Active dumbbells: Dynamics and morphology in the coexisting region

Isabella Petrelli\textsuperscript{1}, Pasquale Digregorio\textsuperscript{1}, Leticia F. Cugliandolo\textsuperscript{2}, Giuseppe Gonnella\textsuperscript{1}, and Antonio Suma\textsuperscript{3,4,}\textsuperscript{a}

\textsuperscript{1} Dipartimento di Fisica, Università degli Studi di Bari and INFN, Sezione di Bari, via Amendola 173, Bari, I-70126, Italy
\textsuperscript{2} Sorbonne Université, Laboratoire de Physique Théorique et Hautes Energies, CNRS UMR 7589, 4 Place Jussieu, 75252 Paris Cedex 05, France
\textsuperscript{3} SISSA - Scuola Internazionale Superiore di Studi Avanzati, Via Bonomea 265, 34136 Trieste, Italy
\textsuperscript{4} Institute for Computational Molecular Science, College of Science and Technology, Temple University, Philadelphia, PA 19122, USA

Received 30 April 2018 and Received in final form 13 September 2018
Published online: 25 October 2018
© EDP Sciences / Società Italiana di Fisica / Springer-Verlag GmbH Germany, part of Springer Nature, 2018

Abstract. With the help of molecular dynamics simulations we study an ensemble of active dumbbells in purely repulsive interaction. We derive the phase diagram in the density-activity plane and we characterise the various phases with liquid, hexatic and solid character. The analysis of the structural and dynamical properties, such as enstrophy, mean-square displacement, polarisation, and correlation functions, shows the continuous character of liquid and hexatic phases in the coexisting region when the activity is increased starting from the passive limit.

1 Introduction

We open this article with a description of a number of noticeable facts of active, but also passive, matter that have motivated our studies of an ensemble of self-propelled dumbbells in purely repulsive interaction confined to move in a two-dimensional space.

1.1 Non-equilibrium dynamics under local bulk energy injection

Active materials are many-body systems composed of self-driven units that convert stored or ambient free energy into systematic movement. They are, typically, living systems, and the size of their elements ranges over many scales, from microorganisms or cells to birds or fish. Artificial realisations, sometimes easier to control in experiments, have also been designed and include self-propelled colloids [1–4], nanorods [5], droplets [6,7] and active gels made by cytoskeleton extracts in the presence of molecular motors [8,9] as well as vibrated mechanical walkers [10–13].

\textsuperscript{a} e-mail: antonio.suma@gmail.com

From the point of view of physics, active materials are novel and very interesting objects of study. They are maintained out of equilibrium by the continuous injection of energy at a microscopic scale within the samples. The microscopic dynamics explicitly break detailed balance and, in consequence, no equilibrium theorem can be used as a guideline to understand the macroscopic behaviour. Nevertheless, since the consumed energy is partly dissipated into the medium, different non-equilibrium steady states establish and they are still amenable to be studied with physics tools. Phase transitions between such steady states are possible even in low-dimensional cases. The full characterisation of the dynamic phase diagram and the various phases is one of the issues that attracts physicists’ attention.

1.2 Density, form, and dimensionality

The dynamics of systems as seemingly unrelated as flocks of birds, swarms of bacteria and vibrated rods share many features in common and, up to a certain extent, they can be treated within the common theoretical framework of active fluids or suspensions [14–24]. Different approaches aimed at a coarse-grained description based on general symmetry arguments are available but fluctuations and phase transitions have been especially analysed in the context of agent-based models. Although many papers study
motion in the dilute limit, much less is known about the behaviour of dense ensembles subject to not so strong activity. In particular, the connection with the passive limit, and their own complex phases and phase transitions, have not been studied in so much detail.

Natural and artificial active matter are rarely constituted by spherically symmetric elements. On the contrary, anisotropic objects are much more common. Examples of living realisations in which the constituents have anisotropic shapes include reconstituted layers of confluent epithelial cells [25] and biofilms formed by dense collections of rod-shaped bacteria [26]. Artificial cases also exist and an example is given by shaken non-spherical grains [27].

For concreteness, let us focus on the example of a bacterial colony, a many-body system made of elongated constituents. Bacterial colonies occupy three-dimensional spaces but they can also be confined to two dimensions when deposited on agar plates. The collective motion of such systems was studied by biologists since long ago with the aim of understanding the chemotaxis and sensory transduction mechanisms [28]. More recently, physicists have entered this field and, in particular, they have identified dynamic phase transitions, order parameters, topological defects, etc. [29–34]. Similar experimental studies were carried out in cell populations [35–37].

The way in which bacteria swim, via a dipolar or multipolar field, is very different from the one in which macroscopic animals do. At micron length scales, inertial effects are negligible compared to viscous forces. Effectively, the swimmers evolve in a zero Reynolds number limit. We can therefore ensure that dense systems of anisotropic self-propelled elements confined to two dimensions, in the low Reynolds number limit, are apt to mimic swarms of bacteria and present interesting challenges from the material science point of view. Before continuing with the description of some general features of active and passive materials, we announce that the subject of this paper is the study of a minimal model for such systems, where the dynamics of the solvent is neglected and only interactions and coarsening phenomena depending on the depth of the critical point at finite Pe, and identified nucleation events within the clusters at high Pe, these authors concluded that the phase separation is between a fluid and an active solid, similar to what is seen near the solid-hexatic transition point in a passive two-dimensional system (see the next section). They pointed towards the existence of a critical point at a lowest non-vanishing Pe with liquid-gas coexistence was conjectured [19,59–62] (e.g., found from a linear stability analysis of the homogeneous fluid replacing all interactions by an effective propulsion that decreases with density). However, to date, there is no clear-cut evidence of it in numerical studies.

More recently, phase separation was demonstrated in models of soft mono-disperse [63] and poly-disperse [64] non-aligning self-propelled disks interacting through harmonic repulsion (with a cut-off) in two dimensions. In these papers the onset of freezing (in the polydisperse case) and phase separation were determined from the analysis of the translational mean-square displacement and the number fluctuations, respectively. A bit later, Redner, Hagan and Baskaran [65] simulated active Brownian isotropic particles with excluded-volume interactions and no alignment. They confirmed phase separation within a coexistence boundary that is similar to the binodal curve of an equilibrium fluid with Pe playing the role of the attraction strength. From the analysis of the hexatic order within the clusters at high Pe, these authors concluded that the phase separation is between a fluid and an active solid, similar to what is seen near the solid-hexatic transition point in a passive two-dimensional system (see the next section). They pointed towards the existence of a critical point at finite Pe, and identified nucleation events and coarsening phenomena depending on the depth of the quenches performed. Finally, we mention that phase separation was also found in a simple model of Brownian disks in which self-propulsion velocities are defined in terms of a persistent Gaussian noise rather than having fixed norms [66–70].

1.3 Motility-induced phase separation (MIPS)

One of the salient features of active matter is its ability to cluster in the absence of cohesive forces.

Clustering in Vicsek-type models [54,55], induced by explicit orientation interactions, and in self-propelled hard rods, favoured by the apolar alignment of the anisotropic particles [56–58], was signalled in the early active matter literature.

Cates et al. [19,59,60] argued that clustering and phase separation are generic properties of systems driven out of equilibrium by a persistent local energy input that breaks detailed balance. As a consequence, motility-induced phase separation is generated in active materials in the absence of any attractive interaction or breaking of the orientational symmetry.

The picture that emerged is based on an analogy with the liquid-gas transition. The homogeneous gaseous phase in the active system would become metastable on a binodal line in the (Pe, φ) phase diagram and next unstable on a spinodal line, (Pe and φ are dimensionless numbers to be defined later in our model; they represent the strength of the active force compared with temperature and particle density, respectively.) Nucleation of the stable (liquid) phase between these two lines would be inhibited by the huge timescales needed, in most cases. The existence of a critical point at a lowest non-vanishing Pe with liquid-gas coexistence was conjectured [19,59–62] (e.g., found from a linear stability analysis of the homogeneous fluid replacing all interactions by an effective propulsion that decreases with density). However, to date, there is no clear-cut evidence of it in numerical studies.

1.4 Two-dimensional melting

The properties of two-dimensional atomic and molecular passive systems and, especially, their melting transition, is a fascinating subject that has attracted the attention of physicists since long ago. Experimental and numerical studies have not yet established, beyond any doubt, which are the mechanisms that lead to melting. Two dimensions is such a special case since the Mermin-Wagner theorem inhibits the spontaneous symmetry breaking of the translational symmetry and, accordingly, positional order can only be quasi-long-range in bi-dimensional systems. A different type of order, orientational order between the imag-
binary bonds that link the centres of neighbouring particles, can instead be truly long range in a 2d solid.

The pioneering numerical study carried out by Alder and Wainwright [71] of the 2d melting of a hard-disk solid lead to the conclusion that the melting transition is first order even in 2d. However, around 15 years later, Halperin and Nelson (HN) [72], and independently Young (Y) [73], worked out a theory for 2d melting mediated by the unbinding of topological defects. This theory is based on the previous ideas of Kosterlitz and Thouless [74] (also Berezinskii [75]) and predicts two infinite-dimensional phase transitions leading from a solid (with quasi–long-range positional and long-range orientational order) to a hexatic phase (with short-range positional and quasi–long-range orientational order) to the liquid (where both orders are only short ranged). The two phase transitions are proposed to be driven by the unbinding of dislocations (solid-hexatic) and the unbinding of disclinations (hexatic-liquid). Needless to say, the experimental and numerical determination of the transition are difficult and this fact makes the distinction between the two pictures a very interesting issue. The literature on this problem is vast; nevertheless, its full resolution has remained elusive.

The two scenarios should, in principle, be distinguished by the finite size behaviour of the hexatic order parameter in the intermediate phase. However, its careful determination is quite impossible to make. The importance of looking at higher moments and correlation functions of this order parameter was stressed in [76] and the analysis in this paper points towards a first-order phase transition in a hard-disk system.

Around 10 years ago, Krauth and collaborators came back to this problem [77–80]. In a series of papers, they proposed that the transition mechanism depends on the range of the potential, characterised by its short distance behaviour, \( U(r) \propto r^{-2n} \). For relatively smooth potentials with \( n \leq 3 \) the HNY scenario was reproduced by their numerical simulations. Instead, for sufficiently hard repulsive potentials (\( n > 3 \)) the transition was found to be of mixed kind, with the solid-hexatic being as predicted by HNY and the hexatic-liquid being first order. These numerical studies were performed with an event driven Monte Carlo algorithm that allows one to reach very long times and equilibrate systems with unprecedented large sizes. The effect of the interaction potential between the elementary constituents was examined in [80]. Spherically symmetric particles were always considered in these papers.

Matter is not only made of symmetric objects. Molecules have non-trivial shapes and the form of the elementary constituents have non-trivial effects on the phase diagram and dynamic behaviour. The crudest model for a diatomic molecule is a dumbbell made by identical beads attached to each other by an edge linking their centres. This is usually considered to be the simplest non-convex body. The liquid theory of ensembles of such molecules and their numerical investigation was the subject of intensive studies, some of them published in refs. [81–84]. Certain experimental systems, such as monolayers of diatomic molecules absorbed on crystalline surfaces correspond, in an idealised representation, to a bidimensional dumbbell system [85]. One can easily imagine that dumbbell systems can jam more or less easily depending on the relation between the bond length and bead diameter, and the rigidity of the link, and these parameters have an effect on the melting/solidification mechanisms.

### 1.5 MIPS & melting?

A natural question to ask is whether MIPS and the coexistence of dense and loose regions, with and without hexatic order, predicted by the two-step melting scenario of Krauth et al. are continuously connected under activity or not at all, that is to say, whether they are independent phenomena. This question was asked in [53] for the active dumbbell system and evidence for there being a continuous relation between the two was given in this paper. The resulting phase diagram is announced in fig. 1 and it will be further discussed in the text. The comparison to recent studies of active disks [86–88] in which similar questions are posed will be given in the concluding section.

This paper contains an extended review of the studies that some of us presented in refs. [47–53] and new results based on further simulations performed for this work, that complement our analysis of the bi-dimensional active dumbbell system in repulsive interaction. In ref. [53] the phase diagram of the active dumbbells fluid was shown only for Pe \( \leq 40 \), while in this paper we complement the study exploring a wider region. We also present a new detailed analysis of structure properties using the orientational and positional correlation functions, and of dynamic properties by analysing the angular velocity of clusters and the enstrophy distribution functions, that we...
correlate to the dumbbells orientation and the polarisation field within these structures. We confirm that we do not see any discontinuity in the behaviour of structural and dynamic observables within the coexistence region in the phase diagram.

The structure of the paper is the following. In sect. 2 we recall the definition of the model and we give values of the parameters that we use to obtain the original results presented in this paper. Section 3 presents a summary of the behaviour of the single active dumbbell. The two next sections are devoted to the presentation of our results. In sect. 4 we focus on the structural properties of the system in its various steady states: in sect. 4.1 we define a handful of instantaneous observables and in sect. 4.2 we show and discuss their numerical measurements. In sect. 5 we define and evaluate several dynamic observables. We take the opportunity to clarify some misconceptions in the literature concerning the identification of effective potential with a cutoff at its minimum $2^{1/n}\sigma_d$ (blue curve), the Mie potential that we used and that is defined in eq. (4) (red curve) and a sharp wall potential (black line).

The total internal potential energy of the system is

$$U = \sum_{i=0}^{2N} \sum_{j=0, j \neq i}^{2N} U_{\text{Mie}}(|\mathbf{r}_i - \mathbf{r}_j|),$$

with $U_{\text{Mie}}$ being a purely repulsive potential defined as

$$U_{\text{Mie}}(r) = \left\{ 4\epsilon \left( \frac{\sigma}{r} \right)^{2n} - \epsilon \left( \frac{\sigma}{r} \right)^n \right\} + \epsilon \theta(2^{1/n}\sigma - r),$$

where $\sigma$ and $\epsilon$ are parameters setting length and energy scales of the potential. The functional form in eq. (4) is abruptly set to zero at its minimum, located at $r = 2^{1/n}\sigma$, by the Heaviside function $\theta$. We considered $2^{1/n}\sigma = \sigma_d$, in order for the minimum to be equal to the disk diameter in such a way that the potential derivative be continuous in $\sigma_d$. A high $n = 32$ value is chosen in order for $U_{\text{Mie}}$ to be as close as possible to the hard-disk case without losing computational efficiency, see fig. 2. Note that eq. (1) also has implicit additional forces that keep the distance between the two spheres of each dumbbell constant. These are taken into account while performing the numerical integration, see sect. 2.2.

The active force $\mathbf{F}_{\text{act}}$ acts on the tail-to-head direction of each dumbbell and has constant modulus $F_{\text{act}}$. Its direction changes following the individual dumbbell’s orientation.

The initial positions and velocities of the dumbbells are specified in sect. 4.2.1. We have made several choices in order to check that, after a sufficiently long transient, the dynamics reach a steady state in which the details of the initial configuration are forgotten.

The dimensionless control parameters are the surface fraction covered by the beads,

$$\phi = \frac{N\pi\sigma_d^2}{2L^2},$$
and the Péclet number,

$$Pe = \frac{2F_{\text{act}}\sigma_d}{k_B T}.$$  \hspace{1cm} (6)

The latter can be seen as the ratio between the advective transport of a fixed amount of material flowing in unit time through a section of a characteristic length $\ell$ induced by the active force, $v_{\text{act}}/\ell = F_{\text{act}}/(\sigma_d\gamma_d)$ and the diffusive transport of the same amount of material due to thermal fluctuations, $D/\sigma_d^2 = k_B T/(2\gamma_d\sigma_d^2)$. It can also be interpreted as the ratio between the work done by the active force when translating the dumbbell by its size, $2\sigma_d F_{\text{act}}$ over the thermal energy scale $k_B T$. An active Reynolds number $Re = m_d F_{\text{act}}/(\gamma_d^2 \sigma_d)$ can be defined in analogy with the hydrodynamic one. It is the ratio between the active force transport $v_{\text{act}}/\ell$ and the viscous transport term estimated as $\gamma_d/m_d$ and takes very small values for the parameters chosen in all our numerical studies. In practice, the active Reynolds number takes in our case values in the range $Re \in [0-0.05]$, and can be considered negligible, a choice that corresponds to the modelling of bacterial colonies, as mentioned in the introductory section. The value of the active force giving a speed value that is compatible with the one observed experimentally for bacteria (of the order of $[20, 50] \mu m/s$) should imply in our simulations Pe values equal to several hundreds.

In the single-dumbbell case eq. (1) can be solved analytically and a summary of the motion in such extreme dilute limit will be given in sect. 3. For a finite density system an analytical solution is out of reach and one has to resort to a numerical integration of the coupled equations. In the next subsection we give some details on the numerical scheme implemented.

### 2.2 Numerical integration

We fixed the number of particles in our system to $N = 128^2, 256^2, 512^2$ and, accordingly, we set $L^2 = N \pi \sigma_d^2/(2\phi)$ in order to have the desired packing fraction $\phi$.

All physical quantities are expressed in reduced units of the sphere’s mass $m_d$, diameter $\sigma_d$ and potential energy $\epsilon$ [90]. The time unit is the standard Lennard-Jones time $\tau_{LJ} = \sigma_d^2/(m_d \epsilon)^{1/2}$. Other important simulation parameters that we used, in reduced units, are $\gamma_d = 10$, $k_B T = 0.05$ and we set $k_B = 1$. The large $\gamma_d$ assures the overdamped limit. Typical simulations took between $10^5$ and $10^6$ simulation time units (MDs), where longer runs are required to ensure stationarity in the co-existence region [53]. On average each simulation lasting $5 \times 10^5$ MDs was run on 16 processors for a total of 100 hours for each cpu.

We used a velocity Verlet algorithm that solves Newton’s equations of motion, plus additional force terms for the Langevin-type thermostat, to numerically integrate the stochastic evolution equation. We kept the bonds rigid with the help of the RATTLE scheme [91]. This is equivalent to considering an additional force in eq. (1), that takes into account the holonomic constraints. The time-step choice is related to the force exerted during the simulation. We adapted it to enforce numerical stability. In this paper, for systems at $Pe \leq 10$ we used a time-step of $0.005$, while for $Pe = 20$ and $Pe = 40$ we used a time-step equal to $0.002$ and, finally, for $Pe = 100$ and $Pe = 200$ the time-step was reduced to $0.001$.

In order to efficiently parallelise the numerical computation we used the open source software Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS), available at github.com/lammps [92].

In addition, we also studied the behaviour of the system using the Weeks-Chandler-Andersen (WCA) potential between the disks, that corresponds to $n = 6$ in the potential $U(r)$ and a truncation at $r_c = 2^{1/6} \sigma_d$, see fig. 2. The dumbbells were still taken to be rigid with a distance $\sigma_d$ between the centres of their beads. We will briefly mention the results found with this other potential in the text.

### 3 Single-dumbbell limit

The analytic solution to the Langevin equation regulating the dynamics of a single active dumbbell was given in refs. [49,50]. We do not repeat the derivation here but we simple summarise some features that will be relevant to identify important timescales and understand the dynamics of the interacting problem.

The individual dumbbell undergoes centre-of-mass displacement and rotational motion, the latter being due to the random noises that act independently on the two disks.

The equation ruling the centre-of-mass motion is a Langevin equation for a point-like particle with mass $2m_d$, under the force $2F_{\text{act}}$ and in contact with a bath with friction coefficient $2\gamma_d$ and temperature $T$.

Four time regimes can be identified from the time-delay $\Delta t$ dependence of the mean-square displacement (MSD) of the centre of mass

$$\Delta^2_{\text{cm}}(\Delta t) = \langle (r_{\text{cm}}(t + \Delta t) - r_{\text{cm}}(t))^2 \rangle.$$  \hspace{1cm} (7)

Here and in what follows the angular brackets denote an average over independent thermal noise histories and initial conditions. The four regimes, shown in fig. 3 with a blue solid line, are:

- I) A ballistic regime at very short time differences, $\Delta t \ll t_I = 2m_d/(2\gamma_d)$, with velocity $v_{\text{cm}} = k_B T/(2m_d)$.
- II) A diffusive regime common to the one for the single passive dumbbell, with diffusion constant, $D = 1/(2d) \lim_{\Delta t \to \infty} d \Delta^2(\Delta t)/dt$ that takes the thermal value $D = k_B T/(2\gamma_d)$. This regime lasts until a crossover dictated by the strength of the active force
In the stiff case in which the molecular vibrations are completely frozen, the stochastic equation for the angle can also be solved and the angular MSD computed. In the single-molecule case, this quantity has less interesting features in the sense that there are only two time-regimes and, for a single molecule, it has only two timescales, separated by $t_I$.

$$t^* \propto t_a/ Pe$$  \hspace{1cm} with \hspace{1cm}  t_a = D_R^{-1} = \gamma_d \sigma_d^2/(2k_BT), \hspace{1cm} and \hspace{1cm} D_R$$

the rotational diffusion constant, see below. A sufficiently strong active force can erase this regime.

- III) A second ballistic regime due to the activity that lasts until $t_a$, with a velocity $v \propto F_{act}/\gamma_d$.

- IV) A final diffusive regime with diffusion constant $D_A = k_BT(1+Pe^2/8)/(2\gamma_d)$, that increases as $Pe^2$ for large active forces.

Obviously, if the parameters are such that the timescales $t_I$, $t^*$ and $t_a$ are not well separated, the regimes overlap. Meanwhile the rotational motion can also be studied by following the angular MSD:

$$\Delta^2_d(\Delta t) = \langle (\theta(t+\Delta t) - \theta(t))^2 \rangle.$$  \hspace{1cm} (8)

In the stiff case in which the molecular vibrations are completely frozen, the stochastic equation for the angle can also be solved and the angular MSD computed. In the single-molecule case, this quantity has less interesting features in the sense that there are only two time-regimes and, for a single molecule, it has only two timescales, separated by $t_I$.

The rotational motion of a single dumbbell is independent of the active force.

The various time regimes described above also manifest in the probability distributions of the single run mean-square and angular displacements of a single dumbbell (not averaged over the noise). The results concerning these quantities can also be found in [49,50]. We will come back to them in sect. 6 where we will discuss the finite density effects on the dynamics of the many-body system.

As is well known [93–99], under certain conditions (slow dynamics, small entropy production, etc.) the long-term dynamics of some non-equilibrium systems (glasses, dense sheared liquids, weakly shaken dense granular matter) can be described (at least partially) in terms of an effective temperature. One way to identify this temperature is to calculate the deviations from the fluctuation dissipation theorem (FDT) linking the integrated linear response

$$\chi(\Delta t) = \frac{1}{\hbar} \sum_{a=1}^{d} \int_{t}^{t+\Delta t} dt' \frac{\delta(v_{cm}^a(t+\Delta t))}{\delta h_a(t)} \bigg|_{h=0}$$  \hspace{1cm} (9)

to a perturbation $h$ applied, in this case, to the centre of mass of the dumbbell, and the mean-square displacement $\Delta^2_{cm}(\Delta t)$ of the same observable, that is to say, the centre of mass of the dumbbell here. Using the expression of the FDT in the equilibrium form,

$$2k_BT \chi(\Delta t) = \Delta^2_{cm}(\Delta t),$$  \hspace{1cm} (10)

one can extract the effective temperature by replacing $T$ with $T_{eff}(\Delta t)$ in the formula above. One has, though, to carefully keep in mind that the dynamics take place in various time-regimes (see the four items listed above) and that the effective temperature may take different values in each of them or even have a meaning only in some of them [93].

In ref. [49] we also computed the effective temperature thus defined of a single active dumbbell and we found that in the last diffusive regime called IV, and defined by $\Delta t \gg t_a$, it takes the form

$$k_BT_{eff} = k_BT(1+Pe^2/8).$$  \hspace{1cm} (11)

This effective temperature characterises the dynamics at long time-delays and should not be confused with, e.g., the one extracted from the velocity fluctuations, $T_{kin}$, that can only access the instantaneous properties (high frequency regime) of the system. For example, for the single dumbbell the kinetic temperature extracted from the averaged kinetic energy and the equipartition assumption is equal to

$$k_BT_{kin} = \langle E_{kin} \rangle = k_BT \left[ 1 + \frac{m_a k_BT}{(2\gamma_d \sigma_d)^2} Pe^2 \right]$$  \hspace{1cm} (12)

and, although the difference with $T_{eff}$ is just in the pre-factor of the term proportional to $Pe^2$, it is clear that the dependence on all the other parameters in the model is not the same. The conceptual and quantitative difference between $T_{eff}$ and $T_{kin}$ is well known in the field of glasses and it will manifest more clearly when dealing with many-body active dumbbells in interaction. We will come back to the interpretation of the effective temperature by the end of this paper, where we will discuss its properties in the many-body active system.

4 Structure

We first present the results of our study of the structural properties of the system.
4.1 Observables

In this subsection we list and explain the definition of the various observables that we used to quantify the degree of order in the dumbbell system and to identify the different phases.

4.1.1 Local density fluctuations

We used two methods to compute the local densities $\phi_j$ and we checked that they give equivalent results.

With the first method, for each bead, we first estimate the local density as the ratio between its surface and the area of its Voronoi region $A^{\mathrm{Vor}}_i$. We next coarse-grain this value by averaging the single-bead densities over a disk with radius $R$ around the $j$-th bead,

$$
\phi_j = \frac{1}{n^{(j)}_R} \sum_{i \in \mathcal{S}^{(j)}_R} \frac{\pi \sigma_i^2}{4A^{\mathrm{Vor}}_i},
$$

where $n^{(j)}_R$ and $\mathcal{S}^{(j)}_R$ are the number of particles and the spherical disk with radius $R$ around the $j$-th bead, respectively. We took, typically, $R = 20 \sigma_d$, and we found equivalent results using radii in the interval $[10,50] \sigma_d$. At $\text{Pe} > 10$ we also used $R = 5 \sigma_d$. For a visual inspection of the density fluctuations in real space, each Voronoi region is then painted with the colour that corresponds to its coarse-grained local density, using a heat map with the usual convention: denser in red, looser in blue.

With the second method, we constructed a square grid on the simulation box. For each point in the grid we calculated a coarse-grained local density over a circle of given radius $R$. We finally assigned this density value to the grid point.

The data stemming from the use of both methods are collected and pdfs of the local densities thus obtained are shown below. Some of them are shown above.

4.1.2 Orientational order

With the aim of exhibiting orientational order, we computed the local hexatic order parameter

$$
\psi_{\theta j} = \frac{1}{N^{\theta}_{\text{nn}}(j)} \sum_{k=1}^{N^{\theta}_{\text{nn}}(j)} e^{i \theta_{jk}},
$$

where $N^{\theta}_{\text{nn}}(j)$ is the number of nearest neighbours of bead $j$ found with a Voronoi tessellation algorithm [100] and $\theta_{jk}$ is the angle between the segment that connects $j$ with its neighbour $k$ and the $x$ axis. For beads regularly placed on the vertices of a triangular lattice, each site has six nearest-neighbours, all the angles are integer multiples of $2\pi/6$, $\theta_{jk} = 2k\pi/6$, and $\psi_{\theta j} = 1$. Deviations from 1 indicate deviations from perfect orientational ordering.

We visualised the local values of $\psi_{\theta j}$ as proposed in ref. [78]: first, we projected the complex local values $\psi_{\theta j}$ onto the direction of their space average. Next, each bead was painted according to this normalised projection. Zones with orientational order have uniform colour, whatever it is. The majority ordering is always coloured in dark red and the hierarchy follows the scale shown at the extreme right of the panels in the figures. Due to the six-fold symmetry of the ordered state, blue regions are hexatically ordered along a lattice which is rotated by $\pi/2$ with respect to the one of the dark red regions. Green spots, which correspond to zero in the colour code, can represent either particles with disordered neighbours or ordered regions rotated by $\pi/4$ from the red ones. Therefore, while the greenish regions surrounding clusters are disordered, also macroscopic green patches appear, which are hexatically ordered.

With the local hexatic order parameter we can compute correlation functions

$$
g_6(r) = \frac{\langle \psi_{\theta j}^* \psi_{\theta k} \rangle |r_{j}-r_{k}|=r}{\langle |\psi_{\theta j}|^2 \rangle}. \quad (15)
$$

The conventional HNY scenario for passive systems predicts that this correlation function should approach a constant in the solid phase, it should vanish as a power law in the hexatic phase, and it should decay exponentially in the liquid phase.

We also considered the modulus of the average per particle and the average per particle of the modulus of the local hexatic order parameter,

$$
2N \psi_6 \equiv \sum_{j=1}^{2N} |\psi_{\theta j}|, \quad 2N \Gamma_6 \equiv \sum_{j=1}^{2N} \langle |\psi_{\theta j}| \rangle \quad (16)
$$

and their variation with the two control parameters $\phi$ and $\text{Pe}$.

4.1.3 Positional order

Positional order is put to the test by the usual two-point correlation functions

$$
C_{q0}(r) = \langle e^{i q \cdot (r_i-r_j)} \rangle \quad (17)
$$

with $r = |r_i-r_j|$ and $q_0$ the wave vector that corresponds to the maximum value of the first diffraction peak of the structure factor:

$$
S(q) = \frac{1}{2N} \sum_{i=1}^{2N} \sum_{j=1}^{2N} \langle e^{i q \cdot (r_i-r_j)} \rangle. \quad (18)
$$

In the arguments of $C$ and $S$ we have already used the translational invariance and isotropy of the system (on average) and we only wrote the absolute values of the distance and wave vector. The quasi-long-range positional order in the solid phase should be evidenced by the algebraic decay of the positional correlation function. Instead, in the hexatic and liquid phases the decay of $C_{q0}(r)$ should be exponential. Note also that the formulas for $C$ and $S$ explicitly need the use of the disks positions,
and not the dumbbells centres of mass to observe hexatic and positional orders. In fact, while the disks constituting the dumbbells pack at high density on a triangular lattice, the dumbbells centres-of-mass form non-periodic structural arrangements when packed [101].

4.1.4 Polarisation

In order to investigate the alignment of head-tail dumbbell orientations inside the clusters, which is absent in the passive case but expected to depend on the strength of the active force, we computed the local polarisation field \( \mathcal{P} \) and we evaluated its probability distributions. For each dumbbell in the dense hexatically ordered phase we considered a unit vector \( \mathbf{p}_i \) in the tail-to-head direction. We constructed a square grid on the simulation box and for each point on the grid we calculated the spatial component of a coarse-grained local polarisation by averaging over the values that this quantity takes for all dumbbells within a disk with a previously chosen radius \( R \). We assigned the averaged outcome to the grid point. Normalisation is such that a perfect polar order in each cell of the grid would give a vector of unit magnitude. We used different radii, ranging from \( R = 5 \sigma_d \) to \( R = 25 \sigma_d \), finding qualitatively similar results. To visualise the local values of \( \mathcal{P} \), we evaluated its modulus and we coloured each zone according to its magnitude.

4.2 Numerical results

In this section we sum up some of the results that we obtained in previous studies of the interacting active dumbbell system [47–53] and we present data for quantities that we had not studied before and complement our analysis. We chose to show data for parameters on curves in the phase diagram that lie within the region with co-existence and on which a fixed area proportion of the two phases, \( 25\%-75\% \) and \( 50\%-50\% \), is maintained. (We will always give the percentage of dense phase first.) The selected points in the phase diagram for which we show more results are taken to be at the parameter values \( \text{Pe} = 0, 10, 20, 40, 100, 200 \) and the corresponding \( \phi \)'s to ensure the chosen repartition of phases.

4.2.1 Initial states

In all our simulations we used three kinds of starting configurations that we call "striped", "random" and "hexatic-ordered". These configuration were obtained as follows.

Random initial conditions. We placed the dumbbells at random positions, with random orientations, in continuous space. At large \( \phi \) these states have, very likely, overlapping dumbbells. Consequently, these configurations can have a very high energy. For the parameters used, the maximum overlap that we accepted is 0.75 \( \sigma_d \). Therefore, we released the excess energy by letting the system equilibrate with a softer Lennard-Jones potential and a smaller time-step equal to 0.001. After this procedure, convenient starting configurations as the one shown in the first figure in the Supplemental Material in [53], were found.

Striped states. We placed the dumbbells in a closed-packed triangular lattice, by placing the dumbbells in sequence starting from the bottom-left end, say. This construction is followed until the dumbbell number satisfies the required global density. An empty slab is therefore left in order to have the right global density. The orientational order in the configuration is then randomised with Monte Carlo moves that take two adjacent dumbbells and exchange their bonds. In this way, the molecules are still placed in a crystalline configuration but their directions are not ordered.

Hexatic ordered states. We took a crystalline configuration in equilibrium with hexatic order like the one depicted, with packing fraction just above the co-existence region at \( \text{Pe} = 0 \), obtained by equilibrating the system starting from a striped configuration with this packing fraction. We then expanded (or contracted) the configuration by multiplying the coordinates with a convenient factor \( \alpha \). Accordingly, the system's linear size is now rescaled to \( \alpha L \), which is at most 20 \( \sigma_d \) larger (or smaller) than the original one, depending on the chosen density that we want to simulate. Finally we applied a short equilibration run with a smaller time-step equal to 0.001. For the densities considered, the rescaling is small enough to preserve the initial hexatic ordering without blowing up the simulation.

The long-time limit of the evolution of initial states of these three types are statistically equivalent for all parameters used. We will not show here results for all these cases, as more details on how the initial configurations are forgotten dynamically can be found in [53].

4.2.2 Phase diagram

Following initial studies in, e.g., [47,48,102] that focused on the boundary enclosing MIPS at high Pe, in [53] we presented the full phase diagram of the active dumbbell model. The main point made in this paper is that

- co-existence between a dense and a dilute phase extends all the way to \( \text{Pe} = 0 \), with no "critical" ending point at a non-zero Pe, in a scenario different from what has been stated in the literature so far.

The dilute phase has no order and behaves as an active liquid or gas while the dense phase has orientational order and it is therefore an active hexatic phase. These results made contact with the recent advances in the understanding of two-dimensional melting of passive disk systems in interaction: namely, the first order kind of the transition between liquid and hexatic phases and the ensuing co-existence between the two [76–80]. Here, we extend the results found for \( \text{Pe} \leq 40 \) in [53] to \( \text{Pe} \leq 200 \).

The phase diagram is displayed in fig. 1 with the grey level indicating the percentage of area occupied by the dense hexatically ordered phase. It is based on a detailed
plan of simulations performed at the values of Pe mentioned above. The various coloured lines are curves of constant area covered by the dense phase. We explain in the next subsection how we obtained this phase diagram using different observables. Notice that since the disks forming the dumbbells are not completely hard, some overlap between them is possible and values of $\phi$ that are slightly larger than the close-packing limit can be accessed in the simulation (recall that the close-packing fraction of disks in two dimensions is achieved by a perfect triangular lattice and it amounts to $\phi_{cp} \approx 0.91$).

### 4.2.3 The local density and hexatic order

One way to estimate the critical lines $\phi(Pe)$ for the upper limit for the pure liquid phase (dotted black curve in fig. 1) and the lower limit for the pure solid phase (white dotted curve in the same plot), is to follow the evolution of the probability distributions (pdf) of the local density and modulus of the hexatic order parameter, and search for the appearance and disappearance of the two-peak structure in these functions [53].

Figure 4 represents the value of the local density at which the high density peak of the pdf is located as a function of the Péclet number, following the curves of 50%-50% and 25%-75% repartition of dense and loose regions in the sample. At Pe = 0 the curves take a non-vanishing value proving that there is co-existence in the passive limit. The growth with Pe is then monotonic, the curves go slowly, and they go a bit beyond the close-packing limit at large Pe due to the fact that the potential is slightly soft.

The data in fig. 4 were obtained from the analysis of the pdfs of $\phi$ shown in figs. 5 and 6, together with snapshots of the configurations using a colour code that represents the different hexatic orders. The data are taken along curves of constant proportion 25%-75% (fig. 5) and 50%-50% (fig. 6) of the two phases. Note that the vertical scale of the different panels in the first row is not the same. We see that

- as Pe increases, the high density peak continuously moves towards higher local densities and its weight increases while the low density peak moves in the opposite direction and its weight decreases, leading to the results already shown in fig. 4. As far as density is concerned we do not find any discontinuity when moving towards higher activities in the coexistence region.

The local hexatic order parameter maps shown in the lower rows in figs. 5 and figs. 6 allow one to understand the mechanism whereby phase separation takes place in the active problem and the reason why the region of the phase diagram with co-existence inflates as Pe increases. Indeed,

- as the activity is turned on, some spatial regions get denser and acquire orientational order, leaving away disordered holes. Depending on the strength of the activity, this process allows the dumbbells in the ordered regions to pack in a single ordered domain, or in polycrystalline arrangements concerning the orientational order.

In the maps in figs. 5 and 6 the fluid disordered regions are greenish while the denser, orientationally ordered ones, take all the other colours. At high Pe values, for instance in the Pe = 200 case, a single macroscopic hexatic phase is not yet formed. At lower area percentage of the higher density phase this is due to the fact that the timescale of clusters diffusion, which rules aggregation, goes well beyond the duration of our simulations. Moreover, the defects between the various orientationally ordered domains in the same clusters, that can be seen when the two phases occur with equal area percentages, should also slowly heal with time, but our simulations do not reach those exceedingly long timescales.

### 4.2.4 Polarisation

We follow the evolution of the probability distributions of the local polarisation modulus in the dense phase, defined in sect. 4.1.4, along the curves of constant area fraction 50%-50% (fig. 7) and 25%-75% (fig. 8) of the liquid and hexatic phase.

For low Pe values (Pe = 0, 10) the pdfs show a single peak at small values of the local polarisation modulus both on the 50%-50% and 25%-75% lines, revealing the absence of any polar order. This is also evident from the coarse-grained polarisation fields of the clusters, that look completely disordered regardless of being inside or at the border of clusters and dominated by low-polarisation values with sporadic negligible ordered parts.

As the activity increases (Pe $\geq$ 20) the probability distributions, both on the 50%-50% and 25%-75% lines, exhibit a two-peak structure showing that activity tends to induce an effective polar interaction between dumbbells, reflected in the emergence of large ordered patterns in the clusters polarisation field. While along the 50%-50% lines the coarse-grained polarisation field remains typically with no polar order (white regions) in the central
Fig. 5. Pdfs of local packing fraction (above) and instantaneous local hexatic order parameter maps (below) at different Pe numbers shown as labels in the upper plots, along the curve with an amount of 25% in area of the dense phase. The pdfs are evaluated by averaging over configurations taken from at least five independent runs in the stationary state. On the second row we show some illustrative configurations taken after $500000 \tau$, when we can consider the systems in a stationary regime. As pointed out in the main text, especially at Pe = 200 a single macroscopic hexatic phase is not formed due to the timescale of clusters diffusion, which rules aggregation, being much longer than the duration of our simulations. Even though in this case we performed up to ten independent runs finding a similar behaviour, we do expect further coalescence, at still longer times. The colour code of the snapshots, present at the right extreme, is such that each bead is coloured according to the local hexatic order parameter $\psi^6$, defined in eq. (14), projected onto the direction of its global average. For this reason, the colour appearing most often is the closest to one (dark red) and zones with the same uniform colour are associated to the same $\psi^6$ value, and therefore have the same kind of orientational order.

Fig. 6. Pdfs of local packing fraction (above) and maps of local hexatic order parameter (below) at different Pe numbers along the curve with equal percentage in area of the dilute and dense phases. The presentation and colour code are the same as in fig. 5 and they are explained in its caption. For these parameters, the total time used in the simulation was enough to obtain a single dense cluster (though polycrystalline from the orientational point of view) in all cases apart from (d). Part of the clusters and ordered on the border, along the 25%-75% line and Pe $\geq 40$ the peak corresponding to random orientation tends to disappear. In these cases, the disordered region at the centre of each cluster is reduced to a small white spot, otherwise stated, a defect for the polarisation field. A deeper look into the local polarisation field, see fig. 9, reveals the emergence at Pe = 40 of an aster structure with a charge equal to $q = 1$, see panel (b). This feature can be better appreciated by showing enlargements of a cluster’s snapshot, overlapped with the coarse-grained polarisation field, close to the interface where dumbbells point inward (fig. 10(a)), and at the centre where the defect appears (fig. 10(b)). At larger activity (Pe = 100, 200), the aster evolves into a spiralling pattern, see panels (c) and (d) in fig. 9 for two enlarged clusters. Notice that this spiralling pattern is associated to an increased angle between the defect and the local polarisation field, that will be characterised later in this paper, see for instance fig. 20.

To clarify the reason for the appearance of the aster pattern with a defect at the cluster’s centre we can trace back the evolution of one of the dumbbells that is located near the defect, at Pe = 40. In this way we can go back to the event that triggered the MIPS in the first place. In fig. 11 a series of snapshots of the formation of the cluster are shown, with the traced dumbbell coloured in orange. The cluster formation process starts with multiple head-to-head dumbbells collisions (thus without polar order).
Fig. 7. Pdfs of the local polarisation modulus $|\mathbf{P}|$ (above) and map of the local polarisation modulus (below) both smoothed over regions with radius $R = 5 \sigma_d$ at different Pe number ranging from Pe = 0 to Pe = 200 along the curve with 50%-50% proportion of hexatically ordered and liquid regions. The appearance in the pdfs of a second peak at high $|\mathbf{P}|$ and increasing Pe, signals that dumbbells prefer to be locally aligned in clusters, especially at the border. The pdfs are evaluated by an average over stationary configurations taken from independent runs, as explained in the caption of fig. 5. For clarity, the snapshots of the local polarisation modulus are taken from the same run in fig. 6, though at different times.

Fig. 8. The same quantities as in fig. 7 here smoothed over regions with radius $R = 10 \sigma_d$ at the same Pe numbers as in the previous figure, now taken along the curve with an amount of 25% in area of the dense phase. Zooms over the clusters surrounded by black boxes in (b), (d), (e), (f) are shown in fig. 9, and over those surrounded by two cyan boxes at the centre and border of the cluster in (d) are shown in fig. 10.

Fig. 9. Local polarisation field $\mathbf{P}$ at Pe = 10, 40, 100, 200 in the 25%-75% line. The four clusters shown are the ones surrounded by black squares in the lower row in fig. 8. Increasing Pe, the clusters exhibit a spiralling pattern with an increasing angle between $\mathbf{P}$ and the vector pointing toward the cluster centres, see also fig. 20.
Fig. 10. Enlargements of the regions surrounded by cyan squares in the snapshot of the system at $Pe = 40$, $\phi = 0.310$ (along the 25%-75% line), shown in fig. 8. The colour scale is associated to modulus of the local polarisation field. The arrow indicates the tail-to-head direction on each dumbbell. (a) Interface between the cluster and the disordered phase. (b) Centre of the cluster.

Fig. 11. Snapshots taken from a single run at four different times, $t = 66000\,\tau$, $67000\,\tau$, $72000\,\tau$, $78000\,\tau$. They show how the cluster is formed around the “central” dumbbell coloured in red that triggers aggregation. The parameters are $Pe = 40$ and $\phi = 0.310$.

After the initial nucleation, other dumbbells start to segregate on the cluster, with a pathway that is necessarily on average directed towards the cluster’s centre. Larger activities allow the initial colliding direction to remain fixed after segregation. This allows the polar field to be ordered in an aster (or spiral) pattern, and not disordered as observed at the lower value $Pe = 10$.

In short, the evolution of the polarisation with $Pe$ is such that:

- there is no polarisation at small $Pe$, the dumbbells within the cluster orient along the radial direction for intermediate $Pe$, and the clusters show a spiralling structure at larger $Pe$. Therefore, in addition to hexatic order, present for all strength of activities in the MIPS region [53], we observe in this paper that polar order is also relevant and differentiates between clusters at small and high activity.

4.2.5 Correlation functions

The orientational and positional correlation functions are shown in figs. 12, 13 and 14, side-by-side, for three activities $Pe = 0, 2, 10$, and various packing fractions indicated in the keys. Below these plots, typical configurations in the stationary regime of various representative pairs $(Pe, \phi)$ are displayed, in order to better understand the behaviour of the correlations. All correlation functions are plotted on a log-lin scale, in order to clearly identify transitions between exponential and power-law decays. Moreover, the highest density with exponential decay and the lowest density with power-law decay (if present) are fitted with a dotted and a dashed line, respectively.

For all the values of the activity, $Pe = 0, 2, 10$, the orientational correlation function changes from an exponential decay at low $\phi$ to an algebraic one at high $\phi$. We interpret this change as being due to the fact that the density parameter is moving across the coexistence region, so that the region above coexistence is characterised by quasi-long-ranged orientational order (see sect. 1.4). This same behaviour was found in the first order liquid-hexatic transition of passive spherical particles [78] and allows us to claim that:

- both in equilibrium and for all positive values of the activity, the orientational correlations suggest that there is coexistence between a disordered phase and a hexatically ordered one, depicted with the grey gradient in the phase diagram of fig. 1.
Fig. 12. Orientational correlation functions (a) and positional correlation functions (b), as defined respectively in eq. (15) and eq. (17) for Pe = 0 and different values of the packing fraction indicated in the keys, in log-lin scale. The highest density with an exponential decay and the lowest density with power-law decay are fitted with a dotted and a dashed line, respectively. For (b) the latter is absent. Above, three snapshots representing local hexatic parameter through the methods described in text, for Pe = 0 and $\phi = 0.740$ (c), 0.750 (d), 0.800 (e).

Fig. 13. Orientational correlation functions (a) and positional correlation functions (b) for Pe = 2, in log-lin scale and with exponential and power-law fits provided. Below, typical snapshots for Pe = 2 and $\phi = 0.890$ in fig. 12(b), which can be still fitted by an exponential).

Fig. 14. Orientational correlation functions (a) and positional correlation functions (b) for Pe = 10, in log-lin scale and with exponential and power-law fits provided. Below, snapshots for Pe = 10 and $\phi = 0.860$ (c), 0.872 (d), 0.874 (e). In the snapshot (c) the two disordered regions are highlighted by the two black squares.

5 Dynamic properties

For the moment we have focused on the analysis of the structural properties of the system, via the analysis of the local density, local hexatic order parameter, their correlation functions, etc. but we have not studied the behaviour of the velocity field. In this section we investigate various aspects of the dumbbell velocities and the translational and rotational MSD in the finite density system. The dynamic observables are also expected to strongly depend on where they are evaluated in the phase diagram. Once again, we choose parameter values on curves of constant proportion of the dense and loose phases in the region of coexistence, the most interesting region of the phase diagram.
5.1 The kinetic energy

In this work, we measured the kinetic energies of the particles that belong to the dense phase separately from the one of those that belong to the dilute phase, along the 50%-50% and the 25%-75% curves in the phase diagram. The data are shown in fig. 15. The kinetic energy of the beads that form dumbbells in the liquid phase grows as a monotonic function of Pe in agreement with the expected behaviour for a single dumbbell, the averaged kinetic energy of which is shown with a solid black line in the same figure (a $Pe^2$ dependence, see eq. (12)). Indeed,

- the dumbbells in the dilute phase are basically free, since their kinetic energy is very close to the one of independent dumbbells all along the $Pe$ values.

- The kinetic energy of the particles in the dense phase barely increases from 0.05 to 0.06.

The kinetic energy of the particles in the dense phase behaves very differently and, on the scale of the figure, their $E_{\text{kin}}$ barely increases from 0.05 to 0.06.

- The kinetic energy of the particles in the dense phase has a very weak growth, due to the fact that the mobility is suppressed inside the clusters and that clusters are massive and move very slowly.

A distinction between the kinetic energy of the two kinds of dumbbells becomes relevant at around $Pe = 30$, in the scale of the figure.

5.2 The velocity field

We evaluated the velocity field by first averaging over a typical time ($\tau_{\text{LJ}} = 1$) the instantaneous velocity of each bead relative to the centre of mass of the full system and then coarse-grained these values over square plaquettes with side $R$. We used different values of $R$ depending on the typical cluster size. For $Pe = 40, 100, 200$ the choice $R = 20 \sigma_d$ seems appropriate, while for smaller $Pe$ we had to use smaller values of $R$ to see clearly the velocity field. Figure 16 shows the polarisation map (in panel (a)), a snapshot of the coarse-grained velocity field (in panel (b)) together with the local hexatic parameter of the same configuration (in panel (c)) of a selected part of a system with $Pe = 200$ on the 50%-50% curve. The system is clearly phase separated between a single cluster made of seven regions with different hexatic order and the rest of the area with very few free dumbbells making this region almost empty. It is also clear from the velocity field that

- the cluster at $Pe = 200$ is turning around its centre and there is no signature of the domains with different hexatic orders, which compose the cluster, in the velocity field. The rotation of the clusters is a general feature at sufficiently high $Pe$ and this suggests a more detailed study that will be done in the following, with the help of the analysis of the enstrophy.

5.3 The enstrophy

The enstrophy is defined as

$$\epsilon = \frac{1}{2} \sum_r |\omega(r)|^2, \quad \text{(19)}$$

where $r$ locates the grid points, defined in the previous section, and $\omega = \nabla \times \mathbf{v}$ is the vorticity vector, with $\mathbf{v}$ the coarse-grained velocity field, calculated as described in the previous section. While the kinetic energy gives a measure of the strength of flow in the system, the enstrophy is a measure of the presence of vortices in the velocity field and it can be used to understand whether clusters rotate driven by activity. In order to measure meaningful values of the enstrophy, the correct coarse-grained scale of the velocity field must be carefully chosen. In particular, it must be larger than the particles size and smaller than the typical size of the aggregated phase. We found appropriate lengths between 10 and $20 \sigma_d$ (we opted for $10 \sigma_d$).

In figs. 17 and 18 we show the pdfs of the enstrophy, evaluated over several independent and stationary configurations. The panels display the results for systems at different $Pe$ numbers, $Pe = 20, 100, 200$, along the 50%-50% and the 25%-75% curves in the phase diagram. The
Fig. 17. Main plots: the enstrophy pdfs at Pe = 20, 100, 200 along the 50%-50% curve. Inserts: coarse-grained velocity field in four representative instantaneous configurations at the same parameters. The peaks appearing in the pdfs at Pe = 100, 200, are associated to clusters undertaking a rotational motion.

Fig. 18. Main plots: the enstrophy pdfs at Pe = 20, 100, 200 along the 25%-75% curve. Inserts: coarse-grained velocity field in four representative instantaneous configurations at the same parameters. The peaks appearing in the pdfs at Pe = 100, 200, are associated to clusters undertaking a rotational motion.

inserts are maps of the velocity field and show its typical behaviour. The probability of finding a non-zero value of the enstrophy decreases smoothly for low Pe number with an exponential tail (not distinguishable in the plot), with the snapshots of the velocity field at Pe = 20, 40 (not shown) suggesting a lack of coherent rotation. It can be checked that the exponential decay at Pe = 20, 40 is very close to the one at Pe = 0, meaning the enstrophy values are purely associated to the thermal noise. Since we do not expect any relevant vortical behaviour in the passive case, the noise level can be approximately estimated by measuring the enstrophy distribution at Pe = 0. Moreover, as noticed in sect. 4.2.4, for low values of the Pe number the coarse-grained polarisation field looks completely disordered, while at Pe = 40 the dumbbells tend to point along the radial direction towards the centre of the cluster without exhibiting a spiralling pattern. In both cases these features seem to prevent a coherent rotation. At higher Pe (see the cases at Pe = 100, 200), the probability distribution function of the enstrophy develops a multi-peak structure depending on the number and size of the clusters on scales clearly distinguishable from the noise contributions. One can say that

- For high values of the activity (Pe ≥ 100) the dumbbells arrange in spirals, the clusters in the system undertake a rotational motion and the probability distribution function of the enstrophy develops a multi-peak structure (without an exponential tail, typical of lower Pe) associated to the rotating clusters.

For intermediate values of the activity such as Pe = 70, we can also observe cases of a coherent, even if less evident, rotation.

In ref. [48] the authors presented a minimal mean-field framework within which it is possible to understand the physics of these rotating aggregates. It was noticed in this Letter that, in contrast with the case of motility-induced clusters observed with spherical particles, steric interactions quench the polarisation of the dumbbells in the rotating cluster. Each particle exerts a local torque which is balanced by the drag on the cluster, sustaining rotations. By comparing the velocity field and the polarisation we found here a substantial distinction:

- the velocity field exhibits a vortex pattern while the polarisation one is a spiral.

In the same context, it was inferred that the angular velocity scales with cluster size as $1/R$ and found that its order of magnitude is around $10^{-4}$. If one accepts the simplifying assumption that the clusters perform a uniform motion, the enstrophy is the angular velocity squared and, in this approximation, its order of magnitude is compatible with the data shown in figs. 17 and 18.
In order to measure the mean angular velocity of the rotating clusters and the orientations of the dumbbells inside each cluster, it is necessary to distinguish between the particles belonging to a cluster, characterised by high-density, and the free particles moving in low-density regions. To perform this task, we used a density-based clustering algorithm, DBSCAN (Density-Based Spatial Clustering of Application with Noise) [103], that gathers points closely packed together and that marks as noise points lying in low-density regions. We associate a particle to a cluster when it has at least 10 nearby particles within a circle of radius $2\sigma_d$ around its centre, and in this way we are able to locate all the cluster aggregates (and the particles contained in each of them).

Once the clusters are identified, we first associate a radius to them, by calculating the radius of a disk with the same moment of inertia. Then we calculate the angular moment with respect to the cluster’s centre of mass. In this way, we easily get an estimate of the angular speed of the rotating clusters around their centre of mass. In fig. 19 we show the results obtained at $Pe = 100, 200$ along the 25%-75% curve.

- The angular velocity scales with cluster size as $1/R$, as found in [48], but we underline that in the case presented here we considered aggregates that spontaneously formed during the motility-induced phase separation, and are not instead created artificially as in [48].

If we define the orientation of each dumbbell inside a cluster as the angle between its head-to-tail vector and the vector pointing from the centre of mass of the dumbbell to the centre of the cluster, we can give a quantitative measure of the different patterns exhibited by the polarisation field in fig. 9 and also understand how, at high $Pe$ values, a net torque can favour the rotation of clusters. In fig. 20 the pdfs of the dumbbells orientations are shown for systems at $Pe = 40, 100, 200$. The distributions are asymmetric and present a non-zero average angle which grows with $Pe$, matching what observed in the polarisation field, see fig. 9.

Therefore, our numerical results for high values of the active force are in perfect agreement with the prediction in [48], where the simulations were initialised with a high-density cluster formed placing the dumbbells in spirals at given orientations. At low $Pe$ numbers the dynamics looks different. When the dumbbells point along the radius of the cluster, towards its centre, or when their directions seem disordered and the polarisation field does not show a spiralling pattern, no rotation is observed for the aggregates.

5.4 Mean-square displacements

The positional mean-square displacement of the centre of mass of the single dumbbell, eq. (7), generalises to

$$\Delta^2_{\text{cm}}(\Delta t) = \frac{1}{N} \sum_{i=1}^{N} (\mathbf{r}_{\text{cm}}(t + \Delta t) - \mathbf{r}_{\text{cm}}(t))^2$$  \hspace{1cm} (20)$$

and the rotational one, eq. (8), to

$$\Delta^2_\theta(\Delta t) = \frac{1}{N} \sum_{i=1}^{N} ((\theta_i(t + \Delta t) - \theta_i(t))^2)$$  \hspace{1cm} (21)$$
in a system with $N$ dumbbells. As we will discuss below, these global measurements can lead to confusing or even misleading results in heterogeneous cases with phase co-existence. When this arises, a distinction between particles in the dense and liquid phases may be necessary to understand the behaviour of these quantities and, more
generally, of the system as a whole. This can be done, for example, by following the probability distribution functions (pdf) of the displacement along a chosen direction, say $x$, at a certain time-lag $\Delta t$:

$$p(\Delta x_{\text{cm}}, \Delta t) = \frac{1}{N} \sum_{i=1}^{N} \delta(\Delta x_{\text{cm}} - \Delta x_{\text{cm},i}(\Delta t)), \quad (22)$$

$$p(\Delta \theta, \Delta t) = \frac{1}{N} \sum_{i=1}^{N} \delta(\Delta \theta - \Delta \theta_{i}(\Delta t)), \quad (23)$$

for a single run, and their average over many repetitions of the dynamics under the same conditions that will naturally lead to smoother functions.

The finite density effects on the dynamics, for relatively low Pe values, that is to say, mostly in the liquid phase, were evaluated in [48, 49]. The four dynamic regimes in the centre-of-mass MSD of a single dumbbell, described in sect. 3, are preserved for sufficiently low densities. The translational diffusion constant in the ultimate diffusive regime, $D_A$, and rotational diffusion constant, $D_R$, once normalised by the thermal energy scale, $k_B T$, depend only on $\phi$ and Pe. $D_A$ diminishes with increasing density while, for sufficiently high Pe, $D_R$ first increases and next decreases with $\phi$. Both are enhanced by Pe and the functional dependence on this parameter is more involved than just quadratic. However,

- at sufficiently large packing fraction the density effects are stronger and a qualitative change in the properties of the MSDs are found, as can be seen in figs. 21 and 22.

The liquid like behaviour, with the four regimes recalled in sect. 3, is progressively modified by the increasing density. For the globally denser system, $\phi = 0.7$, an anomaly in the first diffusive and second ballistic regimes in the centre-of-mass MSD is clear, see fig. 21. The rotational properties are modified even more, with the rotational MSD being strongly affected already at $\phi \approx 0.35$ and the rotational diffusion constant $D_R$ being enhanced by density beyond this value of $\phi$. Another view on the global motion of the system is given by the comparison between the centre-of-mass diffusion constant and the rotational diffusion constant, both measured in the last time-delay regime. This is shown in fig. 23 in the form of the $\phi$ dependence of the ratio $D_A/D_R$ for different Pe numbers, ranging from Pe = 1 to Pe = 50. (We excluded from this figure the case $\phi = 0.7$ for which the last diffusive regime does not establish in the time window of the simulation.) The maximal value is obtained at small $\phi$ and large Pe and it is around 30. At large Pe, say Pe $\leq 50$ the ratio weakly decreases with $\phi$ while at small Pe the trend is the opposite. Observe that, if we consider the typical values of translational and rotational diffusion coefficients for non-tumbling bacterial colonies [23,102], the experimental ratio $D_A/D_R$ appears slightly higher than our results.

The qualitative change in the global dynamics is a consequence of the co-existence of dumbbells that move very differently in the sample, a fact that is confirmed by the
analysis of the probability distributions of the individual dumbbell centre-of-mass and angular displacements displayed in fig. 24 and fig. 25. These figures, constructed for a fixed value of time delay $\Delta t = t_\alpha/2$ (the III regime in the single-molecule limit), present in fact wide tails at high density $\phi = 0.7$, strongly deviating from the Gaussian behaviour.

We can check if the presence of wide tails is a general feature across different Pe inside the co-existence region, by following the distributions of individual centre-of-mass displacements on the 50%-50% curve in the phase diagram, see fig. 26. Indeed, one finds a consistent behaviour at all Pe and even at the longest time delays considered, $\Delta t > t_\alpha = 100$ (in the IV regime). This is also true on the 25%-75% line (not shown).

In order to interpret the wide tails we can measure separately the dumbbells displacements inside and outside clusters. The results of these measurements are shown for the 50%-50% line in fig. 27, where distributions of the modulus of the displacements are plotted. With continuous black lines we plot the data for all beads, with dashed orange lines the data for dumbbells that are in aggregated regions during the whole interval $\Delta t$ and, finally, with dotted red lines the data for dumbbells that are in liquid regions during the whole interval $\Delta t$ (dumbbells that change character in between the two measuring times are excluded from the statistics). We established that a particle belongs to a cluster whenever the averaged hexatic modulus on a disk of radius $R = 10 \sigma_d$ centred around the particle itself is greater than 0.75.

- The global displacement distributions are characterised by two separate peaks at all Pe, that appear to be directly linked to the separate distributions of displacements for particle inside and outside clusters, which are instead single-peaked. The peak value for each separate distribution increases monotonically with Pe, confirming the absence of any discontinuity in the region of phase coexistence for all Pe [53].

5.5 The effective temperature

Figure 28 shows the global effective temperature defined in eq. (11) as a function of the Pe number (scale above) or the active force (scale below) in a system at $T = 0.05$ with around $N = 10^3$ dumbbells [49], leading to the various global densities given in the key. The combined dependence on density and Pe changes for Pe values beyond Pe = 5, approximately, suggesting that a finer analysis of the behaviour of the individual dumbbells is also needed in this case. Indeed, the (very) naive expectation that higher Pe means more agitation and hence a higher effective temperature coincides with the measurements of $T_{\text{eff}}$ for Pe ≤ 5, approximately.

- Beyond Pe ≈ 5 the trend reverses and $T_{\text{eff}}$ diminishes with increasing Pe for $\phi > 0.2$, say. These are roughly parameters where one can expect some aggregates start appearing. No such inversion was found in the study of the FDR $T_{\text{eff}}$ in the homogeneous atomic and molecular active liquids studied in [94, 104, 105] nor in the model used in [99].

Indeed, one can expect strong differences in the value of the effective temperature measured for dumbbells that are and remain within a cluster in the period $\Delta t$ over which the measurement is performed, and those that are free to move as a fluid during the same period. Of course, there are dumbbells that are free at the first measuring time and get trapped at the later one or, vice versa, are trapped and then let free during the internal $\Delta t$. A careful analysis should be carried out to understand the effect of these strong heterogeneities (see, e.g. the discussion in [106] for study of fluctuations in glassy systems and how these may affect the effective temperature defined from the fluctuation dissipation relation).

Certain confusion reigns in the active matter field concerning the identification of an effective temperature.

We first recall that the $T_{\text{eff}}$ defined from the FDR satisfies many thermodynamic properties in glassy systems,
Fig. 26. Distribution of individual centre-of-mass displacements at $Pe = 0, 10, 40$ in the 50%-50% line at $\Delta t = 2500, 5000, 7500$, with $\Delta t > t_a = 100$ in the IV regime. Log-lin scale is used and the pdfs are normalised using the mean-square displacement of the raw histograms.

Fig. 27. Separation of displacement modulus (black) into the cluster contribution (red) and liquid part (orange), at $Pe = 0, 10, 40$ in the 50%-50% at $\Delta t = 2500$. Log-log scale is used and the pdfs are normalised using the mean-square displacement of the raw histograms.

Fig. 28. The global effective temperature as a function of the $Pe$ number (scale above) or the active force (scale below) in a system at $T = 0.05$ with various global densities given in the key. See the text for a discussion of the change in trend at around $Pe = 5$.

at least for simple models that have been studied both analytically and numerically, and it is therefore a good candidate to play a similar role in other non-equilibrium cases.

The effective temperature derived form the fluctuation-dissipation relation (FDR) does not represent an instantaneous configuration of the system. Therefore, it does not make sense to extract $T_{\text{eff}}$ from the comparison between a non-equilibrium configuration and an equilibrium one at $T_{\text{eff}}$ as done in [63]. This fact is well-known in the field of glasses, where it was understood that the relaxation of the system at different timescales has to be taken seriously in the measurement and interpretation of the effective temperature. (Trying a comparison between a glassy and an equilibrium configuration would also lead to meaningless statements in this field.) Instead, the effective temperature defined from the FDR is a dynamic concept, obtained from the comparison of configurations visited at (very) different times with and without perturbation. It is also important to stress that $T_{\text{eff}}$ is not proportional to the variance of an effective thermal noise in a Langevin equation for a reduced system obtained after integrating away some variables, due to the fact that the interactions remaining in the reduced system also play a role in the value that $T_{\text{eff}}$ will take. Typically, the (mean-field) single-variable Langevin equations that one derives from such a reduction have memory, are non-linear and the effective noise correlation depends also on the correlation functions of the selected variables of the reduced system (their mode-coupling version can be found in [107], for example, together with a discussion of how one should proceed to identify the effective temperature from a non-trivial stationary distribution in some complex systems).
an active Brownian model was pathological [63]. The ensuing contradictions found attached to this definition were interpreted as evidence against the existence of an effective temperature for active matter. In our opinion, these contradictions are the simple consequence of having used an erroneous definition, one that is already known to fail in other out of equilibrium systems.

Under which conditions $T_{\text{eff}}$ extracted from the FDR may have a thermodynamic interpretation also in active matter is an important issue. A quite natural condition seems to be that the out of equilibrium dynamics should approach equilibrium either in a long-time limit or by taking a parameter that controls the detailed balance breakdown to zero (e.g., in sheared liquids, taking the shear to vanish). Models that do not have such an equilibrium limit, like the trap model for glasses or various unbounded diffusion problems, have $T_{\text{eff}}$ extracted from the FDRs with numerous anomalies [108]. One can then infer that models of active matter with no equilibrium limit should also have trouble in having a bona fide effective temperature. The model we studied in this paper has a natural passive limit that consists in simply setting the active forces to zero. The model in which particles are propelled by persistent noises studied in, e.g. [70], also allows for equilibrium in the limit of vanishing persistence time. Consequently, these two families of models are good candidates to have meaningful effective temperatures. This is not the case of all models in the literature.

In particular, it is important to evaluate whether the effective temperature found via the use of the (time-delayed) FDR for different observables evolving in the same regime take the same values. This question was answered positively, for some observables, in [94, 104, 105] where very simple active models for particles and polymers were used. The results in [99], however, seem to point in a different direction in a model of self-propelled hard disks. In this paper, a length-dependent effective temperature in the dilute and liquid regimes was found, while these different values are locked to a single unique one in the glassy (high-density) regime. Whether partial equilibriations between different degrees of freedom occur in active systems might then depend on the model and/or on the parameters and associated dynamic regime.

Another property of the effective temperature defined from the FDR in glassy systems is that it is related to the configurational entropy or complexity, as found in mean-field solvable models and also in a number of numerical simulations. A very interesting investigation along these lines in the context of active matter appeared recently in [109].

The effective temperature $T_{\text{eff}}$ should be measurable with a thermometer. The idea of using (tuned) tracer particles as thermometers is described in [93] where references to papers where these tests have been performed in glassy systems can be found. In the context of active matter, following [110], a very recent preprint explores this issue [111]: the kinetic energy of a passive tracer particle is shown to coincide with the $T_{\text{eff}}$ measured in the zero frequency (or long time delay) limit.

Finally, we simply want to mention that several papers in the active matter literature use the FDR to estimate the effective temperature of the system both analytically [112–114] and experimentally [1, 115–117].

6 Discussion

In this work we revisited the structural and dynamical properties of an ensemble of active dumbbells in repulsive interaction through a quite hard potential. With the study of new observables, we confirmed the coexistence between liquid and hexatic order in a narrow interval of packing fraction densities of the passive system, that progressively extends into the Pe > 0 part of the phase diagram, with no discontinuity.

We quantified the separation in dilute and dense regions, the latter with hexatic order, by looking at the probability distribution functions of local densities and local hexatic order and correlating the two. The distribution of the centre-of-mass MSD also showed the existence of two distinct populations of dumbbells.

As new results, we found that the mobility of the dumbbells confined to clusters is highly suppressed. For strong enough Pe, say Pe ≥ 50, we showed that the spontaneously formed clusters turn around their centre of mass with an angular velocity that is proportional to the inverse of their radii. The poly-crystalline nature of the clusters, with respect to the hexatic order, does not seem to play a major role in their rotational properties. Instead, the orientation of the dumbbells inside the clusters is, indeed, important, as a certain amount of disorder is needed to make them turn. The entrophy pdfs allowed us to quantify the rotating properties of the individual clusters.

The possible influence of the systems’ heterogeneity on the effective temperature measurements was also stressed in the text. A more detailed analysis, separating contributions from dumbbells confined to clusters, to freely displacing ones in the dilute phase, would be needed to justify whether the decrease of $T_{\text{eff}}$ with increasing Pe in dense systems, is due to the macroscopic dumbbell aggregation.

We stress the fact that we used a rather hard repulsive potential with power law decay characterised by a short-distance power law decay with $2n = 64$. Recent studies by Klamser et al. [87] on active disks interacting via a repulsive potential with $n = 3$, for which the passive limit complies with the standard HNY melting scenario, point towards a different phase diagram. We shall come back to the behaviour of active Brownian disks interacting via a very hard repulsive potential in a separate publication [88].

Here and in [53] we analysed the structure and dynamics of the active molecular system. Unfortunately, we cannot utilise an equation of state for active dumbbells since no unambiguous definition of pressure exists for these systems. Finding such a definition remains a very interesting open problem.
All our studies concern mono-disperse systems where structural arrest of glassy type is not expected. Recent work by Mandal et al. focuses on the effects of poly-dispersity and how super-cooled liquid facts and glassiness [118–120] manifest in active dumbbell systems. We refer the reader to these references for further details [121,122].

Simulations ran on IBM Nextscale GALILEO at CINECA (Project INF16-fieldturb) under CINECA-INFN agreement and at Bari ReCaS e-Infrastructure funded by MIUR through PON Research and Competitiveness 2007-2013 Call 254 Action LI. GG acknowledges MIUR for funding (PRIN 2012NNRKAF). LFC is a member of Institut Universitaire de France, and thanks the KITP University of Santa Barbara for hospitality during part of the preparation of this work and L. Berthier, P. Choudhuri, C. Dasgupta, M. Dijkstra, J. Klamser, D. Levis, R. Mandal and J. Palacci for very useful discussions.

**Author contribution statement**

All authors conceived the study. PD and IP performed the simulations and produced the figures. All authors commented and wrote the paper.

**References**

1. J. Palacci, C. Cottin-Bizonne, C. Ybert, L. Bocquet, Phys. Rev. Lett. 105, 088304 (2010).
2. I. Theurkauff, C. Cottin-Bizonne, J. Palacci, C. Ybert, L. Bocquet, Phys. Rev. Lett. 108, 268303 (2012).
3. A. Bricard, J.-B. Caussin, N. Desreumaux, O. Dauchot, D. Bartolo, Nature 503, 95 (2013).
4. I. Buttinoni, J. Bialke, F. Kümmel, H. Löwen, C. Bechinger, T. Speck, Phys. Rev. Lett. 110, 238301 (2013).
5. W.F. Paxton, K.C. Kistler, S.J. DeCamp, M. Heymann, T. Surrey, F. Nedelec, S. Leibler, E. Karsenti, Science 329, 431 (2012).
6. S. Thakur, R. Kapral, Phys. Rev. E 85, 056603 (2007).
7. T. Sanchez, D.T.N. Chen, J. Harder, C. Valeriani, A. Cacciuto, Soft Matter 38, 45005 (2014).
8. J.T. Siebert, J. Letz, T. Speck, P. Virnau, Soft Matter 12, 555 (2016).
9. J.A. Park et al., Nat. Mater. 14, 1040 (2015).
10. G.I. Menon, Active matter, in Rheology of Complex Fluids, edited by Abhijit P. Deshpande et al. (Springer, 2010) pp. 193–218.
11. S. Ramaswamy, Annu. Rev. Condens. Matter Phys. 1, 323 (2010).
12. M.C. Marchetti, J.F. Joanny, S. Ramaswamy, T.B. Liverpool, J. Prost, M. Rao, R.A. Simha, Rev. Mod. Phys. 85, 1143 (2013).
13. M.E. Cates, Rep. Prog. Phys. 75, 042601 (2012).
14. D.A. Fletcher, P.L. Geissler, Ann. Rev. Phys. Chem. 60, 469 (2009).
117. E. Ben-Isaac, Y.-K. Park, G. Popescu, F.L. Brown, N.S. Gov, Y. Shokef, Phys. Rev. Lett. 106, 238103 (2011).
118. S.-H. Chong, A.J. Moreno, F. Sciortino, W. Kob, Phys. Rev. Lett. 94, 215701 (2005).
119. A.J. Moreno, S.-H. Chong, W. Kob, F. Sciortino, J. Chem. Phys. 123, 204505 (2005).
120. S.-H. Chong, W. Kob, Phys. Rev. Lett. 102, 025702 (2009).
121. R. Mandal, P.J. Bhuyan, P. Chaudhuri, M. Rao, C. Dasgupta, Phys. Rev. E 96, 2 (2017).
122. S.K. Nandi, R. Mandal, P.J. Bhuyan, C. Dasgupta, M. Rao, N.S. Gov, Proc. Natl. Acad. Sci. U.S.A. 115, 7688 (2018).