



Cite this: *Soft Matter*, 2021,  
17, 10447

Received 3rd September 2021,  
Accepted 12th October 2021

DOI: 10.1039/d1sm01276b

[rsc.li/soft-matter-journal](http://rsc.li/soft-matter-journal)

# Turbulence-induced clustering in compressible active fluids†

Vasco M. Worlitzer,<sup>a</sup> Gil Ariel,<sup>b</sup> Avraham Be'er,<sup>cd</sup> Holger Stark,<sup>e</sup> Markus Bär<sup>a</sup>  
and Sebastian Heidenreich<sup>a</sup>

We study a novel phase of active polar fluids, which is characterized by the continuous creation and destruction of dense clusters due to self-sustained turbulence. This state arises due to the interplay between self-advection of the aligned swimmers and their defect topology. The typical cluster size is determined by the characteristic vortex size. Our results are obtained by investigating a continuum model of compressible polar active fluids, which incorporates typical experimental observations in bacterial suspensions, in particular a non-monotone dependence of speed on density.

## 1 Introduction

Active matter has become a central topic of contemporary physics with applications ranging from collective motion of animals and cells<sup>1</sup> to new forms of soft matter<sup>2</sup> and the dynamics of cells and tissues.<sup>3</sup> Among active systems, bacterial colonies provide an intriguing and experimentally well-controllable class of systems. Many aspects of their behavior can be understood by analogies to fundamental non-equilibrium statistical physics.<sup>4</sup> Several bacterial species provide showcases for collective motion and ordered patterns, that can be described by simple models of self-propelled particles dominated by alignment<sup>5</sup> or repulsive interactions and particle shape.<sup>6</sup> Suspensions of swimming bacteria may be governed, in addition, by longer range hydrodynamic interactions, which can lead to a new dynamic state referred to as meso-scale turbulence. This dynamical state is characterized by the presence of self-sustained vortices and jets and is reminiscent of turbulence.<sup>7–12</sup> In contrast to inertial turbulence there exists a characteristic vortex size,<sup>7,9,13,14</sup> which can be explained by the competition between alignment and hydrodynamic

interactions.<sup>14–17</sup> However, models of meso-scale turbulence commonly assume incompressibility,<sup>14–17</sup> and are therefore not able to describe clustering phenomena.

A common description of compressible active fluids is based on self-propulsion and purely repulsive interactions showing intriguing non-equilibrium phenomena like motility-induced phase separation (MIPS).<sup>18–20</sup> For this phenomenon it is crucial that the self-propulsion speed of the particle slows down with increasing density.<sup>19</sup> In this context, experiments of bacterial suspensions exhibit a puzzling situation: On the one hand, the speed of bacteria increases with density for low and intermediate densities<sup>21</sup> and on the other hand, clusters or density variations are commonly present and often emerge from dilute initial conditions.<sup>22</sup> This behavior has not been explained so far. While there exist approaches to model bacterial turbulence incorporating density variations,<sup>23,24</sup> those models do not feature a characteristic length scale observed experimentally.<sup>7,9,13,14</sup>

Based on our previous investigations,<sup>25</sup> where a continuum model was proposed that combines meso-scale turbulence with MIPS, we extend here this concept and incorporate typical experimental observations in bacterial suspensions by assuming a non-monotone dependence of speed on density, see Fig. 1(a). Our previous study<sup>25</sup> considered the dynamics in a simplistic regime (monotone decrease of speed with density) and identified several interesting phases, including turbulence-induced clustering, which is characterized by the continued formation, reshaping and fracture of dense clusters. Here, we focus on this novel phase in a more realistic setup, analyze its properties, their dependence on different physical parameters and scope within parameter space. In particular, we show that the dynamic arises due to self-advection of the aligned swimmers and point out that the nucleation of clusters is organized by the defect topology in the polar order field of the swimmers. In general, topological defects often dominate the dynamics of

<sup>a</sup> Department of Mathematical Modelling and Data Analysis, Physikalisch-Technische, Bundesanstalt Braunschweig und Berlin, Abbestrasse 2-12, D-10587 Berlin, Germany. E-mail: [vasco.worlitzer@ptb.de](mailto:vasco.worlitzer@ptb.de), [sebastian.heidenreich@ptb.de](mailto:sebastian.heidenreich@ptb.de)

<sup>b</sup> Department of Mathematics, Bar-Ilan University, 52900 Ramat Gan, Israel

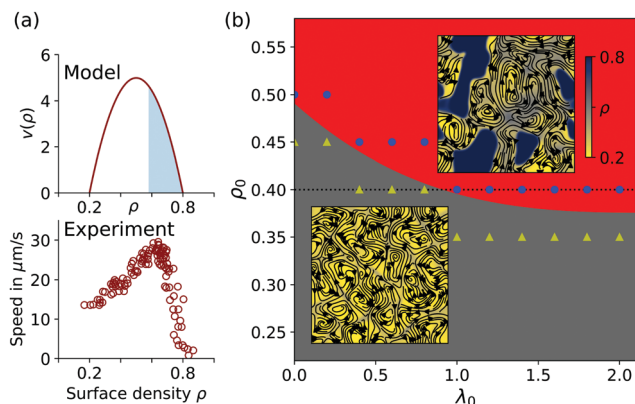
<sup>c</sup> Zuckerberg Institute for Water Research, The Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Sede Boqer Campus, 84990 Midreshet Ben-Gurion, Israel

<sup>d</sup> Department of Physics, Ben-Gurion University of the Negev, 84105 Beer Sheva, Israel

<sup>e</sup> Institute of Theoretical Physics, Technische Universität Berlin, Hardenbergstrasse 36, D-10623 Berlin, Germany

† Electronic supplementary information (ESI) available: Movie showing the time evolution of the density and polarization density obtained by solving our model of a compressible active fluid. See DOI: 10.1039/d1sm01276b





**Fig. 1** (a) Density-dependent speed  $v(\rho)$  in our model (see eqn (2)) and obtained from experiments (reproduced from ref. 21). Linearly unstable region due to MIPS is colored in blue, see Section A.1. (b) Phase diagram for  $\Gamma_2 = -\Gamma_0 = 1$  while varying  $\rho_0$  and  $\lambda_0$  exhibiting meso-scale turbulence with homogeneous density (MST, grey) and turbulence-induced clustering (TIC, red). Symbols represent simulation data, from which the phase boundary is estimated (yellow triangles = MST, blue circles = TIC). Insets: Typical snapshots of the respective states. Streamlines calculated from  $\mathbf{v}$ , while the background color is obtained from  $\rho$  (see colorbar). For better visualization no streamlines are plotted for areas with  $\rho = \rho_{\max}$ , i.e. inside dense clusters.

active biological matter: they govern cell death and extrusion in epithelial tissues,<sup>26</sup> control the dynamics in neural cell cultures<sup>27</sup> and promote new layer formation in dense bacterial colonies.<sup>28</sup> Moreover, we show that the average cluster size depends linearly on the characteristic vortex size. Overall, our results point out an alternative way how clusters can emerge in systems far from equilibrium.

## 2 Model

The proposed model for compressible active fluids consists of coupled equations for the density  $\rho$  and the polarization density  $\mathbf{p}$ , quantifying polar order, see<sup>25</sup>

$$\partial_t \rho = -\nabla \cdot [v(\rho)\mathbf{p}] + D\Delta\rho, \quad (1a)$$

$$\begin{aligned} \partial_t \mathbf{p} + \lambda_0 (\mathbf{p} \cdot \nabla) \mathbf{p} = & -\frac{1}{2} \nabla [v(\rho)\rho] - [A(\rho) + C|\mathbf{p}|^2] \mathbf{p} \\ & + \Gamma_0 \Delta \mathbf{p} - \Gamma_2 \Delta^2 \mathbf{p}. \end{aligned} \quad (1b)$$

The evolution of the density is governed by the continuity equation eqn (1a), where the current density consists of a diffusive part and a term stemming from the directed motion of the particles. Eqn (1b) is reminiscent of the equation to model meso-scale turbulence.<sup>16,17</sup> The terms arise due to self-propulsion, alignment and hydrodynamic interactions. The term involving the gradient of  $\rho$  couples the polarization to the evolution of the density. The velocity field  $\mathbf{v}(\mathbf{x}, t)$  is obtained from  $\mathbf{v} = v(\rho)\mathbf{p}/\rho$ . Consequently, the polarization is parallel to the velocity. This is a restriction of our model, as experiments have shown that cells in bacterial suspensions are not fully aligned with the flow.<sup>4,29</sup> A derivation of a continuum model of meso-scale turbulence that treats the polarization and solvent

flow fields separately, but assumes a constant density, can be found in ref. 17. We model the density-dependent speed  $v(\rho)$  with the quadratic function

$$v(\rho) = -c[\rho - (\rho_{\max} + \rho_{\min})/2]^2 + v_0, \quad (2)$$

where  $c = 4v_0/(\rho_{\max} - \rho_{\min})^2$ . This form is motivated by experimental studies of bacterial suspensions,<sup>21,30</sup> see Fig. 1(a). We remark that the behaviour of bacterial suspensions at the onset of meso-scale turbulence, i.e. in the limit  $\rho \rightarrow \rho_{\min}$ , is poorly understood. Hence, we do not expect to accurately capture the dynamics with our model in this limit.

Furthermore, density  $\rho$  and polarization  $\mathbf{p}$  couple to each other *via* the polar alignment strength  $A(\rho)$  in eqn (1b). Inspired by Landau theory, one often assumes  $A(\rho) = A_0(\rho_c - \rho)$  to model the polar-isotropic phase transition. For simplicity, we assume that the onset of collective motion coincides with the polar-isotropic phase transition, i.e.  $\rho_c = \rho_{\min}$  and hence  $A(\rho) < 0$ .

In our numerical simulations, we choose  $\rho_{\min} = 0.2$  and  $\rho_{\max} = 0.8$  in line with experimental results.<sup>21</sup> Moreover, we fix  $A_0 = 0.3$ ,  $C = 0.5$ ,  $D = 5$  and  $v_0 = 5$ . We vary the mean density  $\rho_0$ , and use  $\Gamma_0$  and  $\Gamma_2$  to control the characteristic vortex size. Finally, the strength of self-advection of the polarization is tuned by  $\lambda_0$  (for details on the numerics see Section B.1).

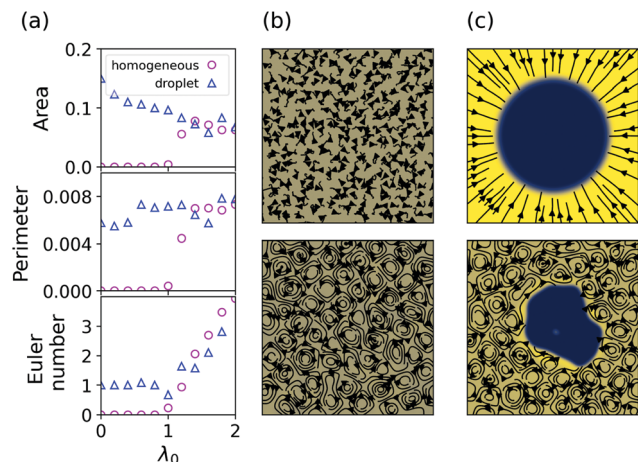
## 3 Phase diagram

We initiate simulations with a homogeneous density profile and choose  $\Gamma_0 = -1$  and  $\Gamma_2 = 1$ , such that the system is in a meso-scale turbulent state.<sup>25</sup> This state is characterized by vortices with a characteristic length scale as well as the presence of dynamical jets.<sup>14,15,31</sup> By varying the mean density  $\rho_0$  and the strength of self-advection  $\lambda_0$ , we sketch a phase diagram, see Fig. 1(b). Note that the mean density  $\rho_0$  is varied outside the spinodal regime of MIPS, i.e. between  $\rho_{\min}$  and  $\rho_s \approx 0.58$ , see Fig. 1(a) and Section A.5.

Meso-scale turbulence (MST) or vortex lattices accompanied by an almost constant density  $\rho \approx \rho_0$  are observed for low values of  $\rho_0$  and  $\lambda_0$ . For high values of  $\rho_0$  and  $\lambda_0$ , we discover a novel phase, which we refer to as turbulence-induced clustering (TIC). In this phase dense clusters with  $\rho = \rho_{\max}$ , hence  $\mathbf{v} = 0$ , emerge throughout the simulation domain, while dilute areas are still characterized by vortices. The dense clusters are not fixed in space but rather get reshaped and stretched. Furthermore, clusters commonly fracture and smaller parts of the fracture might vanish. We emphasize that the emergence, reshaping and fracture of clusters continues indefinitely. That is, while a cluster might fracture in one part of the simulation domain, a new cluster might emerge at another place, see the supplemental movie. Quantitatively, the two phases can be distinguished by determining the modality of the density distribution, see ref. 25 for details.

To quantify the dynamical evolution of dense clusters, we use Minkowski functionals. They provide a mathematical framework to completely characterize the morphology of patterns.<sup>32,33</sup> In two dimensions, the Minkowski functionals





**Fig. 2** (a) Time-averaged Minkowski functionals when varying  $\lambda_0$  along the dotted line in Fig. 1(b), *i.e.* for  $\rho_0 = 0.4$ . Circles represent runs started from a homogeneous density profile, while triangles represent runs started from an initial dense droplet. Surface area and perimeter are rescaled by the domain size to yield fractions. (b and c) Snapshots of the initial condition (upper) and at  $t = 150$  (end of the simulation, lower) for a homogeneous density profile (b) and an initial droplet (c) in the MST phase, *i.e.* for  $\rho_0 = 0.4$  and  $\lambda_0 = 0$ . Respective snapshots of the TIC phase are provided in Section A.2. In all subfigures  $\Gamma_2 = -\Gamma_0 = 1$ .

coincide with the area  $A$ , perimeter  $U$  and Euler number  $\chi$ . The Euler number can be calculated by counting the number of connected finite domains, in our case the dense clusters, and subtracting the number of holes. Hence, the Euler number is a topological quantity related to the number of clusters. For our purposes, Minkowski functionals can robustly quantify whether there is coarsening, and they allow to detect geometrical changes. These insights cannot be reliably gained from the more commonly studied quantities such as the number of clusters and the average cluster size. For example, if the dense phase is more abundant than the dilute phase, the dense phase forms a connected domain interspersed with holes, see Section A.5 for an illustration. If the system coarsens, these holes vanish over time, eventually leaving only one large hole. Such a coarsening process cannot be quantified by the number of clusters or the size of the cluster as both do not change over time (up to fluctuations of course). In contrast, Minkowski functionals can discriminate whether the systems coarsens or not. If there is coarsening, the Euler number would increase, eventually reaching a value of  $-1$ , while the perimeter would decrease.<sup>33</sup> Moreover, Minkowski functionals can quantify geometrical changes. For example, a deformation of a spherical cluster leads to an increase of the perimeter while the area remains constant. Such observations are not possible from the number of clusters and cluster sizes. Besides, the computation of Minkowski functionals is straightforward and fast as one only must count pixels, see Section B.3. In contrast, the determination of cluster properties in a system with periodic boundary conditions is much more tedious and time consuming, see Section B.2. Therefore, Minkowski functionals allow for a computationally inexpensive and precise analysis of the geometrical and topological features of our system. For details on Minkowski functionals refer to Section B.3.

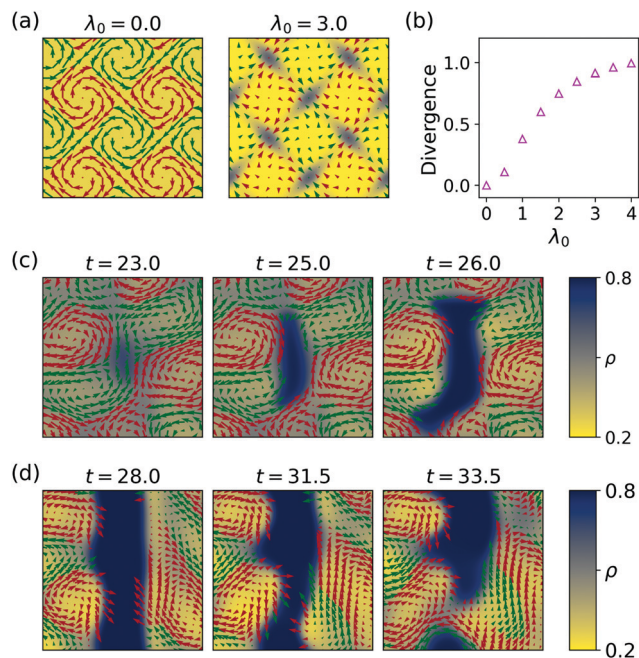
Analyzing the temporal evolution of the Minkowski functionals in the TIC phase reveals that  $A$ ,  $U$  and  $\chi$  fluctuate around a non-zero mean value, see Section A.5, due to continuous creation and destruction of clusters, which indicates a statistical steady state. We rule out finite-size effects by varying the linear size  $L$  of the system, see Section A.3. To further test the robustness of our results, we initiate simulations with a single dense circular domain and vary the self-advection strength  $\lambda_0$  along the dotted line in Fig. 1(b). For low values of  $\lambda_0$  (*i.e.* in the MST phase) the mean values of the Minkowski functionals are well separated (Fig. 2(a)): an initial circular cluster preserves its Euler number 1 and maintains positive area and perimeter, while the Minkowski functionals are zero for the homogeneous density profile. Hence, the time evolution in the MST phase depends heavily on the initial conditions, see Fig. 2(b) and (c). For high values of  $\lambda_0$  (*i.e.* in the TIC phase) differences in the time-averaged Minkowski functionals cannot be attributed to the initial conditions. After a transient, the dynamics and statistical properties of the system for different initial conditions are indistinguishable, see Fig. 2(a) and Section A.2. From this observation we can draw two conclusions about the TIC phase: (1) the novel statistical steady state is independent of the initial conditions and (2) the completely phase separated state is unstable. The second conclusion in turn shows that another process than MIPS must be responsible for the TIC phase, as MIPS predicts a complete phase separation on long time scales. Our results indicate that self-advection through the non-linear term  $\lambda_0(\mathbf{p} \cdot \nabla)\mathbf{p}$  of the polarization triggers the nucleation of clusters as well as the fracture of already existing domains.

## 4 Dynamical creation and destruction of clusters due to meso-scale turbulence

The nucleation of clusters due to meso-scale turbulence can be understood by the compressibility of the system and the local defect topology. We illustrate this by initiating the system with a vortex lattice at constant density  $\rho = \rho_0$ . In this situation, the evolution of the density eqn (1a) reduces to  $\partial_t \rho = -v(\rho_0) \nabla \cdot \mathbf{p}$ . Hence, density changes are only possible if sinks ( $\nabla \cdot \mathbf{p} < 0$ ) or sources ( $\nabla \cdot \mathbf{p} > 0$ ) of the polarization are present. We remark that this argument is only valid at the early stages, when density gradients are negligible. We detect sources and sinks of the polarization by computing the integral  $1/L^2 \int (\nabla \cdot \mathbf{p})^2 dx$ . Fig. 3(b) shows that the maximum of this integral increases with the strength of self-advection  $\lambda_0$ , leading to density gradients within the system.

Analyzing the snapshots in Fig. 3(a) shows that the density changes are organized by the defect topology of the polarization. In a vortex lattice, a topological defect with charge 1 is located at the vortex center and the point in between four adjacent vortices is a topological defect with charge  $-1$ . Moreover, the points in between four adjacent vortices are saddle-points of the polarization field. While for  $\lambda_0 = 0$  the vortex lattice is only perturbed marginally, high values of  $\lambda_0$  result in



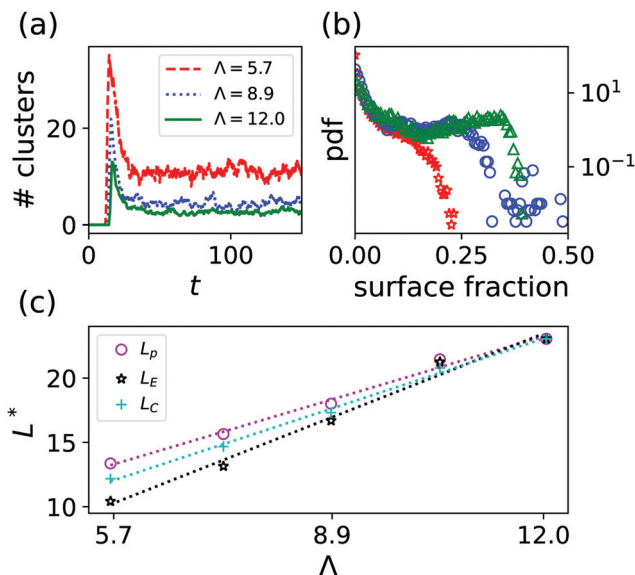


**Fig. 3** (a) Snapshots at  $t = 20$  of runs started from a vortex lattice with  $\rho_0 = 0.25$ ,  $\Gamma_0 = -2$ ,  $\Gamma_2 = 1$  for  $\lambda_0 = 0$  and  $\lambda_0 = 3$ . (b) Maximal absolute divergence, i.e.  $\max_{[0,\delta t]} 1/L^2 \int (\nabla \cdot \mathbf{p})^2 \mathrm{d}\mathbf{x}$  for the initial stage of the vortex lattice, i.e. for  $\delta t = 10$ , when varying  $\lambda_0$ . Parameters as in (a). (c) Snapshots of a cluster growing from random initial conditions with  $\rho_0 = 0.45$ ,  $\lambda_0 = 0.5$ . (d) Evolution of a flat interface between a dense cluster and a vortex array in the dilute region with  $\rho_0 = 0.45$  and  $\lambda_0 = 3$ . In (a, c and d) arrows represent the  $\mathbf{p}$  field, with arrow color indicating clockwise (red) and counterclockwise (green) rotating vortices.

the appearance of sources of the polarization at the vortices. Clearly, combining a source with a vortex results in an outward spiral. This is in line with simulation results from.<sup>34</sup> Therein, elongated pushers (like *Bacillus subtilis*) show a propensity to form outward spirals, when increasing self-advection. Density shifted away from the vortices accumulates in between four adjacent vortices (Fig. 3(a)) due to mass conservation. This is reminiscent of the results of,<sup>35</sup> where swimmers forced by a velocity field in the form of a vortex lattice leave the vortices and accumulate in between them. Hence, there is a density shift from topological defects with charge 1 to those with topological charge  $-1$ , which represent possible nucleation sites for dense clusters. Indeed, in simulations started from random initial conditions, clusters preferentially nucleate in between four adjacent vortices, see Fig. 3(c). Furthermore, initiating the system with a flat interface between a dense cluster and a vortex array in the dilute region, shows that the surrounding vortices modulate and eventually break-up the interface, see Fig. 3(d).

## 5 Cluster size

Finally, we address the mechanism which determines the average cluster size in the TIC phase. Trivially, the cluster size depends on the mean density  $\rho_0$ , see Section A.5. More



**Fig. 4** (a) Number of clusters, (b) cluster size distribution and (c) characteristic domain size  $L^*$  for different characteristic vortex sizes  $\Lambda = 2\pi\sqrt{2\Gamma_2/\Gamma_0}$ . Details for the choice of  $\Gamma_0, \Gamma_2$  can be found in Section A.6. Data is obtained for  $\rho_0 = 0.5$  by averaging over 100 time steps after a transient of  $\delta t = 50$  for 10 runs ((a) and (b)) and 5 runs (c) with identical parameters but different random initial conditions. Colors in (b) correspond to those in (a). In (c) the characteristic domain size obtained from the perimeter ( $L_p$ , magenta circles), the Euler number ( $L_E$ , black stars) and the correlation function ( $L_C$ , cyan crosses) is shown. As a visual guide linear fits for all three cases are provided as dashed lines.  $L_p$  and  $L_E$  are rescaled with the maximum of  $L_C$ .

interestingly, for fixed  $\rho_0$ , the cluster size changes with the characteristic vortex size. The characteristic vortex size can be estimated from the fastest growing mode  $k_c$  obtained from the stability analysis as  $\Lambda = 2\pi/k_c = 2\pi\sqrt{2\Gamma_2/\Gamma_0}$ , see Section A.1.

Increasing  $\Lambda$  results in a decrease in the number of clusters (Fig. 4(a)), but a higher probability to find larger clusters, which can be deduced from the cluster size distribution plotted in Fig. 4(b). Moreover, we compute a characteristic domain size  $L^*$  either from the density correlation function ( $L_C$ ), the perimeter ( $L_p$ ) or the Euler number ( $L_E$ ), see Section A.6. All three quantities increase linearly with  $\Lambda$ , see Fig. 4(c). This shows that the mean value, around which the Minkowski functionals fluctuate, is determined by the characteristic vortex size. A possible explanation is provided by our previous observation that clusters preferentially nucleate in between adjacent vortices. Increasing the vortex size leaves more space in between vortices (implies larger distance between topological defects with charge +1), while the number of possible nucleation sites reduces, resulting in fewer but larger clusters. Hence, we conjecture that the novel TIC phase can be understood as fluctuations around a vortex lattice with dense cluster situated in between vortices.

## 6 Conclusions

We presented an extension of a continuum model that combines MIPS with meso-scale turbulence by including a realistic



self-propulsion speed dependence on density. By numerical investigations, we showed that meso-scale turbulence induces dynamical clustering if incompressibility is not enforced. Independent of the system size and initial conditions, a novel dynamical state emerges, which is characterized by the continuous nucleation, deformation and destruction of clusters. These antagonistic dynamics arise due a common cause: Hydrodynamic interactions lead to self-advection of the polarization governing the dynamics. This points to the importance of hydrodynamic interactions, which are usually excluded in the literature,<sup>36,37</sup> when studying clustering phenomena of elongated self-propelled particles. Moreover, our prediction of a linear dependence of the cluster size on the characteristic vortex size can be easily examined experimentally, where these quantities are readily accessible. Hence, our model allows for a quantitative comparison with real systems exhibiting TIC.

Furthermore, the TIC phase extends to densities  $\rho_0 > \rho_s$ , *i.e.* to the regime where spinodal decomposition through MIPS is expected: Clusters form initially through a process reminiscent of spinodal decomposition, while on long time scales we observe continued emergence, reshaping and fracture of clusters, see Section A.5. Hence, turbulence-induced clustering is a generic phenomenon for compressible polar active fluids. However, our results depend on the assumption that the polarization and velocity fields are parallel. In general, the topological defects of the polarization and velocity fields do not coincide if this assumption is violated. As the topological defects play a central role in the formation of clusters, we cannot claim that turbulence-induced clustering is expected, if a more general class of compressible active fluids is considered.

Our results bear some reminiscence to passive systems quenched into the spinodal regime in the presence of turbulence. Therein, thermodynamic forces drive coarsening, while existing domains are broken up due to the turbulent motion. The competition between these effects leads to a coarsening arrest at a certain length scale.<sup>38–41</sup> However, there are two notable novel aspects of our observations: Firstly, nucleation of clusters due to inertial turbulence is not reported for passive systems,<sup>38–41</sup> where clusters emerge spontaneously as the system is quenched into the spinodal regime. Our simulations are performed outside the spinodal regime. Consequently, clustering is initiated and governed by the turbulent fluid motion in the active system. Secondly, the average cluster size in the active system is controlled by the characteristic vortex size, rather than by opposing forces. Hence, turbulence-induced clustering illustrates a novel route to pattern formation which is unique to active systems. In the context of biological systems, it points to a possible functional benefit of meso-scale turbulence, *e.g.* as a driver for aggregation of bacteria at low and intermediate densities.

## Conflicts of interest

There are no conflicts to declare.

## A Appendix

### A.1 Stability analysis

A linear stability analysis of the model eqn (1) can be performed as detailed in ref. 25. Since we set  $\rho_{\min} = \rho_c$ , the only attainable, trivial steady state is given by  $(\rho, \mathbf{p}) \equiv (\rho_0, \mathbf{p}_0)$  with  $|\mathbf{p}_0| = \sqrt{-A(\rho_0)/C}$  (global polar order). Introducing perturbations  $\mathbf{p} \rightarrow \mathbf{p}_0 + \delta\mathbf{p}$  and  $\rho \rightarrow \rho_0 + \delta\rho$  and linearizing eqn (1) close to the steady state  $(\rho, \mathbf{p}) \equiv (\rho_0, \mathbf{p}_0)$  yields the equations of motion for the perturbations

$$\begin{aligned}\partial_t \delta\rho &= -v'(\rho_0)\mathbf{p}_0 \cdot \nabla \delta\rho + D\Delta\delta\rho - v(\rho_0)\nabla \cdot \delta\mathbf{p}, \\ \partial_t \delta\mathbf{p} &= -\frac{1}{2}(v(\rho_0) + v'(\rho_0)\rho_0)\nabla \delta\rho - A'(\rho_0)\mathbf{p}_0\delta\rho \\ &\quad - \lambda_0(\mathbf{p}_0 \cdot \nabla)\delta\mathbf{p} - A(\rho_0)\delta\mathbf{p} \\ &\quad - C(2\mathbf{p}_0\mathbf{p}_0 + |\mathbf{p}_0|^2 I)\delta\mathbf{p} + \Gamma_0\Delta\delta\mathbf{p} - \Gamma_2\Delta^2\delta\mathbf{p}.\end{aligned}\quad (3)$$

Next, we expand the perturbations in Fourier modes as

$$\begin{pmatrix} \delta\rho \\ \delta\mathbf{p} \end{pmatrix} = \begin{pmatrix} \hat{\rho} \\ \hat{\mathbf{p}} \end{pmatrix} \exp(\sigma t + i\mathbf{k} \cdot \mathbf{x}), \quad (4)$$

where  $\mathbf{k} = (k_1, k_2)$ . Inserting this expression into eqn (3) leads to

$$\begin{aligned}\sigma \hat{\rho} &= -(iv'(\rho_0)\mathbf{p}_0 \cdot \mathbf{k} + D|\mathbf{k}|^2)\hat{\rho} - iv(\rho_0)\mathbf{k} \cdot \hat{\mathbf{p}}, \\ \sigma \hat{\mathbf{p}} &= -\left(\frac{1}{2}(v(\rho_0) + v'(\rho_0)\rho_0)\mathbf{k} + A'(\rho_0)\mathbf{p}_0\right)\hat{\rho} \\ &\quad - (i\lambda_0(\mathbf{p}_0 \cdot \mathbf{k}) + A(\rho_0) + C|\mathbf{p}_0|^2 + \Gamma_0|\mathbf{k}|^2 + \Gamma_2|\mathbf{k}|^4)\hat{\mathbf{p}} \\ &\quad - 2C\mathbf{p}_0\mathbf{p}_0 \cdot \hat{\mathbf{p}}.\end{aligned}\quad (5)$$

As the expression eqn (5) is not very handy, we solve the dispersion relation numerically. Fig. 5(a) shows the dispersion relation in the direction  $\mathbf{k}_\perp$  perpendicular to  $\mathbf{p}_0$  for the parameters used in Fig. 1. While  $\sigma_1$  and  $\sigma_2$  do not show any instabilities,  $\sigma_3$  exhibits a finite-wavelength instability. The critical wavenumber  $k_c$  is given by  $k_c = \sqrt{-\Gamma_0/2\Gamma_2}$ . Hence, this instability is the one expected from meso-scale turbulence.

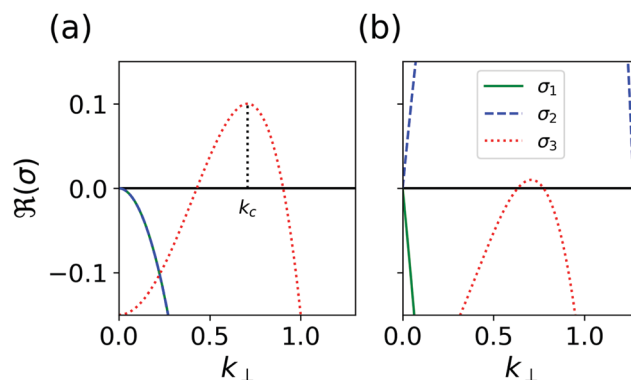


Fig. 5 Dispersion relations for the polar state in the direction  $\mathbf{k}_\perp$  perpendicular to  $\mathbf{p}_0$ . Parameters are chosen as described in the main text with  $\Gamma_0 = -\Gamma_2 = -1$  and  $\lambda_0 = 3$ . (a)  $\rho_0 = 0.45$  and (b)  $\rho_0 = 0.6$ . The position of the critical wavenumber  $k_c$  is shown by a black dashed line in (a).



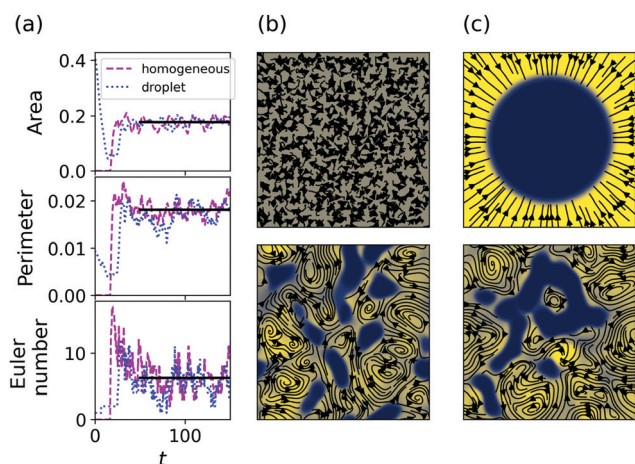
Fig. 5(b) shows the dispersion relation for  $\rho_0 = 0.6$ . In addition, to the finite-wavelength instability, a long-wavelength instability in  $\sigma_2$  appears. This instability is reminiscent of MIPS, see also.<sup>25</sup> The appearance of the long-wavelength instability marks the spinodal  $\rho_s$  of the system, *i.e.*, the point when the homogeneous density profile loses stability. The numerically computed value of the spinodal is  $\rho_s \approx 0.58$ . See also the discussion in Section A.5.

For the data shown in Fig. 4, the parameters  $\Gamma_0, \Gamma_2$  are chosen such that the amplitude  $\sigma(k_c)$  is fixed when varying  $\Lambda$ . The amplitude  $\sigma(k_c)$  can be seen as a measure for the distance to the instability. Hence, fixing  $\sigma(k_c)$  guarantees a similar energy input of the finite-wavelength instability. This is important as it is known that close to the instability,  $\sigma(k_c) \approx 0$  no meso-scale turbulence is observed, but a vortex lattice<sup>25,31</sup> instead.

## A.2 Independence of initial conditions

As mentioned in the main text, in the TIC phase the Minkowski functionals fluctuate around a non-zero mean value. This is depicted in Fig. 6(a). The mean value is independent of the initial conditions, which we illustrate by initializing the system with a dense droplet as well as a homogeneous density profile.

We used the independence of initial conditions to distinguish between the MST and TIC phases when varying the strength of self-advection  $\lambda_0$ , see Fig. 2 in the main text. However, the determination of a critical  $\lambda_0^c$  depends sensitively on the parameters and initial conditions of the system. For large values of  $\lambda_0$ , clusters nucleate immediately at the onset of meso-scale turbulence, *i.e.* the time of emergence of the first cluster is independent of  $\lambda_0$ . However, the nucleation time of the first cluster increases when reducing  $\lambda_0$  to  $\lambda_0 \approx 1$ .

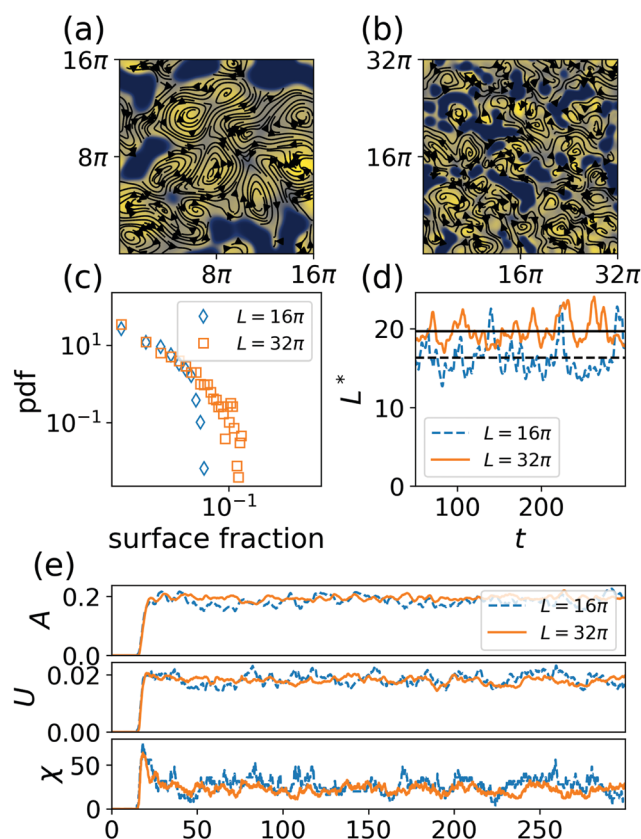


**Fig. 6** (a) Time evolution of the Minkowski functionals for  $\Gamma_2 = -\Gamma_0 = 1$ ,  $\rho_0 = 0.45$  and  $\lambda_0 = 3$  for a system initiated with a homogeneous density profile (dashed purple line) and a dense droplet (dotted blue line). The solid black line shows the average after an initial transient. Surface area and perimeter are rescaled by the domain size to yield fractions. (b and c) Snapshots of the initial condition (upper) and at  $t = 150$  (end of the simulation, lower) for a homogeneous density profile (b) and an initial droplet (c) with parameters as in (a).

Hence, we cannot rule out the nucleation of clusters due to meso-scale turbulence for smaller values of  $\lambda_0$ , as the nucleation time might lie outside our simulated time frame. Therefore, details about the sensitivity on initial conditions remain unclear.

## A.3 Finite size scaling

To check for finite size effects, we increase the size of the simulation domain. For the data provided in the main text, we use a simulation domain with linear size  $L = 16\pi$  and  $256 \times 256$  grid points. We now additionally use a simulation domain with linear size  $L = 32\pi$  and  $512 \times 512$  grid points. The results are shown in Fig. 7. The snapshots provided in Fig. 7(a) and (b) indicate that the dynamics do not differ qualitatively. This can be quantified by measuring the Minkowski functionals, which are plotted in Fig. 7(e). The average value of the rescaled area  $A$  and perimeter  $U$  agree for both simulation domain sizes as expected. Due to the additivity of Minkowski functionals, see Section B.3, one would expect the Euler number to quadruple when the linear size of the simulation domain is doubled. We measure an increase of the Euler number by a factor of 3.2, slightly below the expected value. Furthermore, the cluster size distribution (Fig. 7(c)) and characteristic domain size  $L^*$



**Fig. 7** (a) and (b) show simulation snapshots at  $t = 300$  for simulation domain sizes of  $L = 16\pi$  (a) and  $L = 32\pi$  (b). For the two different simulation domain size the cluster size distribution (c), the evolution of the characteristic domain size obtained from the density correlation function  $L^*$  (d) and the Minkowski functionals over time (e) are plotted. The Euler number  $\chi$  for  $L = 16\pi$  is multiplied by a factor of 4, which represents the expected scaling due to additivity. Parameters as in Fig. 6.

(Fig. 7(d)) show qualitatively similar behaviour for the different simulation domain sizes. The characteristic domain size is obtained from the density correlation function, see Section B.2 for details. We observe a slightly larger characteristic domain size for the larger simulation domain. This is also reflected in the cluster size distribution. Together with our observation about the Euler number, we conclude that for the larger simulation domain size there are slightly fewer but larger clusters. However, as the differences are small, we conjecture that they can be attributed to fluctuations rather than to systematic finite size effects.

#### A.4 Comparison between TIC and MIPS

As our model eqn (1) incorporates MIPS as a special case,<sup>25</sup> one might expect that the emergence of dense clusters can be explained by MIPS. Similar to equilibrium phase separation, dense clusters emerge in MIPS either through spinodal decomposition, when increasing  $\rho_0$  past the spinodal  $\rho_s$ , or *via* nucleation in the metastable regime between the spinodal and binodal ( $\rho_b < \rho_0 < \rho_s$ ), if sufficiently large perturbations are provided.<sup>42–44</sup> The spinodal  $\rho_s$  marks the point, when the homogeneous density profile becomes linearly unstable against perturbations. Hence, the position of the spinodal can be computed by a linear stability analysis, which is done in Section A.1. Since we choose  $\rho_0 < \rho_s$  in our simulations, spinodal decomposition can be ruled out. Nucleation through MIPS appears unlikely as the initial small-scale fluctuations vanish before clusters appear. This is represented by the time lag between the start of the simulation and the increase of the Euler number  $\chi$  (Fig. 6(a)). Furthermore, if clusters are nucleated through MIPS, they grow and coarsen until a completely phase separated state is reached. Related to this is the presence of hysteresis:<sup>42–44</sup> Depending on the initial conditions either a homogeneous density profile or a phase separated state is observed. As shown earlier, we do not observe hysteresis in the TIC phase and clusters fracture before reaching phase separation. Altogether, those differences indicate that another dynamical process other than MIPS is responsible for the dynamics in the TIC phase. Nevertheless, metastability of the homogeneous density profile might be a necessary condition for the emergence of dense clusters, as no clusters emerge below the estimated binodal  $\rho_b$  independent of  $\lambda_0$ , as explained in the following Section A.5.

#### A.5 Extension of the TIC phase to the spinodal regime

So far, we focused on the regime  $\rho_0 < \rho_s$ , *i.e.* on densities below the spinodal of MIPS. We now vary  $\rho_0$  between  $\rho_{\min}$  and  $\rho_{\max}$  while fixing  $\lambda_0 = 3$ . The resulting phase diagram is pictured in Fig. 8(a).

For low mean densities (see  $\rho_0 = 0.3$ ), we observe meso-scale turbulence without the presence of clusters, *i.e.* the MST phase. This can be quantified by the evolution of the Minkowski functionals plotted in Fig. 8(b). Again, we initiated the system either with a homogeneous density profile (solid lines) or with a dense droplet (dashed lines).

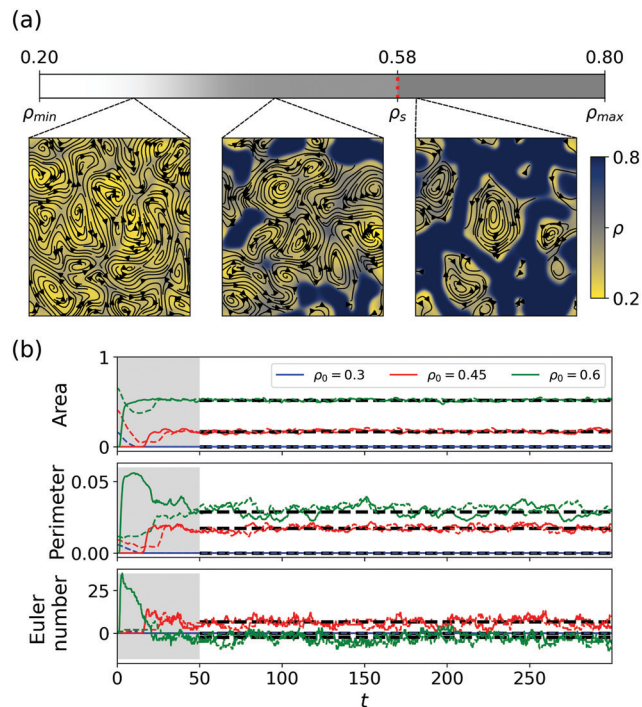


Fig. 8 Dynamics when varying  $\rho_0$  for  $\Gamma_2 = -\Gamma_0 = 1$  and  $\lambda_0 = 3$ . (a) Snapshots with mean densities  $\rho_0 = 0.3$ ,  $\rho_0 = 0.45$  and  $\rho_0 = 0.6$  at  $t = 300$  (from left to right). (b) Evolution of the Minkowski functionals for the three cases shown in (a). Simulations are either started from a homogeneous density (full lines) or a dense droplet (dashed lines). Initial transient is marked in gray, while the average after the transient is provided by black dashed lines.

Increasing the mean density, we observe the nucleation of dense clusters due to meso-scale turbulence, *i.e.* the TIC phase (see  $\rho_0 = 0.45 < \rho_s$ ). We speculate that the transition between MST and TIC coincides with the binodal  $\rho_b$ . In equilibrium systems, the binodal can be obtained by a common tangent construction of the free energy functional.<sup>45</sup> While such a common tangent construction was also applied to non-equilibrium systems *via* a mapping to an effective free energy,<sup>45</sup> we are not aware of such a mapping in our system. Numerically, the position of the binodal can be estimated by the values of the coexisting densities. Estimating the lower coexisting density from the peak of the density distribution yields  $\rho_b \approx 0.35$ . We note though that the distribution around this peak is fairly wide. Hence, it can be debated whether there exists a lower coexisting density at all in our systems. Nevertheless, the computed value  $\rho_b \approx 0.35$  agrees reasonably well with the observed onset of clustering. A more systematic study might be in order to support our hypothesis. Hence, we do not provide a transition point in Fig. 8(a) but a transition zone.

Interestingly, increasing the mean density above the spinodal (see  $\rho_0 = 0.6 > \rho_s$ ) leads to similar long-time dynamics as in the TIC phase. That is, we observe the constant formation, reshaping and fracture of dense clusters independent of the initial conditions. We note that the average Euler number is negative. This is to be expected as the proportion of dense domains is larger than the proportion of dilute domains, see ref. 33.

However, notable differences for mean densities below and above the spinodal are present in the short time dynamics (marked in gray in Fig. 8(b)). For  $\rho_0 < \rho_s$ , the nucleation of clusters follows the onset of meso-scale turbulence. As this dynamical state needs some time to evolve from the initial conditions there is a time lag until the nucleation of clusters. For  $\rho_0 > \rho_s$ , dense clusters form through spinodal decomposition immediately. This is represented by the peak of the Euler number in Fig. 8(b). This is to be expected, as the spinodal marks the point when the homogeneous density profile loses linear stability. That is, small perturbations of the initial density profile are enough to trigger the formation of clusters. However, the linear stability analysis of the homogeneous density profile in general bears no significance once that particular state is destabilized. Hence, a linear stability analysis around a homogeneous density profile does not provide hints at the long-time dynamics.

#### A.6 Dependence of the characteristic domain size on the vortex size

Fig. 4 of the main text shows the dependence of the characteristic domain size on the characteristic vortex size  $\Lambda$ , which is obtained from the stability analysis (Section A.1). Note though, that the characteristic vortex size measured by the velocity correlation function differs from  $\Lambda$  in the incompressible system due to energy transport between scales for  $\lambda_0 \neq 0$ .<sup>15,46</sup> However, the velocity correlation function is computed by averaging over the entire simulation domain. Hence, information about the clusters are mixed with information about the dilute phases. Therefore, the velocity correlation function is not suited to provide information about the characteristic vortex size in the TIC phase.

Fig. 9 shows the temporal evolution of the characteristic domain size. Therein, the characteristic domain size is computed from the density correlation function ( $L_c$ ), the perimeter ( $L_p$ ) and the Euler number ( $L_E$ ). Details on the different

**Table 1** Parameters used to study the dependence of the characteristic domain size on the vortex size

$\Lambda$	$\Gamma_0$	$\Gamma_2$
5.7	-0.42	0.17
7.3	-0.68	0.46
8.9	-1.0	1.0
10.5	-1.38	1.92
12.0	-1.83	3.36

approaches can be found in Sections B.2 and B.3. We remark that the density correlation function provides a characteristic domain size in simulation units. Characteristic domain sizes derived from the Minkowski functionals only show a proportionality, but might lack the correct units. This explains the different scales in Fig. 9 and the need to rescale the data in Fig. 4. The time evolution of  $L_c$  and  $L_p$  appears quite similar. For  $L_E$ , fluctuations are more pronounced. This is expected as the Euler number is a topological measure that varies on an integer scale. Furthermore, we fixed  $\Lambda$  while varying  $\Gamma_0$  and  $\Gamma_2$ , which does not influence the results significantly. Values of  $\Gamma_0$  and  $\Gamma_2$  are provided below (Table 1).

## B Methods

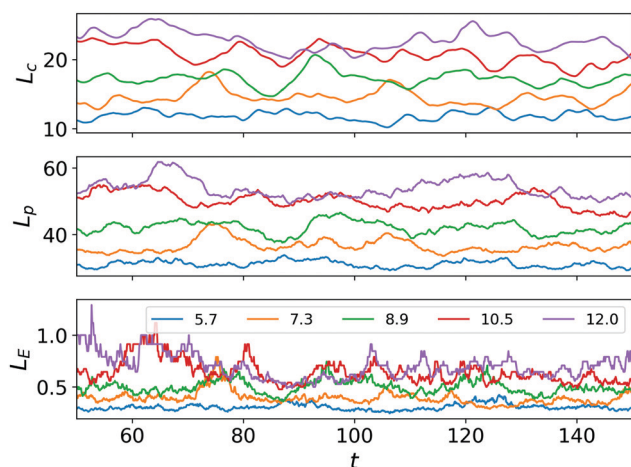
### B.1 Numerical Implementation

We apply a pseudo-spectral method to numerically solve the evolution equations eqn (1). For the time integration an operator splitting approach is used,<sup>47</sup> where nonlinearities are accounted for by a 2/3 dealias rule.<sup>48</sup> To circumvent the well-known Gibbs phenomena when applying spectral methods to conservation laws,<sup>49,50</sup> we use a large diffusion constant  $D$  to smooth phase boundaries. This allows us to study the dynamics on long timescales for large system sizes without the need for very fine grids.

The equations eqn (1) are solved on a quadratic domain of linear size  $L = 16\pi$  with periodic boundary conditions and  $256 \times 256$  spatial grid points. To check for finite size effects we additionally use a domain size  $L = 32\pi$  with  $512 \times 512$  grid points, see Section A.3. The time step is chosen as  $\Delta t = 10^{-3}$ . Data is stored for every 100th time step. Simulations are run until  $T = 150$ , unless indicated otherwise. When taking statistics, an initial transient of  $\Delta t = 50$  is neglected.

### B.2 Cluster size analysis

Besides the Minkowski functionals, we also analyse our system with the more common tools of cluster analysis. To do so we identify dense clusters by thresholding, where we use a threshold value  $\rho_{th} = 0.65$ . To account for periodicity, each snapshot is copied four times and taped together periodically. From this enlarged snapshot, the largest cluster is determined and removed from the original snapshot. This procedure is repeated iteratively until no cluster is left. This ensures that each cluster is only accounted for once and identified correctly. From this data the number of clusters at each time and the



**Fig. 9** Characteristic length scale over time for different values of  $\Lambda$  (color coded). The characteristic length scale is either computed from the correlation function (upper panel), perimeter (middle) or Euler number (lower panel).



cluster size distribution can be computed. These quantities are shown in Fig. 4 in the main text.

Moreover, we use the density  $\rho(\mathbf{x}, t)$  to compute the spatial density correlation at each time step. From the spatial density correlation function a characteristic domain size can be computed by taking the first moment of its Fourier transform.

### B.3 Minkowski functionals

Minkowski functionals represent a rigorous mathematical framework to study the morphology (*i.e.* the topological and geometric properties) of a structure. Minkowski functionals can be defined by requiring additivity, motion invariance and continuity for union of compact and convex sets. From these intuitive but abstract requirements one can characterize the Minkowski functionals in  $d$  dimensions. The following overview is based on.<sup>32</sup>

**B.3.1 Mathematical definition.** A subset  $S \subset \mathbb{R}^d$  is in the class  $\mathcal{R}$  if and only if there exist compact and convex sets  $K_i$  such that  $S = \bigcup_{i=1}^N K_i$ . We call a functional  $M: \mathcal{R} \rightarrow \mathbb{R}$  a Minkowski functional if the following properties hold:

1. Additivity: for  $S, T \in \mathcal{R}$

$$M(S \cup T) = M(S) + M(T) - M(S \cap T). \quad (6)$$

2. Motion invariance: for  $S \in \mathcal{R}$  and a  $g \in \mathcal{G}$ , where  $\mathcal{G}$  is the group of translational and rotational motions,

$$M(gS) = M(S). \quad (7)$$

3. Continuity: for convex sets  $S, S_n \in \mathcal{R}$  with  $S_n \rightarrow S$

$$M(S_n) \rightarrow M(S). \quad (8)$$

One can proof that there are only  $d + 1$  independent Minkowski functionals  $M_0, \dots, M_d$  in  $d$  spatial dimensions. That is, all other functionals fulfilling eqn (6)–(8) can be written as a linear combination of the  $d + 1$  functionals  $M_0, \dots, M_d$  (Hadwiger theorem). Furthermore, in two dimensions we can give a direct characterization of the functionals by simple properties of a structure as its area  $A$ , perimeter  $U$  and Euler characteristic  $\chi$

$$M_0(S) = A, \quad M_1(S) = U, \quad M_2(S) = \chi. \quad (9)$$

**B.3.2 Application of Minkowski functionals to numerical data.** To apply Minkowski functionals to numerical or experimental data we have to define what constitutes a structure. This is usually done by thresholding. For example, defining

$$\tilde{\rho}(\mathbf{x}, t, \rho_{\text{th}}) = \begin{cases} 1, & \text{if } \rho(\mathbf{x}, t) \geq \rho_{\text{th}} \\ 0, & \text{if } \rho(\mathbf{x}, t) < \rho_{\text{th}} \end{cases} \quad (10)$$

for the density  $\rho(\mathbf{x}, t)$  given by eqn (1a) creates a binary picture. For our numerical data, we use a threshold value  $\rho_{\text{th}} = 0.65$ . Areas with  $\tilde{\rho}(\mathbf{x}, t, \rho_{\text{th}}) = 1$  (white) make up the structures to be measured, while  $\tilde{\rho}(\mathbf{x}, t, \rho_{\text{th}}) = 0$  (black) is considered as background. This provides an additional characterization of the Euler characteristic in two dimensions as the difference of white and black connected finite domains

$$\chi = N_w - N_b. \quad (11)$$

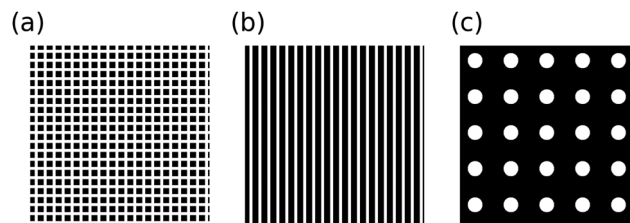


Fig. 10 Examples for structures with Euler characteristic  $\chi < 0$  (a),  $\chi = 0$  (b) and  $\chi > 0$  (c). Exact values are  $\chi = -400, 0, 25$  respectively.

From this classification one can easily derive the Euler characteristic of a circle (1), a hole (−1) and a flat interface (0) in 2D (Fig. 10). These structures represent the final state of equilibrium systems quenched into the spinodal regime.<sup>33</sup>

We apply a customized algorithm to compute the Minkowski functionals with periodic boundary conditions. Clearly, the (total) area  $A$  is not affected by the boundary and can be obtained straight forwardly. For the perimeter  $U$ , we first compute the interface as for non-periodic images.<sup>51</sup> Then, pixels at the boundary of the computation box that are falsely identified as interface are subtracted. The computation of the Euler number is outlined in the following.

**B.3.3 Computation of the Euler characteristic.** Our method to compute the Euler characteristic is based on the algorithm presented in ref. 52. This algorithm suffers from two drawbacks. Firstly, objects have to be individually separated from the background to obtain the right Euler characteristics. This is tedious, especially for periodic boundary conditions. Moreover, the algorithm only works for 4-connectivity. Both of these shortcomings can be overcome by a simple additional step. We quickly present the basic idea of the algorithm.

Assume a binary picture as shown in Fig. 11 (left side). Each corner is assigned a value based on the number of neighbouring white pixels. Let  $N_0, \dots, N_4$  denote the number of corners with value 0, ..., 4. Then, assuming 4-connectivity and separating the object from possible additional objects in the image, the Euler characteristic is given by<sup>52</sup>

$$E = \frac{N_1 - N_3}{4}. \quad (12)$$

However, assume a situation as depicted in Fig. 11(b). If the two objects are not treated individually, then the computed Euler

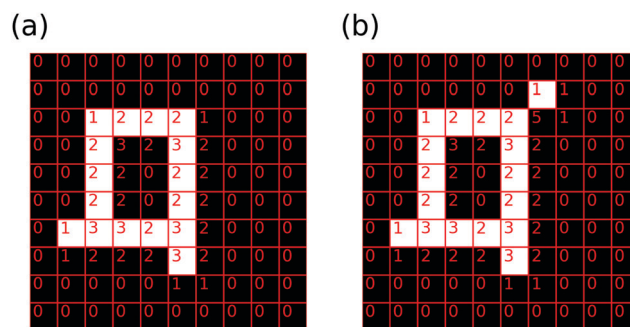


Fig. 11 Original binary picture with signed corners (a) and after adding a pixel while applying the new algorithm (b).



characteristic is wrong due to two white pixels touching each other diagonally. This can be rectified by an additional step in the algorithm: The corners that carry the value 2 and have identical diagonal neighbours are assigned a new number, for simplicity 5. Then, the Euler characteristic is obtained by either adding  $\frac{N_5}{2}$  to eqn (12) for 4-connectivity or subtracting it for 8-connectivity. Hence, eqn (12) needs to be modified to

$$E_{4,8} = \frac{N_1 - N_3 \pm 2N_5}{4}. \quad (13)$$

This can be understood from a simple topological argument. First, assuming 4-connectivity, a diagonal connection does not change the topology. Hence, a diagonal connection is the same as two corners with value 1. Hence, we have to add twice the amount of such corners. Next, assume 8-connectivity. If there is a diagonal connection, then the topology is not changed by adding an additional pixels. This increases the amount of  $N_3$  by 1 and decreases the amount of  $N_1$  by 1.

**B.3.4 Scaling of the Minkowski functionals with the characteristic domain size.** We now want to connect the Minkowski functionals  $M_i$  to the characteristic domain size  $L^*$  (see Section B.2). Due to the properties of the Minkowski functionals one usually assumes the relationship

$$M_i \sim (L^*)^{-i} \quad (14)$$

for systems exhibiting dynamic scaling.

To illustrate the idea behind this scaling hypothesis, we consider the example depicted in Fig. 12. Clearly, the pattern  $S_1$  is a scaled version of the pattern  $S_0$  with scaling factor  $\kappa = 2$ , while the morphology is unchanged. The pattern at  $S_1$  can be constructed from the pattern at  $S_0$  by cutting out the highlighted part in Fig. 12(a), which we will denote by  $S_c$ , and enlarging  $S_c$  to the domain size. This idea can be used to compute the values of the Minkowski functionals of  $S_1$  by only utilizing the scaling factor  $\kappa$  and the values of the Minkowski functionals for  $S_0$ . To do so, we use the additivity of Minkowski functionals, see eqn (6), and their homogeneity, *i.e.*

$$M_i(\kappa S) = \kappa^{d-i} M_i(S). \quad (15)$$

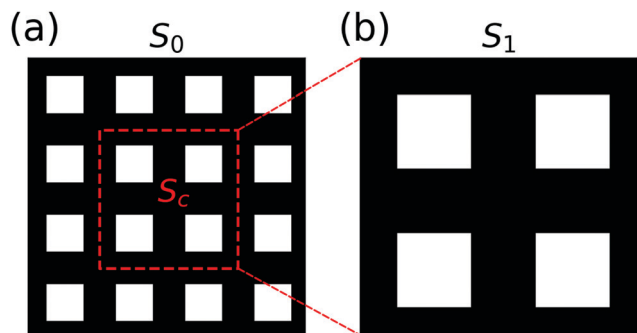


Fig. 12 Patterns  $S_0$  (a) and  $S_1$  (b). Red dotted line illustrates that  $S_1$  can be generated by cutting out  $S_c$  and enlarging it to the domain size.

Due to additivity, it holds

$$M_i(S_c) = \frac{1}{\kappa^d} M_i(S_0), \quad (16)$$

while clearly

$$S_1 = \kappa S_c. \quad (17)$$

Combining those two steps results in

$$\begin{aligned} M_i(S_1) &= M_i(\kappa S_c) \\ &= \kappa^{d-i} M_i(S_c) \\ &= \frac{\kappa^{d-i}}{\kappa^d} M_i(S_0) \\ &= \kappa^{-i} M_i(S_0). \end{aligned} \quad (18)$$

If the system exhibits dynamic scaling for  $t > t_0$ , *i.e.* the pattern at any point  $t$  is a scaled version of the pattern at some point  $t_0$ , it holds  $\kappa(t) = L^*(t)/L^*(t_0)$ , where  $L(t)$  and  $L(t_0)$  are the characteristic domain sizes at  $t_0$  and  $t$  respectively. Hence, we deduce

$$M_i(S[t]) \sim (L^*[t])^{-i}, \quad (19)$$

which provides the desired relationship between the Minkowski functionals and the characteristic domain size.

## Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) through grants HE 5995/3-1 (SH, VMW and AB), BA 1222/7-1 (MB and GA) and SFB 910 (projects B4 (HS) and B5 (MB)). GA and AB are thankful for partial support from the Israel Science Foundation grant 373/16.

## Notes and references

- 1 T. Vicsek and A. Zafeiris, *Phys. Rep.*, 2012, **517**, 71–140.
- 2 M. C. Marchetti, J. F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao and R. A. Simha, *Rev. Mod. Phys.*, 2013, **85**, 1143–1189.
- 3 J. Prost, F. Jülicher and J.-F. Joanny, *Nat. Phys.*, 2015, **11**, 111–117.
- 4 A. Be'er and G. Ariel, *Movement Ecol.*, 2019, **7**, 1–17.
- 5 H. Chaté, *Annu. Rev. Condens. Matter Phys.*, 2020, **11**, 189–212.
- 6 M. Bär, R. Großmann, S. Heidenreich and F. Peruani, *Annu. Rev. Condens. Matter Phys.*, 2020, **11**, 441–466.
- 7 C. Dombrowski, L. Cisneros, S. Chatkaew, R. E. Goldstein and J. O. Kessler, *Phys. Rev. Lett.*, 2004, **93**, 2–5.
- 8 L. H. Cisneros, R. Cortez, C. Dombrowski, R. E. Goldstein and J. O. Kessler, *Exp. Fluids*, 2007, **43**, 737–753.
- 9 T. Ishikawa, N. Yoshida, H. Ueno, M. Wiedeman, Y. Imai and T. Yamaguchi, *Phys. Rev. Lett.*, 2011, **107**, 1–4.
- 10 K. A. Liu and I. L. Lin, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2012, **86**, 1–6.



- 11 E. B. Steager, C. B. Kim and M. J. Kim, *Phys. Fluids*, 2008, **20**, 073601.
- 12 R. Großmann, P. Romanczuk, M. Bär and L. Schimansky-Geier, *Phys. Rev. Lett.*, 2014, **113**, 1–5.
- 13 H. P. Zhang, A. Be'er, R. S. Smith, E. L. Florin and H. L. Swinney, *EPL*, 2009, **87**, 48011.
- 14 H. H. Wensink, J. Dunkel, S. Heidenreich, K. Drescher, R. E. Goldstein, H. Löwen and J. M. Yeomans, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 14308–14313.
- 15 J. Dunkel, S. Heidenreich, M. Bär and R. E. Goldstein, *New J. Phys.*, 2013, **15**, 045016.
- 16 S. Heidenreich, J. Dunkel, S. H. L. Klapp and M. Bär, *Phys. Rev. E*, 2016, **94**, 29–31.
- 17 H. Reinken, S. H. L. Klapp, M. Bär and S. Heidenreich, *Phys. Rev. E*, 2018, **97**, 1–20.
- 18 Y. Fily and M. C. Marchetti, *Phys. Rev. Lett.*, 2012, **108**, 1–5.
- 19 J. Bialké, H. Löwen and T. Speck, *EPL*, 2013, **103**, 1–6.
- 20 M. E. Cates and J. Tailleur, *Annu. Rev. Condens. Matter Phys.*, 2015, **6**, 219–244.
- 21 G. Ariel, M. Sidortsov, S. D. Ryan, S. Heidenreich, M. Bär and A. Be'er, *Phys. Rev. E*, 2018, **98**, 1–10.
- 22 A. Be'er, B. Ilkanaiv, R. Gross, D. B. Kearns, S. Heidenreich, M. Bär and G. Ariel, *Commun. Phys.*, 2020, **3**, 1–6.
- 23 C. W. Wolgemuth, *Biophys. J.*, 2008, **95**, 1564–1574.
- 24 A. E. Hamby, D. K. Vig, S. Safonova and C. W. Wolgemuth, *Sci. Adv.*, 2018, **4**, eaau0125.
- 25 V. M. Worlitzer, G. Ariel, A. Be'er, H. Stark, M. Bär and S. Heidenreich, *New J. Phys.*, 2021, **23**, 033012.
- 26 T. B. Saw, A. Doostmohammadi, V. Nier, L. Kocgozlu, S. Thampi, Y. Toyama, P. Marcq, C. T. Lim, J. M. Yeomans and B. Ladoux, *Nature*, 2017, **544**, 212–216.
- 27 K. Kawaguchi, R. Kageyama and M. Sano, *Nature*, 2017, **545**, 327–331.
- 28 K. Copenhagen, R. Alert, N. S. Wingreen and J. W. Shaevitz, *Nat. Phys.*, 2021, **17**, 211–215.
- 29 S. D. Ryan, G. Ariel and A. Be'er, *Biophys. J.*, 2016, **111**, 247–255.
- 30 A. Sokolov, I. S. Aranson, J. O. Kessler and R. E. Goldstein, *Phys. Rev. Lett.*, 2007, **98**, 1–4.
- 31 A. U. Oza, S. Heidenreich and J. Dunkel, *Eur. Phys. J. E: Soft Matter Biol. Phys.*, 2016, **39**, 4–11.
- 32 K. R. Mecke, *Statistical Physics and Spatial Statistics*, Berlin, Heidelberg, 2000, pp. 111–184.
- 33 V. Sofonea and K. R. Mecke, *Eur. Phys. J. B*, 1999, **8**, 99–112.
- 34 J. Elgeti, M. E. Cates and D. Marenduzzo, *Soft Matter*, 2011, **7**, 3177–3185.
- 35 C. Torney and Z. Neufeld, *Phys. Rev. Lett.*, 2007, **99**, 078101.
- 36 A. Jayaram, A. Fischer and T. Speck, *Phys. Rev. E*, 2020, **101**, 022602.
- 37 R. Großmann, I. S. Aranson and F. Peruani, *Nat. Commun.*, 2020, **11**, 5365.
- 38 A. M. Lacasta, J. M. Sancho and F. Sagués, *Phys. Rev. Lett.*, 1995, **75**, 1791–1794.
- 39 L. Berthier, J.-L. Barrat and J. Kurchan, *Phys. Rev. Lett.*, 2001, **86**, 2014–2017.
- 40 S. Berti, G. Boffetta, M. Cencini and A. Vulpiani, *Phys. Rev. Lett.*, 2005, **95**, 224501.
- 41 P. Perlekar, R. Benzi, H. J. H. Clercx, D. R. Nelson and F. Toschi, *Phys. Rev. Lett.*, 2014, **112**, 014502.
- 42 T. Speck, J. Bialké, A. M. Menzel and H. Löwen, *Phys. Rev. Lett.*, 2014, **112**, 1–5.
- 43 T. Speck, A. M. Menzel, J. Bialké and H. Löwen, *J. Chem. Phys.*, 2015, **142**, 224109.
- 44 D. Levis, J. Codina and I. Pagonabarraga, *Soft Matter*, 2017, **13**, 8113–8119.
- 45 R. Wittkowski, A. Tiribocchi, J. Stenhammar, R. J. Allen, D. Marenduzzo and M. E. Cates, *Nat. Commun.*, 2014, **5**, 1–9.
- 46 V. Bratanov, F. Jenko and E. Frey, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 15048–15053.
- 47 R. Kozlov, *J. Comput. Appl. Math.*, 2008, **222**, 592–607.
- 48 C. Canuto, M. Y. Hussaini, A. Quarteroni and T. A. Zang, *Spectral methods: fundamentals in single domains*, Springer Science & Business Media, 2007.
- 49 S. Ghosh, M. Hossain and W. H. Matthaeus, *Comput. Phys. Commun.*, 1993, **74**, 18–40.
- 50 D. Gottlieb and C.-W. Shu, *SIAM Rev.*, 1997, **39**, 644–668.
- 51 S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart and T. Yu, scikit-image contributors, *PeerJ*, 2014, **2**, e453.
- 52 J. Sossa-Azuela, R. Santiago-Montero, M. Pérez-Cisneros and E. Rubio-Espino, *J. Appl. Res. Technol.*, 2013, **11**, 360–370.

