\right) + \left(\frac{\tau}{2} + \frac{\tau^2 k}{\sigma^2}\right) (\mathbf{p}_1^2 + \mathbf{p}_2^2) - 2 \frac{\tau^2 k}{\sigma^2} \mathbf{p}_1 \mathbf{p}_2 \quad \text{for } |r_1 - r_2| < \sigma$$

$$w(r_1, r_2, \mathbf{p}_1, \mathbf{p}_2) = \frac{\tau}{2D\beta} (\mathbf{p}_1^2 + \mathbf{p}_2^2) \quad \text{for } |r_1 - r_2| > \sigma \quad (24)$$

which is a velocity dependent potential. Although we have followed the definitions derived from equilibrium statistical mechanics, the potential that is obtained cannot be easily used to study the effective 2-body problem as there exists a dependence in the module of the velocity and on the angle they form, which goes back to

the idea of the slow down of particles when they collide due to their persistent nature. We will have to manipulate some definitions to produce an analytical result like what we would expect when facing an equilibrium statistical physics problem, a potential depending only on the space coordinates.

B. Reduced formalism

The probability distribution of having particle 1 at \mathbf{r}_1, \dots , particle n at \mathbf{r}_n and the correlation function for a N -particle systems are respectively [19],

$$\begin{aligned} \Psi^{(n)}(\mathbf{r}_1, \dots, \mathbf{r}_n) &= \frac{1}{Z_N} \int e^{-\beta U_N} d\mathbf{r}_{n+1} \dots d\mathbf{r}_N \\ g^{(n)}(\mathbf{r}_1, \dots, \mathbf{r}_n) &= \frac{A^n N!}{N^n (N-n)!} \frac{1}{Z_N} \int e^{-\beta U_N} d\mathbf{r}_{n+1} \dots d\mathbf{r}_N \\ &= \frac{A^n N!}{N^n (N-n)!} \langle \delta(\bar{\mathbf{r}}_1 - \mathbf{r}_1) \dots \delta(\bar{\mathbf{r}}_n - \mathbf{r}_n) \rangle \end{aligned} \quad (25)$$

where in general $U_N = U_N(\mathbf{r}_1, \dots, \mathbf{r}_N)$ is the interaction potential of the complete system. As we saw, in our case the interaction depends on the velocities. We would like to define new magnitudes that maintain the same relations that these ones, as we would like to interpret the new pair correlation function, when multiplied by the density, as the probability of having a particle at distance r , and the exponent of the probability distribution as the effective potential.

If we look at the potential for N particles in the probability distribution, we know that the third derivative of Φ is 0, and so, $\partial_{\mathbf{r}_i} \partial_{\mathbf{r}_j} = C(i, j)$ where $C(i, j)$ is a constant (in the sense that there is no dependence in space or velocity) that depends on the particles we are taking the derivatives on. This makes possible to write our potential in two terms separating the dependencies in position and velocity $U_N(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1, \dots, \mathbf{p}_N) = U_1(\mathbf{r}_1, \dots, \mathbf{r}_N) + U_2(\mathbf{p}_1, \dots, \mathbf{p}_N)$. The pure velocity dependent potential can be expressed now as a sum over the sets of proximity conditions e.g. particles 1 and 2 are close and so interact and the rest of particles do not interact, then, pairs 1-2 and 2-3 interact but the rest of particles do not, and so on. We will denote it as $U_2(\mathbf{p}_1, \dots, \mathbf{p}_N) = \sum_{c_i}^M u_2^i(\mathbf{p}_1, \dots, \mathbf{p}_N)$ where c_i are the conditional sets. From here, it is now possible to define reduced magnitudes (with a bar) that will only depend on position coordinates by integrating over the velocities. What we obtain is

$$\begin{aligned} \overline{g^{(n)}}(\mathbf{r}_1, \dots, \mathbf{r}_n) &= \int g^{(n)}(\mathbf{r}_1, \dots, \mathbf{r}_n, \mathbf{p}_1, \dots, \mathbf{p}_n) d\mathbf{p}_1 \dots d\mathbf{p}_n = \\ &= \frac{A^n N!}{N^n (N-n)!} \frac{1}{Z_N} \int e^{-\beta U_2} d\mathbf{p}_1 \dots d\mathbf{p}_N \int e^{-\beta U_1} d\mathbf{r}_{n+1} \dots d\mathbf{r}_N \\ &= \frac{A^n N!}{N^n (N-n)!} \frac{1}{Z_N} \int e^{-\beta U_1} d\mathbf{r}_{n+1} \dots d\mathbf{r}_N \\ &\quad \exp \left(\ln \left(\alpha^{-1} \sum_{c_i} \int e^{-\beta \sum_{c_i}^M u_2^i} d\mathbf{p}_1 \dots d\mathbf{p}_N \right) \right) \end{aligned} \quad (26)$$

where $\alpha = \int e^{-\beta \sum_{c_i}^M u_2^i} d\mathbf{p}_1 \dots d\mathbf{p}_N$ is a constant, and we have redefined the partition function also to its reduced version as

$$\overline{Z_N} = Z_N \alpha^{-1} \quad (27)$$

here, M denotes the set for which the particles are in the dilute disposition. We take this term because if not, we would obtain a pair correlation function that at the limit of large separations does not tend to one, and so, it is not normalized when multiplying by the density and integrating. The rest of the dependencies on velocity have been integrated, and for each domain can be introduced in the exponential and becoming a new contribution to the zero-point potential. With this reduced partition function, we can obtain the new set of variables that maintain the relations

$$\begin{aligned} \overline{g^{(n)}}(\mathbf{r}_1, \dots, \mathbf{r}_n) &= \frac{A^n N!}{N^n (N-n)!} \overline{\langle \delta(\bar{\mathbf{r}}_1 - \mathbf{r}_1) \dots \delta(\bar{\mathbf{r}}_n - \mathbf{r}_n) \rangle} \\ \overline{\Psi^{(n)}}(\mathbf{r}_1, \dots, \mathbf{r}_n) &= \overline{\langle \delta(\bar{\mathbf{r}}_1 - \mathbf{r}_1) \dots \delta(\bar{\mathbf{r}}_n - \mathbf{r}_n) \rangle} \end{aligned} \quad (28)$$

C. Reduced 2-body problem

Now we would want to obtain the reduced pair correlation function in the 2 body problem and from there the effective potential. We follow the previous definitions to obtain

$$\begin{aligned} \overline{g^{(2)}}(r) &= \frac{A^2}{2} \frac{1}{\overline{Z_2}} \frac{\tau^2}{4D^2 \pi^2} e^{-\beta \left(1 - \frac{r}{\sigma}\right)^2 \left(k + 4 \frac{\tau k^2}{\sigma^2}\right)} \times \\ &2\pi \int_0^\infty dp_1 p_1 \int_0^\infty dp_2 p_2 \int_0^{2\pi} d\theta_{12} e^{-\beta \left(\frac{\tau}{2} + \frac{\tau^2 k}{\sigma^2}\right) (p_1^2 + p_2^2)} \\ &\quad \times e^{\frac{2\beta \tau^2 k}{\sigma^2} p_1 p_2 \cos(\theta_{12})} \quad \text{for } r < \sigma \\ \overline{g^{(2)}}(r) &= \frac{A^2}{2} \frac{1}{\overline{Z_2}} \quad \text{for } r > \sigma \end{aligned} \quad (29)$$

where $r = |\mathbf{r}_1 - \mathbf{r}_2|$. In this case, we have a piecewise function, and the velocity dependence is different in each region, this corresponds to the sum over domains of conditional sets. Then, we can follow exactly the previous exposition and redefine the partition function as

$$\overline{Z}_2 = Z_2 \frac{\tau^2}{4D^2\pi^2} \quad (30)$$

$$\overline{w}(r) = \left(1 - \frac{r}{\sigma}\right)^2 \left(k + 4\frac{\tau k^2}{\sigma^2}\right) - \frac{\ln}{\beta} \left(\frac{\tau^2}{2D^2\pi} \int_0^\infty dp_1 p_1 \int_0^\infty dp_2 p_2 \int_0^{2\pi} d\theta_{12} e^{-\beta\left(\frac{\tau}{2} + \frac{\tau^2 k}{\sigma^2}\right)(p_1^2 + p_2^2)}\right) \quad \text{for } r < \sigma \quad (31)$$

$$\overline{w}(r) = 0 \quad \text{for } r > \sigma$$

that at large distances as expected is 0. It is interesting to see that at the 2-body problem, what we obtain in the separation dependence is an harmonic potential with a new energy scale that has linear dependence in the activity $k' = k(1 + 4\frac{\tau k}{\sigma^2})$. Then, for the space dependence the activity makes more repulsive the potential. But the constant that appears at $r = \sigma$ and defines a discontinuity in the potential that arises from the velocity dependence has also a non-trivial contribution.

The sign of this constant cannot be obtained directly, and we need to numerically integrate the second logarithm for specific values to see the behavior that arises from it. If we look at the case $\tau = D = \beta = k = \sigma = 1$

and the reminder dependence in the velocity (integrated) is taken as $\exp(\ln())$ and absorbed in the exponential. This is a well defined pair correlation function, and from here we can obtain the effective 2-body potential from the exponential that in general is

we obtain

$$w_2(r) = 5 \left(1 - \frac{r}{\sigma}\right)^2 - 0.4 \quad \text{for } |r_1 - r_2| < \sigma \quad (32)$$

$$w_2(r) = 0 \quad \text{for } |r_1 - r_2| > \sigma$$

This negative sign in the constant creates an effective attractive regime near the discontinuity. Now we can run simulations for systems with two particles and measure the pair correlation function and the effective potential to compare it with these analytical results.

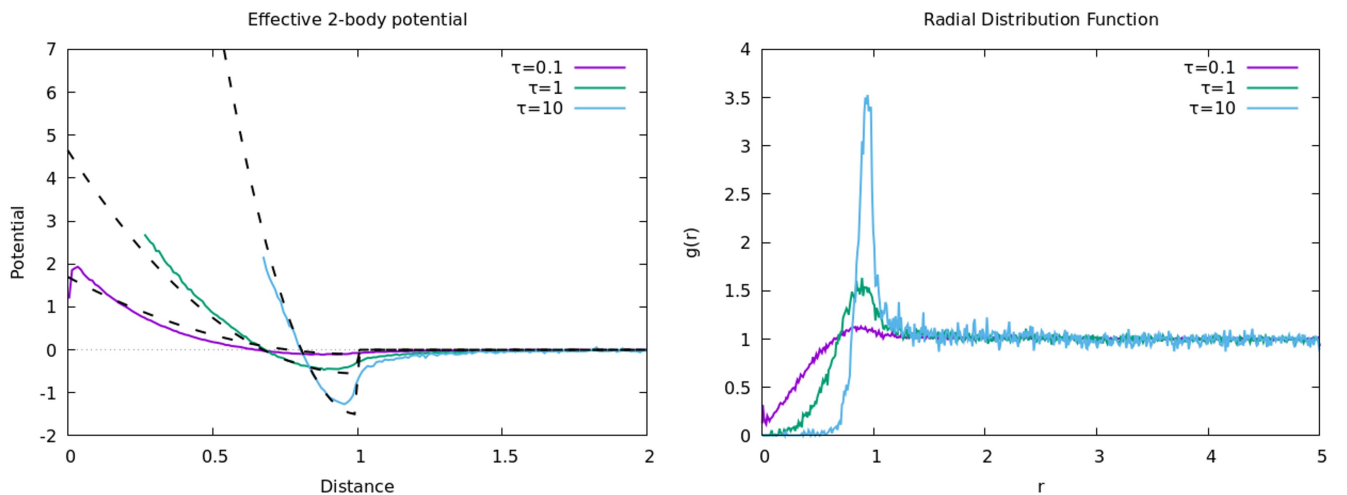


FIG. 7. Pair correlation function and Effective 2-body potential for different τ and $D = \beta = k = \sigma = 1$. The dashed lines correspond to the analytical result from (31).

We can see that, as the selfpropulsion increases the attractive well gets deeper. But at the same time the region where the attractive regime exists gets narrower. From (31) we can see that for high τ we obtain a sharper repul-

sive inner region. In relation to Fig.6 we would expect to obtain a region that permits overlap as we increase even more to $\tau = 100$, but at the 2-body level the overlap becomes more and more unlikely. We would expect peaks

that resemble not a homogeneous liquid but a crystal phase. With the results of the reduced effective potential and the reduced pair distribution function, the analysis of the 2-body problem is closed.

D. 3-body problem

To inspect the contributions that appear beyond the 2-body ones that have been studied, we will look at the restricted three body problem. To do so, we start as before from (22) to obtain the three body probability distribution in the most general form

$$\begin{aligned} \Psi_3 = & \frac{1}{Z_3} \exp \left[-\beta \left(\Phi(r_1 - r_2) + \Phi(r_2 - r_3) + \Phi(r_1 - r_3) \right) - \frac{\tau}{2D} (\mathbf{p}_1^2 + \mathbf{p}_2^2 + \mathbf{p}_3^2) \right. \\ & - \frac{\tau\beta^2 D}{2} \left(|\partial_{r_1}(\Phi(r_1 - r_2) + \Phi(r_2 - r_3) + \Phi(r_1 - r_3))|^2 + |\partial_{r_2}(\Phi(r_1 - r_2) + \Phi(r_2 - r_3) + \Phi(r_1 - r_3))|^2 \right. \\ & \left. \left. + |\partial_{r_3}(\Phi(r_1 - r_2) + \Phi(r_2 - r_3) + \Phi(r_1 - r_3))|^2 \right) - \frac{\beta\tau^2}{2} \left((\mathbf{p}_1)^2 \partial_{r_1}^2 + (\mathbf{p}_2)^2 \partial_{r_2}^2 + (\mathbf{p}_3)^2 \partial_{r_3}^2 + \right. \\ & \left. \left. 2\mathbf{p}_1\mathbf{p}_2\partial_{r_1}\partial_{r_2} + 2\mathbf{p}_2\mathbf{p}_3\partial_{r_2}\partial_{r_3} + 2\mathbf{p}_1\mathbf{p}_3\partial_{r_1}\partial_{r_3} \right) \left(\Phi(r_1 - r_2) + \Phi(r_2 - r_3) + \Phi(r_1 - r_3) \right) \right] \end{aligned} \quad (33)$$

From here, to obtain magnitudes independent on the velocity, we could apply the reduced formalism and obtain the reduced effective potential and 3-body correlation function. But the integrals over the velocity and the equations become heavy. For our purpose, that is to see what effects go beyond 2-body contributions, it is enough to obtain the potential with the velocity dependences. To

do so, we need to differentiate 4 cases corresponding to what pairs interact, this is explained in appendix A.

These four possibilities have 4 different effective potentials that have to be computed. After some calculation we can obtain the complete result for the effective 3-body potential for each region. Here $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$, $\mathbf{r}' = \mathbf{r}_2 - \mathbf{r}_3$ and the results are shown in the order of the table.

$$\begin{aligned} w_3(r, r', \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3) = & \left(k + 4\frac{\tau k^2}{\sigma^2} \right) \left(\left(1 - \frac{r}{\sigma} \right)^2 + \left(1 - \frac{r'}{\sigma} \right)^2 + \left(1 - \frac{|r+r'|}{\sigma} \right)^2 \right) \\ & + \frac{4k^2\tau}{\sigma^2} \left(\left(1 - \frac{r}{\sigma} \right) \left(1 - \frac{|r+r'|}{\sigma} \right) - \left(1 - \frac{r}{\sigma} \right) \left(1 - \frac{r'}{\sigma} \right) + \left(\left(1 - \frac{r'}{\sigma} \right) \left(1 - \frac{|r+r'|}{\sigma} \right) \right) \right) + \\ & \left(\frac{\tau}{2} + \frac{2\tau^2 k}{\sigma^2} \right) (\mathbf{p}_1^2 + \mathbf{p}_2^2 + \mathbf{p}_3^2) - \frac{2\tau^2 k}{\sigma^2} (\mathbf{p}_1\mathbf{p}_2 + \mathbf{p}_1\mathbf{p}_3 + \mathbf{p}_2\mathbf{p}_3) \end{aligned} \quad (34)$$

$$\text{for } r < \sigma \wedge r' < \sigma \wedge (r+r') < \sigma.$$

$$\begin{aligned} w_3(r, r', \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3) = & \left(k + 4\frac{\tau^2}{\sigma^2} \right) \times \left(\left(1 - \frac{r}{\sigma} \right)^2 + \left(1 - \frac{r'}{\sigma} \right)^2 \right) - \frac{4k^2\tau}{\sigma^2} \left(\left(1 - \frac{r}{\sigma} \right) \left(1 - \frac{r'}{\sigma} \right) \right) \\ & + \left(\frac{\tau}{2} + \frac{\tau^2 k}{\sigma^2} \right) (\mathbf{p}_1^2 + \mathbf{p}_3^2) + \left(\frac{\tau}{2} + \frac{2\tau^2 k}{\sigma^2} \right) \mathbf{p}_2^2 - \frac{2\tau^2 k}{\sigma^2} (\mathbf{p}_1\mathbf{p}_2 + \mathbf{p}_2\mathbf{p}_3) \end{aligned} \quad (35)$$

$$\text{for } r < \sigma \wedge r' < \sigma \wedge (r+r') > \sigma.$$

$$w_3(r, r', \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3) = \left(1 - \frac{r}{\sigma} \right)^2 \left(k + 4\frac{\tau^2}{\sigma^2} \right) + \frac{\tau}{2} \mathbf{p}_3^2 + \left(\frac{\tau}{2} + \frac{\tau^2 k}{\sigma^2} \right) (\mathbf{p}_1^2 + \mathbf{p}_2^2) - 2\frac{\tau^2 k}{\sigma^2} \mathbf{p}_1\mathbf{p}_2 \quad (36)$$

$$\text{for } r < \sigma \wedge r' > \sigma \wedge (r+r') > \sigma.$$

$$w_3(r, r', \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3) = 0 \quad (37)$$

$$\text{for } r > \sigma \wedge r' > \sigma \wedge (r+r') > \sigma.$$

If we look at the first case, we obtain once more a term equal to the sum of harmonic (original) potentials

but with a rescaled energy scale and a velocity dependent term. But a new term that couples the interactions between pairs is founded. This arises from the nonlinear term in the probability distribution. We have a square of the derivative of the potential. When this potential has more than one term depending on a particle position, a cross term arises that cannot be seen in the 2-body case. This is even more clear when comparing the second and third case of the potential. In the third case, up to the term of \mathbf{p}_3 , we obtain the 2-body potential, as it corresponds to an interacting pair and a particle far from this pair, it is a separable problem. But in the second case we have the pairs 1-2 and 2-3. A physical disposition of this case would be to have them in a line as 1-2-3, where the distance between 1 and 3 is bigger than σ . Here we do not have just the harmonic term but also a cross term that cannot be obtain without the nonlinear contribution. This is exactly the part that distinguish between the 2-body and the 3-body problem.

From here we could go on performing the same path to look at contribution above the 3-body ones. But knowing that there is a nonlinear contribution, and with the

results obtained, nontrivial effects will arise when looking at the effective n-body potential.

VI. N-BODY PARTITION FUNCTION

As we are in thermodynamic equilibrium, in principle, if we can calculate the partition function, we can apply equilibrium statistical mechanics and obtain the thermodynamics of the interacting active system. Of course, to solve the complete problem analytically is most likely impossible. But it is possible to obtain a methodology to consider the N-body terms systematically in the partition function. We know that in (22) (the probability distribution of N AOUP), the normalization constant that appears, is the desired partition function. So, if we integrate over the phase space, as the probability is normalized to 1, we have the path to obtain the N-body partition function, and formally the thermodynamics of the system. Then the partition function can be computed as

$$Z_N = \frac{1}{h^{2N} N!} \int d\mathbf{r}^N \int d\mathbf{p}^N \exp \left[-\beta\Phi - \frac{\tau}{2} \sum_i \left(\frac{\mathbf{p}_i^2}{D} + \beta^2 D |\partial_i \Phi|^2 \right) - \frac{\beta\tau^2}{2} \left(\sum_i \mathbf{p}_i \partial_i \right)^2 \Phi \right] \quad (38)$$

Where we have included the term for indistinguishable particles and the elemental volume. To perform the multibody expansion, it is very useful and meaningful to rewrite the potential as

$$\begin{aligned} \Phi(r_{ij}) &= \sum_i \sum_{j < i} \epsilon_{ij} \phi(r_{ij}) \\ \epsilon_{ij} &= \begin{cases} 1 & \text{if } r_{ij} < \sigma \\ 0 & \text{if } r_{ij} > \sigma \end{cases} \end{aligned} \quad (39)$$

The ϵ_{ij} matrix, can be understood with a complete analogy to the adjacency matrix in network theory. It expresses the connections(interactions) established between the nodes(particles). From the definition we can obtain two properties, first, that the diagonal elements are zero as the particle do not interact with themselves. The second one is that it is a symmetric matrix, if i interacts with j , j interacts with i . Specific configurations of the matrix correspond to different conditional sets as they were defined in the reduced functions exposition. Also, the density of non-zero values, express the local density in the system.

A. Dilute gas

As a first approach to the problem, we could think in a limit case. The easiest one is the "completely" dilute

gas. The density is low enough that particles do not interact between them. For low enough densities we can suppose that $\sigma \rightarrow 0$ and obtain an ideal gas of N AOUP (N-IG-AOUP) $\Phi = 0$

$$\begin{aligned} Z_N &= \frac{1}{h^{2N} N!} \int d\mathbf{r}^N \int d\mathbf{p}^N e^{-\frac{\tau}{2} \sum_{i=1}^N \frac{\mathbf{p}_i^2}{D}} \approx \\ &= \frac{A^N}{h^{2N} N!} \left[\int d\mathbf{p} e^{-\frac{\tau \mathbf{p}^2}{2D}} \right]^N = \frac{1}{N!} \left[\frac{2\pi AD}{h^2 \tau} \right]^N \end{aligned} \quad (40)$$

B. Multibody expansion

Now, we start to attack the problem for non-vanishing potential. In general, the equilibrium condition tells us $\partial_i^3 \phi = 0 \rightarrow \partial_i^2 \phi = \bar{k} \rightarrow \partial_i \phi = \bar{k}r + \alpha$, and for any configuration, all the possible derivatives that can be used in the calculations are

$$\begin{aligned} \partial_i \Phi &= \sum_{j=1}^N \epsilon_{ij} (\bar{k}r_{ij} + \alpha) \\ \partial_i^2 \Phi &= \bar{k} \sum_{j=1}^N \epsilon_{ij} \\ \partial_i \partial_j \Phi &= -\bar{k} \epsilon_{ij} \end{aligned} \quad (41)$$

We take here this form of the potential to obtain the result for the general equilibrium with short range interaction and not our specific potential. With these results, we can separate the problem into two independent problems. First an integration over velocities and a second one over positions as was shown in the reduced formalism. But before we considered the short-range charac-

teristic by separating in a sum over conditional sets. In this case, this is already considered in the adjacency matrix and no decomposition of the Hamiltonian is needed, instead a sum over the constructable matrices can be taken.

Using the previous results, the solution to the velocities integrals is

$$\mathcal{Z} = \sum_{\{\epsilon_{ij}\}} \prod_i \int \exp\left(-\frac{\beta\tau}{2}\left(\mathbf{p}_i^2 + \tau\mathbf{p}_i^2\bar{k} \sum_j \epsilon_{ij} - 2\tau\bar{k} \sum_{j<i} \mathbf{p}_i\mathbf{p}_j\epsilon_{ij}\right)\right) d\mathbf{p}_1\dots d\mathbf{p}_N \quad (42)$$

That has an analytical solution only for $\epsilon_{ij} = 0 \forall i, j$, that is the dilute gas. For all the other cases, the cross term makes the function not integrable. So, not much more can be done for the velocities. We have then, that the partition function is

$$Z_N = \frac{\mathcal{Z}}{h^{2N}N!} \int d\mathbf{r}^N \exp\left[-\beta\Phi - \frac{\tau}{2} \sum_i \beta^2 D |\partial_i \Phi|^2\right] \quad (43)$$

Now we can make the following change of variables $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$, and integrate over the coordinates of one particle. We start with N coordinates $\{\mathbf{r}_i\}_{i=1}^N$ and end up with $N-1$, $\{\mathbf{r}_{i,i+1}\}_{i=1}^{N-1}$. All the other vectors can be computed by differences of the $N-1$ ones as $\mathbf{r}_{ij} = \sum_{k=i}^{j-1} \mathbf{r}_{k,k+1}$, for $\mathbf{r}_{ii} = 0 \forall i$ and $\mathbf{r}_{ij} = 0$ if $i, j < 1$. There is an explanation for this base and not other possibilities such as $\{\mathbf{r}_{1j}\}_{j=2}^N$ with $\mathbf{r}_{ij} = \mathbf{r}_{1j} - r_{1i}$, but can only be seen at a later stage in the calculations. By the properties previously exposed and $\Phi = \sum_{i<j} \dots = \frac{1}{2} \sum_{i,j} \dots$, we can rewrite

$$Z_N = \frac{\mathcal{Z}A}{h^{2N}N!} \prod_{i=1}^{N-1} \int d\mathbf{r}_{i,i+1} \exp\left\{\frac{-\beta}{2} \left[\sum_j \epsilon_{ij} \phi(r_{ij}) + \frac{\tau\beta D}{2} \sum_{j,l} \epsilon_{ik} \epsilon_{il} (\bar{k}r_{ij} + \alpha)(\bar{k}r_{il} + \alpha) \right]\right\} \quad (44)$$

where $\epsilon_{ik} \epsilon_{il} = \delta_{\epsilon_{ik} + \epsilon_{il}, 2}$. Here the integrals go over

the separations between the N particles. This is not just interesting because we have integrated one coordinate, but because the integrals can be separated and directly correlated with the ϵ matrix and a multibody expansion.

Of course, the lowest density configuration, and so, the first term of the sum is when all particles are disconnected, the N-IG-AOUP region. Now we won't take the limit $\sigma \rightarrow 0$. We have for $\epsilon_{ij} = 0 \forall i, j$

$$Z_N^1 = \frac{\mathcal{Z}_N^1 A}{h^{2N}N!} \prod_{i=1}^{N-1} \int_{r_{i,i+1} \in [\sigma, L]} d\mathbf{r}_{i,i+1} \quad (45)$$

$$= \frac{\mathcal{Z}}{h^{2N}N!} A (A - \pi\sigma^2)^{N-1}$$

We have obtained the partition function of an ideal gas with excluded volume. From here on, the calculations become less transparent, although the physical intuitions are clear. Now we need to allow interactions between particles. The first idea is of course to relax the condition on the separation for a pair of particles, that without losing generality will be $1-2$. This problem now becomes separable in the sense that the other $N-2$ particles have decoupled dynamics, and $1-2$ are governed by a 2-body effective interaction. This is clear when looking at ϵ_{ij} , there is only 1 position different than 0, and so it cannot couple with anybody. We can create $N(N-1)/2$ pairs. The contribution then becomes,

$$\overline{Z}_N^2 = \overline{Z}_N^2 \frac{N(N-1)}{2N!h^{2N}} A (A - \pi\sigma^2)^{N-2} \int_0^\sigma dr_{12} r_{12} \exp\left\{-\frac{\beta}{2} \left[\bar{k} (1 + \bar{k}\tau) r_{12}^2 + (\alpha + \tau\beta\bar{k}D\alpha)r_{12} + \left(\gamma + \frac{\tau\alpha^2}{2}\right) \right]\right\} \quad (46)$$

where $\gamma = \phi(0)$ is the zero point energy of the potential. The next contribution will be to have two separable pairs of particles that interact through the effective two body potential. We will have an ideal AOUP gas of $N-4$ particles and two 2-body problems separable. After we can have

$3, 4, \dots, N(N-1)/2$ pairs. The degeneracy of each contribution is different. For the first one it was the number of pairs possible. The second will be the numbers of pairs times the number of pairs possible with 2 particles less, and so on. In general,

$$\mathcal{D}_k = \prod_{i=1}^k \frac{(N-2(i-1))(N-1-2(i-1))}{2} \quad (47)$$

To obtain the contribution that combines all the pairs with 2-body interactions we will take the pairs to be 1-2, 3-4, 5-6... $(N-2) - (N-1)$ for N odd. Now the choice of the base of $N-1$ variables is evident. The pair particles distances are the coordinates to integrate, and so separable. Which are the variables to integrate with a

$$Z_N = \frac{\mathcal{Z}_N^1 A}{h^{2N} N!} (A - \pi\sigma^2)^{N-1} + \sum_{n=1}^{\frac{N-1}{2}} \frac{\mathcal{Z}_n^2 \mathcal{D}_n A}{h^{2N} N!} (A - \pi\sigma^2)^{N-2-(n-1)} \left(\int_0^\sigma dr_{12} r_{12} \exp \left\{ -\frac{\beta}{2} \left[\frac{\bar{k}}{2} (1 + \bar{k}\tau) r_{12}^2 + (\alpha + \tau\bar{k}\alpha) r_{12} + \left(\gamma + \frac{\tau\alpha^2}{2} \right) \right] \right\} \right)^n + (3, \dots, N) - \text{body interactions} \quad (48)$$

where the subindex in the configurational integrals express that the adjacency matrix has a particular form for each configuration. Why haven't we considered double pairs as (1-2) and (2-3)? Because the effective interaction obtained for three bodies, explicitly shows that this kind of pseudo-triplets couple the interaction, and 3-body non-trivial effects arise as was discussed in previous sections. And so, the integrals won't be separable as the interaction would have 3-body contributions. In terms of the adjacency matrix, for the upper triangle, we would have one column with a one and the next one without ones. How can we proceed in a systematic manner to go beyond this result? The first thing to do is to include 3-body effects. We would have N-3 ideal gas of AOUP, and then 1-2-3 could be in two states that have contributions beyond 2-body. First, interactions could be between 1-2,1-3 and 2-3, and so the general 3-body effective potential should be considered that is equation

$$Z_N = Z_N^1 + Z_N^2(Z_N^1) + Z_N^3(Z_N^1, Z_N^2) + \dots$$

$$Z_N^1 = \frac{\mathcal{Z}_N^1 A}{h^{2N} N!} (A - \pi\sigma^2)^{N-1}$$

$$Z_N^2 = \sum_{n=1}^{\frac{N-1}{2}} \frac{\mathcal{Z}_n^2 \mathcal{D}_n A}{h^{2N} N!} (A - \pi\sigma^2)^{N-2-(n-1)} \left(\int_0^\sigma dr_{12} r_{12} \exp \left\{ -\frac{\beta}{2} \left[\frac{\bar{k}}{2} (1 + \bar{k}\tau) r_{12}^2 + (\alpha + \tau\bar{k}\alpha) r_{12} + \left(\gamma + \frac{\tau\alpha^2}{2} \right) \right] \right\} \right)^n \quad (49)$$

This multibody expansion resembles a well known result in equilibrium statistical physics, the Virial expansion. The Virial term of order n has contributions of $n, n-1, n-2 \dots$ 1-body, and the relation between the n term

potential and which not is summarized in Table 2 in the appendix.

As N is odd we integrate the same number of interacting coordinates as free ones, $(N-1)/2 = k$, $k \in \mathbb{R}$. As the pairs are separable, all of them are equivalent, and we can replace them by r_{12} . Finally, we can recall all the results concerning two body interaction in the N-body problem and obtain the partition function up to 2-body interactions.

(34). But they could also be in a disposition such that 1-2 and 2-3 will interact and 1-3 no. Then a partial 3-body effective potential should be used (35). Starting from the second case, we will have the N-3 ideal gas of AOUP plus 1 pseudo triplet. Then, we would vary the ideal gas by introducing separable pairs as with the 2-body term until the maximum possible. Then we will have 2 pseudo-triplets and play with the ideal gas and the pairs and so on. After, we should take the complete triplet and play with pairs and pseudo-triplets, then add a new pure triplet and play once more with all the possible pairs and pseudo-triplets, and so on. Every new term is the previous one (varying the maximum number of N-body clusters) and an additional separable (N+1)-body term recursively. In principle this could be carried on obtaining formally the exact N-body partition function. This result is a multibody expansion for the partition function

and $n-1$ term gives the pure n-body contribution.

VII. CONCLUSIONS

We started this project with the aim of unravelling the nonequilibrium nature of MIPS. Recovering the result that AOUP that interact with a harmonic potential fulfill detailed balance[10][11] gave the possibility to see if MIPS could be seen in a system in equilibrium. First, it was obtained that for densities between 0.4 and 0.8, the system clusterize for persistence time around and above 10 FIG.4. Regions of high density and low motility are observed. When looking at the effects of finite size in the simulations, the results where robust, and for a system of $N = 6000$ particles and deep in the phase space ($\phi = 0.8, \tau = 100$) a large and stable high-density phase is formed FIG.5. Looking at the nature of the dense state, it was obtained that overlap between particles is important. The radial distribution function for stationary configurations of our simulations showed peaks around 0.9σ and $1,6\sigma$, as opposed to the ones that appear in the hexatic order founded in MIPS of active Brownian disks [20]. What we found resembles much more a homogeneous liquid state.

Knowing that MIPS can appear in thermodynamic equilibrium, we could study the multibody effects that appear in the system departing from the equilibrium probability distribution that defined an effective potential. We wanted to discern if the transition could be explained by effective 2-body interactions or as previous research showed[12], non-trivial n-body effects are needed to account for the transition. The results obtained by performing 2-body simulations show that as activity is increased the inner region of the particles becomes more impenetrable. But also, an attractive well is developed near the edge of the particles. The width of the attractive regime decreases with the increase of the persistence time, and so we would expect that overlap for high activity in the 2-body problem would be unlikely. Also, as the equilibrium probability distribution was known, the reduced effective potential and radial distribution function could be defined and compared with the simulations. The analytical calculations unveiled that the coupling between velocity and position(potential) was the origin of a discontinuity in σ that created the attractive regime. When looking at the three-body problem, the non-linear contribution to the effective potential created

3-body terms unseen in the 2-body problem. These new terms coupled the interactions between different pairs if there was one particle in common. This new contribution acts on the inner region and could be a possible explanation of the overlap.

Finally, as the system is strictly in thermodynamic equilibrium as detailed balance is fulfilled and the entropy production is 0, the calculation of the partition function should lead to a complete characterization of the thermodynamics of the system. A recursive method was developed to calculate the partition function. This was possible because the effective 1,2,3.. potentials can be analytically obtained and separated. With the help of the adjacency matrix, we have been able to express Z_N as a sum over terms that account for n-body effects hierarchically, a multibody expansion. Up to 2-body contributions the partition function was explicitly obtained.

These results open more interesting questions concerning this system to continue the research. First, to close down the importance of the 2-body effects in the transition, a simulation of Brownian particles that interact with the 2-body effective potential in a regime where the active system phase separates could be done. First, we could see if the mapped system phase separates, and then if the overlap exists. Also, we could characterize the phase transition by systematically simulating the system along the phase space and looking at local densities. Finally, if a wall was included, by means of the generalized virial theorem, the pressure on the wall could be obtained, and numerically integrating the partition function up to the n-body contribution could be performed to obtain the analytical pressure and the equation of state.

ACKNOWLEDGMENTS

I would like to thank Sara Dal Cengio and Matteo Paoluzzi for their kindness in helping me in the development of this project. To my advisor, Demian Levis, for his time, dedication, patience, and guidance that I have had unconditionally throughout these months. To my family, for their heroic love, every day.

-
- [1] M. C. Marchetti, J. F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao, and R. A. Simha, "Hydrodynamics of soft active matter," *Rev. Mod. Phys.*, vol. 85, pp. 1143–1189, Jul 2013. [Online]. Available: <https://link.aps.org/doi/10.1103/RevModPhys.85.1143>
 - [2] C. Bechinger, R. D. Leonardo, H. Löwen, C. Reichardt, G. Volpe, and G. Volpe, "Active particles in complex and crowded environments," *Reviews of Modern Physics*, vol. 88, no. 4, nov 2016. [Online]. Available: <https://doi.org/10.1103%2Frevmodphys.88.045006>
 - [3] J. O’Byrne, Y. Kafri, J. Tailleur, and F. van Wijland, "Time irreversibility in active matter, from micro to macro," *Nature Reviews Physics*, vol. 4, no. 3, pp. 167–183, Jan. 2022.
 - [4] G. Szamel, "Self-propelled particle in an external potential: Existence of an effective temperature," *Physical Review E*, vol. 90, no. 1, jul 2014. [Online]. Available: <https://doi.org/10.1103%2Fphysreve.90.012111>
 - [5] A. P. Solon, M. E. Cates, and J. Tailleur, "Active brownian particles and run-and-tumble particles: A

- comparative study,” *The European Physical Journal Special Topics*, vol. 224, no. 7, pp. 1231–1262, jul 2015. [Online]. Available: <https://doi.org/10.1140%2Fepjst%2F2015-02457-0>
- [6] D. Martin, J. O’Byrne, M. E. Cates, É. Fodor, C. Nardini, J. Tailleur, and F. van Wijland, “Statistical mechanics of active ornstein-uhlenbeck particles,” *Physical Review E*, vol. 103, no. 3, mar 2021. [Online]. Available: <https://doi.org/10.1103%2Fphysreve.103.032607>
- [7] M. E. Cates and J. Tailleur, “Motility-induced phase separation,” *Annual Review of Condensed Matter Physics*, vol. 6, no. 1, pp. 219–244, mar 2015. [Online]. Available: <https://doi.org/10.1146%2Fannurev-conmatphys-031214-014710>
- [8] Y. Fily and M. C. Marchetti, “Athermal phase separation of self-propelled particles with no alignment,” *Physical Review Letters*, vol. 108, no. 23, jun 2012. [Online]. Available: <https://doi.org/10.1103%2Fphysrevlett.108.235702>
- [9] É. Fodor, C. Nardini, M. E. Cates, J. Tailleur, P. Visco, and F. van Wijland, “How far from equilibrium is active matter?” *Physical Review Letters*, vol. 117, no. 3, jul 2016. [Online]. Available: <https://doi.org/10.1103%2Fphysrevlett.117.038103>
- [10] L. L. Bonilla, “Active ornstein-uhlenbeck particles,” *Physical Review E*, vol. 100, no. 2, aug 2019. [Online]. Available: <https://doi.org/10.1103%2Fphysreve.100.022601>
- [11] S. D. Cengio, D. Levis, and I. Pagonabarraga, “Fluctuation–dissipation relations in the absence of detailed balance: formalism and applications to active matter,” *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2021, no. 4, p. 043201, apr 2021. [Online]. Available: <https://doi.org/10.1088/1742-5468/abee22>
- [12] F. Turci and N. B. Wilding, “Phase separation and multibody effects in three-dimensional active brownian particles,” *Phys. Rev. Lett.*, vol. 126, p. 038002, Jan 2021. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.126.038002>
- [13] M. E. Tuckerman, *Statistical mechanics: Theory and molecular simulation*. Oxford University Press, 2018.
- [14] R. G. Winkler, A. Wysocki, and G. Gompper, “Virial pressure in systems of spherical active brownian particles,” *Soft Matter*, vol. 11, pp. 6680–6691, 2015. [Online]. Available: <http://dx.doi.org/10.1039/C5SM01412C>
- [15] M. Paoluzzi, C. Maggi, and A. Crisanti, “Statistical field theory and effective action method for scalar active matter,” *Phys. Rev. Research*, vol. 2, p. 023207, May 2020. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevResearch.2.023207>
- [16] M. P. Allen and D. J. Tildesley, *Computer simulation of liquids*. Oxford University Press, 2017.
- [17] D. Ashton and N. Wilding, “Three-body interactions in complex fluids: Virial coefficients from simulation finite-size effects,” *The Journal of chemical physics*, vol. 140, 03 2014.
- [18] U. M. B. Marconi, M. Paoluzzi, and C. Maggi, “Effective potential method for active particles,” *Molecular Physics*, vol. 114, no. 16-17, pp. 2400–2410, 2016. [Online]. Available: <https://doi.org/10.1080/00268976.2016.1155777>
- [19] D. A. McQuarrie, *Statistical Mechanics*, ser. Harper’s chemistry series. New York: Harper Collins, 1976.
- [20] P. Digregorio, D. Levis, A. Suma, L. F. Cugliandolo, G. Gonnella, and I. Pagonabarraga, “Full phase diagram of active brownian disks: From melting to motility-induced phase separation,” *Physical Review Letters*, vol. 121, no. 9, aug 2018. [Online]. Available: <https://doi.org/10.1103%2Fphysrevlett.121.098003>

Appendix: 3-body contributions

Region	Degeneracy
$ \mathbf{r}_1 - \mathbf{r}_2 < \sigma \wedge \mathbf{r}_2 - \mathbf{r}_3 < \sigma \wedge \mathbf{r}_1 - \mathbf{r}_3 < \sigma$	1
$ \mathbf{r}_1 - \mathbf{r}_2 < \sigma \wedge \mathbf{r}_2 - \mathbf{r}_3 < \sigma \wedge \mathbf{r}_1 - \mathbf{r}_3 > \sigma$	3
$ \mathbf{r}_1 - \mathbf{r}_2 < \sigma \wedge \mathbf{r}_2 - \mathbf{r}_3 > \sigma \wedge \mathbf{r}_1 - \mathbf{r}_3 > \sigma$	3
$ \mathbf{r}_1 - \mathbf{r}_2 > \sigma \wedge \mathbf{r}_2 - \mathbf{r}_3 > \sigma \wedge \mathbf{r}_1 - \mathbf{r}_3 > \sigma$	1

Table 1. Connectivity regions for the interacting particles. As the particles are indistinguishable, when one of the three conditions in the first case is broken we have three possibilities to obtain an equivalent system. That creates a multiplicity on the sets of 3 for the second and third case.

For these 4 cases the interaction potential that we have are respectively

$$\begin{aligned}
 \frac{\Phi_1(r_{12}, r_{13}, r_{23})}{k} &= \left(1 - \frac{r_{12}}{\sigma}\right)^2 + \left(1 - \frac{r_{13}}{\sigma}\right)^2 + \left(1 - \frac{r_{23}}{\sigma}\right)^2 \\
 \frac{\Phi_1(r_{12}, r_{23})}{k} &= \left(1 - \frac{r_{12}}{\sigma}\right)^2 + \left(1 - \frac{r_{23}}{\sigma}\right)^2 \\
 \frac{\Phi_1(r_{12})}{k} &= \left(1 - \frac{r_{12}}{\sigma}\right)^2 \\
 \frac{\Phi_1}{k} &= 0
 \end{aligned} \tag{A.1}$$

Appendix: Multibody expansion

In this appendix we detailed the calculations that were made from (44) to obtain multibody expansion of the partition function. The first term in the expansion corresponding to the dilute system yields

$$\begin{aligned}
 Z_N^1 &= \frac{\mathcal{Z}A}{h^{2N}N!} \prod_{i=1}^{N-1} \int_{r_{i,i+1} \in [\sigma, L]} dr_{i,i+1} = \\
 &= \frac{\mathcal{Z}A}{h^{2N}N!} \prod_{i=1}^{N-1} 2\pi \int_{\sigma}^L dr_{i,i+1} r_{i,i+1} = \\
 &= \frac{\mathcal{Z}A}{h^{2N}N!} \prod_{i=1}^{N-1} 2\pi \left(\int_0^L dr_{i,i+1} r_{i,i+1} - \int_0^{\sigma} dr_{i,i+1} r_{i,i+1} \right) \\
 &= \frac{\mathcal{Z}A}{h^{2N}N!} \prod_{i=1}^{N-1} 2(A - \pi\sigma^2) = \frac{\mathcal{Z}}{hr^{2N}N!} A(A - \pi\sigma^2)^{N-1}
 \end{aligned} \tag{A.1}$$

Then the complete 2-body contributions, equation (48), is presented as a sum over separable pairs plus an IG-AOUP. The steps to the first term, equation (46) and

the last one omitted from the body of the project can be reconstructed from the previous equation and the following intermediate results,

$$\begin{aligned} \overline{Z}_N^2 &= \mathcal{Z} \frac{N(N-1)}{2N!h^{2N}} A \prod_{i=2}^{N-1} \int_{r \in [\sigma, L]} d\mathbf{r}_{i,i+1} \int_{r \in [0, \sigma]} d\mathbf{r}_{12} \exp \left\{ -\frac{\beta}{2} \left(\phi(r_{12}) + \frac{\tau}{2} (\bar{k}r_{12} + \alpha)^2 \right) \right\} = \mathcal{Z} \frac{N(N-1)}{2N!h^{2N}} \times \\ &A (A - \pi\sigma^2)^{N-2} \int_0^\sigma dr_{12} r_{12} \exp \left\{ -\frac{\beta}{2} \left[\frac{\bar{k}}{2} (1 + \bar{k}\tau) r_{12}^2 + (\alpha + \tau\bar{k}\alpha) r_{12} + \left(\gamma + \frac{\tau\alpha^2}{2} \right) \right] \right\} \end{aligned} \quad (\text{A.2})$$

$$\begin{aligned} \overline{\overline{Z}}_N^2 &= \frac{\mathcal{Z} \mathcal{D}_{\frac{N-1}{2}} A}{h^{2N} N!} \left(\prod_{i=1}^{N-1} \int_{r \in [\sigma, L]} d\mathbf{r}_{2i, 2i+1} \right) \left(\prod_{i=1}^{N-1} \int_{r \in [0, \sigma]} d\mathbf{r}_{2i-1, 2i} \exp \left\{ -\beta \phi(r_{2i-1, 2i}) - \frac{\tau\beta^2 D}{2} |\partial_{r_{2i-1, 2i}} \phi(r_{2i-1, 2i})|^2 \right\} \right) \\ &= \frac{\mathcal{Z} \mathcal{D}_{\frac{N-1}{2}} A}{h^{2N} N!} (A - \pi\sigma^2)^{\frac{N-1}{2}} \left(\prod_{i=1}^{N-1} 2\pi \int_0^\sigma dr_{2i-1, 2i} r_{2i-1, 2i} \exp \left\{ -\frac{\beta}{2} \left[\frac{\bar{k}}{2} (1 + \bar{k}\tau) r_{2i-1, 2i}^2 + (\alpha + \tau\bar{k}\alpha) \times \right. \right. \right. \\ &\left. \left. \left. r_{2i-1, 2i} + \left(\gamma + \frac{\tau\alpha^2}{2} \right) \right] \right\} \right) \end{aligned} \quad (\text{A.3})$$

As the pairs are separable, all of them are equivalent, and

we can replace them by r_{12} that is a not dependent on the productory, and we obtain

$$\overline{\overline{Z}}_N^2 = \frac{\mathcal{Z} \mathcal{D}_{\frac{N-1}{2}} A}{h^{2N} N!} (A - \pi\sigma^2)^{\frac{N-1}{2}} \left(\int_0^\sigma dr_{12} r_{12} \exp \left\{ -\frac{\beta}{2} \left[\frac{\bar{k}}{2} (1 + \bar{k}\tau) r_{12}^2 + (\alpha + \tau\bar{k}\alpha) \times r_{12} + \left(\gamma + \frac{\tau\alpha^2}{2} \right) \right] \right\} \right)^{\frac{N-1}{2}} \quad (\text{A.4})$$

Finally, a table that expresses which variables in this last contribution of the sum are integrated and which not is added.

Pair	Interaction	Integration with potential
r_{12}	Yes	Yes
r_{23}	No	No
r_{34}	Yes	Yes
r_{45}	No	No
...
$r_{N-2, N-1}$	Yes	Yes
$r_{N-1, N}$	No	No

Table 2. Variables to integrate in the configurational partition function for it to be separable. The variables represent the separation between the particles corresponding to the index.