

Giant fluctuations and structural effects in a flocking epithelium

Fabio Giavazzi¹, Chiara Malinverno², Salvatore Corallino², Francesco Ginelli³, Giorgio Scita^{4,2} and Roberto Cerbino¹

¹ Università degli Studi di Milano, Dipartimento di Biotecnologie Mediche e Medicina Traslazionale, I-20090 Segrate, Italy

² IFOM-FIRC Institute of Molecular Oncology, I-20139 Milano, Italy

³ SUPA, Institute for Complex Systems and Mathematical Biology and Department of Physics, University of Aberdeen, Aberdeen, AB24 3UE, United Kingdom

⁴ Università degli Studi di Milano, Dipartimento di Oncologia e Emato-Oncologia, I-20133 Milano, Italy

E-mail: fabio.giavazzi@unimi.it

Abstract. Epithelial cells cultured in a monolayer are very motile in isolation but reach a near-jammed state when mitotic division increases their number above a critical threshold. We have recently shown that a monolayer can be reawakened by over-expression of a single protein, RAB5A, a master regulator of endocytosis. This reawakening of motility was explained in term of a flocking transition that promotes the emergence of a large-scale collective migratory pattern. Here we focus on the impact of this reawakening on the structural properties of the monolayer. We find that the unjammed monolayer is characterised by a fluidisation at the single cell level and by enhanced non-equilibrium large-scale number fluctuations at a larger length scale. Also with the help of numerical simulations, we trace back the origin of these fluctuations to the self-propelled active nature of the constituents and to the existence of a local alignment mechanism, leading to the spontaneous breaking of the orientational symmetry.

1. Introduction

Fluctuations are of key importance in statistical physics, a paradigmatic example being the erratic motion displayed by a colloidal particle as a consequence of the random collisions with the molecules of the solvent. In systems at equilibrium, however, thermal fluctuations are usually small and most importantly obey a normal statistics, so that in many instances they can be easily accounted for or even safely neglected when the system is described at the macroscopic level. A familiar but remarkable exception is represented by thermodynamic systems close to a critical point [1], such as a binary mixture close to its consolution point. Due to the vanishingly small osmotic compressibility, concentration fluctuations become so large to be visible macroscopically, opening to the possibility of accessing important thermodynamic quantities with light scattering [2, 3] or quantitative microscopy [4, 5] methods. Far from equilibrium, fluctuations can become anomalously large in a generic region of the phase diagram, thus without the need of fine tuning the system close to a critical point. This can occur for instance in inert systems that are in a non-equilibrium state because of the presence of a macroscopic gradient [6, 7] or also for *active matter* systems, composed by individual entities intrinsically kept far from equilibrium by energy being injected and dissipated at the microscopic, single component level [8, 9]. A notable class of active matter is represented by living matter, typically fueled by internal biochemical processes. Living active matter examples range from sub-cellular biofilaments displaced by motor proteins [10, 11, 12, 13] to bacterial colonies [14, 15], insects swarms [16] and superior social vertebrates [17, 18, 19, 20].

All these systems exhibit remarkable fluctuations in significant variables, the most paradigmatic example being represented by the so called giant number fluctuations (GNF), a hallmark of the inhomogeneous spatial distribution of the individual components in the system [9]. A very direct way to quantify GNF is to consider the fluctuation $\delta N = N - \langle N \rangle$ in the number N of moving entities contained in a region of prescribed size around its mean value $\langle N \rangle$. If the entities are randomly and independently distributed in space (the so-called spatial Poisson process) the hypotheses of the central limit theorem (CLT) are fulfilled and the mean squared amplitude of the number fluctuation is expected to be equal to $\langle \delta N^2 \rangle \sim \langle N \rangle$. If, on the other hand, long-ranged positional correlations are present, significant deviations from this simple behaviour are expected and different scaling exponents are observed for $\langle \delta N^2 \rangle$ as a function of $\langle N \rangle$ [21]. GNF have been shown to be a simple yet powerful and universal indicator of mesoscopic scale fluctuations in active matter systems [9], such as for instance active granular materials [22], bacteria [14] and cells [23, 24].

Truly asymptotic GNF, on the other hand, characterise active matter systems that exhibit long ranged order, such as polar *flocking*, a form of collective motility displayed, for instance, by birds flying in flocks and that can be understood through a minimal model (known as the Vicsek model) combining self-propulsion and effective alignment interactions [25]. Hydrodynamic theory results [26, 27], and numerical simulations based on the Vicsek model suggest that flocking systems exhibit anomalous fluctuations $\langle \delta N^2 \rangle \sim \langle N \rangle^{8/5}$ [28, 29]. However, the experimental observation of GNF in flocking biological systems has been so far quite scarce and limited to systems showing nematic order [30]. Also, while the flocking transition has been thoroughly investigated in low density systems [25, 26, 27, 28, 29, 31], it is not yet fully understood how it occurs in systems characterised by a large density, such as confluent cellular monolayers. In particular, little is known about how the flocking transition affects the structural, physical properties from the microscopic to the mesoscopic scales. If from a biological standpoint, one is interested in better understanding the microscopic features that control the transition to collective motion, the physical properties emerging from the interplay between jamming [32], individual active deformation [24, 33, 34] non-equilibrium clustering effects

[35, 36] and the spontaneous symmetry breaking underlying the flocking transition are still not completely understood [37], especially at the finite (and often relatively small!) length and time scales of biological relevance.

In this work, we investigate experimentally the anomalous number fluctuations and spatial correlations occurring in highly confluent epithelial cell monolayers undergoing a flocking transition to polar order. As a model system, we study doxycycline-inducible, RAB5A-expressing populations of human mammary epithelial MCF-10A cells. RAB5A is a master regulator of the endocytic activity of the cell and it plays an important role in the dissemination of aggressive breast tumours [38, 39, 40]. In control monolayers, cell density increases due to mitotic division, which causes a near complete kinetic arrest akin to a jamming or rigidity transition [32, 41]. We have recently shown that, under these latter conditions [42], the elevation of RAB5A reawakens the motility of kinetically arrested monolayer, by promoting a flocking transition accompanied by the emergence of large-scale, collective, directed migration patterns [43, 25].

Here we study in more detail this flocking transition and its dramatic impact on the structural properties of the monolayer. In agreement with low density theoretical predictions, the onset of an ordered migrating state is found to be accompanied by the emergence of GNF that we also monitor as a low- q divergence in the monolayer static structure factor $S(q)$. We also find that in the reawakened monolayer, the small scale behaviour of $S(q)$ shows that the positional ordering of the cells is less pronounced compared to the near-jammed control, indicating significative fluidisation effects.

We interpret these results through the well-known Toner and Tu hydrodynamic theory of flocking [26, 27], and by making use of a minimal model of self-propelled soft disks, first introduced in [44, 45]. Our numerical results show that a simple self-propelled mechanical model with no adhesion forces is able to reproduce qualitatively and even quantitatively the static structure factor and the number fluctuations of both the highly confluent reawakened samples and the near-jammed controls. Furthermore, in agreement with experimentally motivated conjectures on the effects of RAB5A [42], the transition between these two phases can be triggered, at constant cell density, by decreasing the time scale by whom cells orient their self-propulsion with their local velocity. Finally, our numerical simulations suggest that RAB5A over-expression could also indirectly overcome the feedback mechanisms, such as contact-inhibition of locomotion [46], suppressing individual cell motility in the high-density, near-jammed phase.

2. Methods

2.1. Cell culture and transfection

MCF-10A cells were a kind gift of J. S. Brugge (Department of Cell Biology, Harvard Medical School, Boston, USA) and were maintained in Dulbecco's Modified Eagle Medium: Nutrient Mixture F-12 (DMEM/F12) medium (Invitrogen) supplemented with 5% horse serum, 0.5 mg/ml hydrocortisone, 100 ng/ml cholera toxin, 10 μ g/ml insulin and 20 ng/ml EGF. The cell line was authenticated by cell fingerprinting and tested for mycoplasma contamination. Cells were grown at 37 °C in humidified atmosphere with 5% CO₂. MCF-10A cells were infected with pSLIK-neo-EV (empty vector control) or pSLIK-neo-RAB5A lentiviruses and selected with the appropriate antibiotic to obtain stable inducible cell lines. Constitutive expression of EGFP-H2B was achieved by retroviral infection of MCF-10A cells with pBABE- puro-EGFP-H2B vector.

2.2. Microscopy observation

Cells were seeded in six-well plates (1.5×10^6 cells/well) in complete medium and cultured until a uniform monolayer had formed. RAB5A expression was induced, where indicated, 16 h before performing the experiment by adding fresh complete media supplemented with 2.5 $\mu\text{g/ml}$ doxycycline to cells. At the time of recording, fresh media containing EGF was added. An Olympus ScanR inverted microscope with $\times 10$ objective was used to acquire images with a frame rate of 10 min over a 24 h period. For each condition (RAB5A and control) 5 independent fields of view (FOVs) (much smaller than the entire culture plates and far from the boundaries) are captured. Each FOV is imaged both in fluorescence and in phase-contrast (Fig.1(a)-(b)). The assay was performed using an environmental microscope incubator set to 37 $^\circ\text{C}$ and 5% CO_2 perfusion. After cell induction, doxycycline was maintained in the media for the total duration of the time-lapse experiment.

2.3. Image processing, monolayer static structure factor and number fluctuations

To quantitatively characterise the structural properties of the cell spatial arrangement, we measure the isotropically-averaged static structure factor $S(q) \equiv \langle S(\mathbf{q}) \rangle_{|\mathbf{q}|=q}$ of the monolayer, obtained as the orientational average $\langle \cdot \rangle_{|\mathbf{q}|=q}$ of the full static structure factor

$$S(\mathbf{q}) = \frac{1}{N_0} \left\langle \sum_{n,m} e^{-j\mathbf{q} \cdot (\mathbf{x}_n - \mathbf{x}_m)} \right\rangle. \quad (1)$$

Here N_0 is the total number of cells, \mathbf{x}_n the position of the n -th cell (identified as the centroid of its nucleus) and $\langle \cdot \rangle$ indicates an average over different experimental realisations and, in the case of a stationary system, over a prescribed time window. Introducing the number density: $\rho(\mathbf{x}) = \sum_n \delta(\mathbf{x} - \mathbf{x}_n)$, and the fluctuation $\delta\rho(\mathbf{x}) = \rho(\mathbf{x}) - \rho$ around the average density ρ , the static structure factor can be also written as

$$S(\mathbf{q}) = \frac{1}{N_0} \left\langle |\hat{\delta\rho}(\mathbf{q})|^2 \right\rangle, \quad (2)$$

where $\hat{\cdot}$ indicates the 2D Fourier transform operation. As it can be appreciated in Fig. 1(a), there is a certain degree of variability both in the size and, more importantly, in the fluorescence intensity of the nuclei, which makes a fully automated tracking of all the nuclei rather difficult. To overcome this problem, we adopted an operator-supervised, semi-automatic procedure. Each image is first processed with a standard, centroid-based particle tracking algorithm [47], whose results are then carefully checked cell by cell and manually corrected with the help of a custom software developed in MATLAB. The data considered in this study consist of image sequences, acquired over a time window of 24 hours with a frame rate of six images per hour, of five independent FOVs for each of the two conditions (control and RAB5A). For each FOV, three images, evenly distributed in the time window comprised 5 and 15 h, are selected and analysed. As can be appreciated also from Fig. 2, within this time window the macroscopic parameters describing cell motility maintain an almost constant value and the system could be assumed to be in an approximately stationary state. In total, 30 images (with an average of about 2500 cells each) are used to increase the statistical accuracy of our analysis.

Beside evaluating the static structure factor, a very direct way to quantify the degree of spatial homogeneity in the distribution of discrete entities is to consider the fluctuation $\delta N = N - \langle N \rangle$ in the number N of cells contained in a region of prescribed size around its mean value $\langle N \rangle$. We

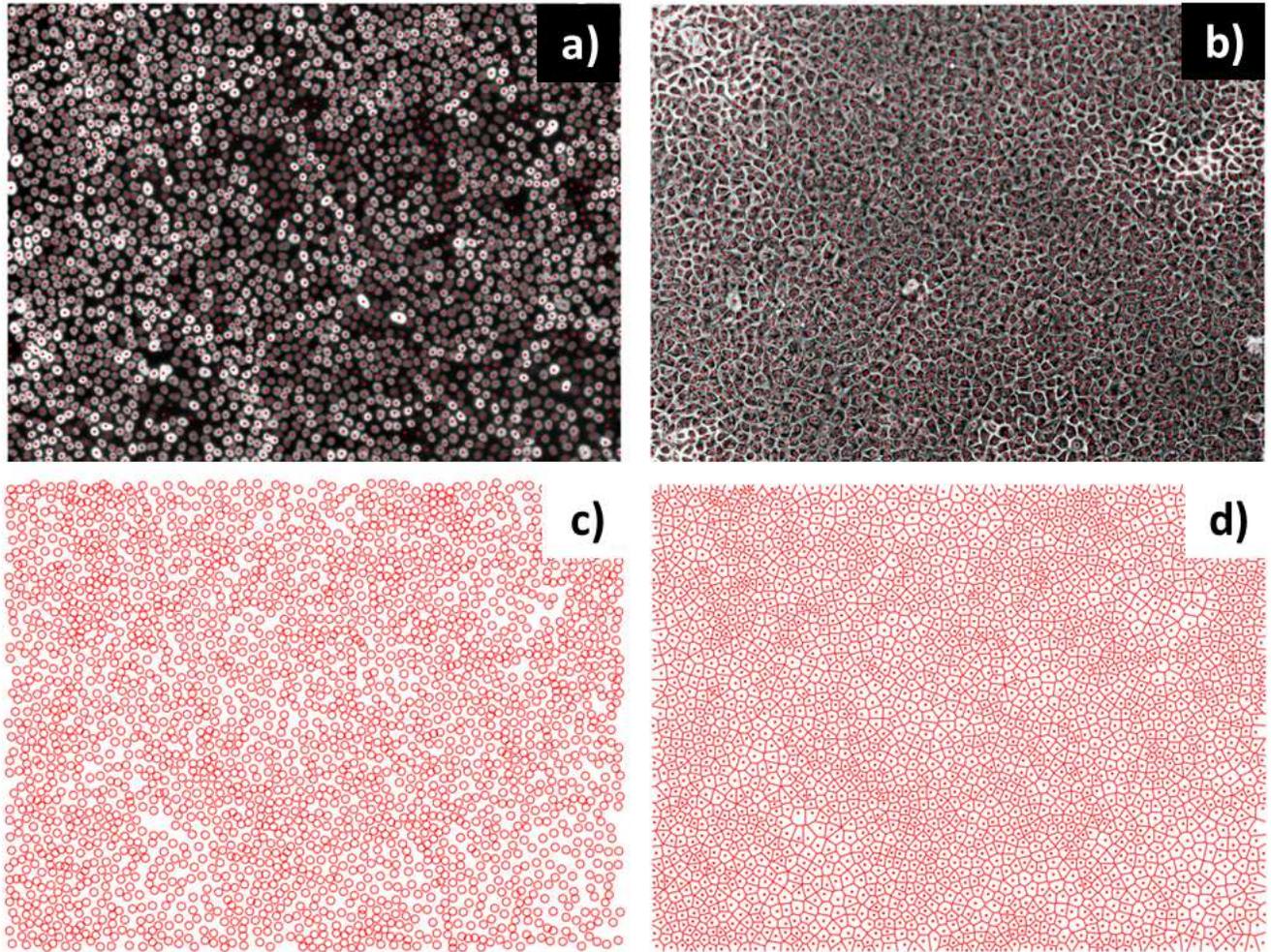


Figure 1. a) Representative fluorescence microscopy image of a highly-confluent control epithelial cell monolayer. The horizontal size of the image corresponds to 1.3 mm. Cell nuclei are fluorescently tagged and their positions can be determined with high accuracy (red dots). b) Same field of view of panel a) under phase-contrast imaging. (c) Reconstruction of the nuclei distribution, represented as disk having approximately the same size of the nuclei. (d) Voronoi diagram of the nuclei centres.

measure $\langle \delta N^2 \rangle$ as a function of $\langle N \rangle$ by counting the cells enclosed in non-overlapping square boxes of linear size a variable in the range $[1.3, 300] \mu\text{m}$.

The statistics of number fluctuations is intimately related to the static structure factor. In fact, the number of cells in a given region A (of area a^2) of the plane can be written as $N = \int_A d^2\mathbf{x} \rho(\mathbf{x})$. Correspondingly, the mean value of the cell number is $\langle N \rangle = a^2 \rho$ and the mean squared value is given by

$$\langle \delta N^2 \rangle = \frac{\langle N \rangle}{a^2} \int d^2\mathbf{q} |\hat{\chi}_A(\mathbf{q})|^2 S(\mathbf{q}), \quad (3)$$

where $|\hat{\chi}_A(\mathbf{q})|^2$ is the Fourier spectrum of the characteristic function of the region A . For example, if A is a square of side a , we have: $|\hat{\chi}_A(\mathbf{q})|^2 = a^4 \text{sinc}^2(\frac{q_x a}{2}) \text{sinc}^2(\frac{q_y a}{2})$.

From Eq. 3 it is clear that the behavior of $\langle \delta N^2 \rangle$ for large $\langle N \rangle$ is determined by the behaviour of $S(q)$ in the $q \rightarrow 0$ limit, that is

$$\frac{\langle \delta N^2 \rangle}{\langle N \rangle} \sim S(q \rightarrow 0), \quad (4)$$

in the large particles number limit. In fact, if a finite limit exists for $S(q)$ for $q \rightarrow 0$, one has normal fluctuations i.e. $\langle \delta N^2 \rangle \sim \langle N \rangle$. This is the case for instance of equilibrium systems, for which $S(q \rightarrow 0) = \rho k_B T \chi_T$, where k_B is the Boltzmann constant, T is the absolute temperature and χ_T is the isothermal compressibility [9].

On the contrary, if $S(q)$ diverges algebraically for small q , $S(q) \sim q^{-\alpha}$ with $\alpha > 0$, the scaling of the mean squared amplitude of the number fluctuations is no longer linear in the mean number, and anomalously large number fluctuations are present [48]: $\langle \delta N^2 \rangle \sim \langle N \rangle^{1+\frac{\alpha}{2}}$. As shown by the Toner and Tu hydrodynamic theory [26, 27], this is the case of active matter systems displaying collective motion. The spontaneously broken rotational invariance of the flocking state is, in fact, subject to large wavelength orientation fluctuations which are easily excited and decay very slowly (these are the celebrated Nambu-Goldstone modes [49, 50]). Due to the non-equilibrium, self-propelled nature of active matter, density fluctuations are coupled to the orientational ones, so that not only the velocity fluctuations correlations, but also the density fluctuation ones become long-ranged. Correspondingly, the density static structure factor $S(q)$ shows a divergence for $q \rightarrow 0$. While in the hydrodynamic theory $S(\mathbf{q})$ is characterised by a somehow complicated anisotropic behaviour, it is possible to compute its isotropically averaged small wavenumber behaviour. For polar aligning systems in two spatial dimensions, it is predicted to diverge as $S(q) \sim q^{-\alpha}$ with an exponent $\alpha = 6/5$ [29]. This finally gives the anomalous fluctuations $\langle \delta N^2 \rangle \sim \langle N \rangle^{8/5}$ [29], as confirmed by numerical simulations of the Vicsek model [28]. In the disordered non-flocking phase, a finite $S(0)$ and normal fluctuations are expected from the theory in the large scale, hydrodynamic limit. However, in active systems with repulsive interactions one expects non-equilibrium clustering effects, which could also lead either to true phase segregation phenomena or to anomalously large fluctuations at some mesoscopic, intermediate scale [35, 36, 51].

2.4. Quantification of cell motility and orientational order

Coarse-grained maps of the instantaneous velocity within the monolayer are obtained by performing a Particle Image Velocimetry (PIV) analysis on the phase-contrast image sequences, as described in detail in [42]. For each image, vectorial velocities are obtained over a square grid with spacing of 32 pixels, corresponding to $20.6 \mu\text{m}$. We indicate with $\mathbf{v}_j^{(k)}(t)$ the vectorial velocity measured at

time t in the j -th grid point of the k -th FOV. The overall motility of the cells within the monolayer at a given time t is well captured the mean squared velocity $v_{RMS}(t)$, defined as:

$$v_{RMS}(t) = \sqrt{\left\langle \frac{1}{M} \sum_{j=1}^M |\mathbf{v}_j^{(k)}(t)|^2 \right\rangle_k}, \quad (5)$$

while the degree of orientational order is quantified by the instantaneous order parameter $\psi(t)$, calculated as [25]:

$$\psi(t) = \left\langle \left| \frac{1}{M} \sum_{j=1}^M \frac{\mathbf{v}_j^{(k)}(t)}{|\mathbf{v}_j^{(k)}(t)|} \right| \right\rangle_k. \quad (6)$$

In the above expressions $\langle \cdot \rangle_k$ denotes an average over different (and statistically independent) FOVs (five in the present case) and M is the number of grid points per image ($M = 1024$ in our case). According to Eq.6, ψ is comprised between 0 and 1. If all the cells within each FOV coherently migrate in the same direction one has $\psi \simeq 1$, while $\psi \simeq 1/\sqrt{M}$ is expected for the case of a randomly oriented velocity field.

2.5. Numerical active matter model

To simulate the epithelial monolayer we use a collisional Vicsek-like model (CVM), first introduced in [44, 45]. In this agent-based model, N_0 soft disks with position \mathbf{r}_i and interaction radius σ_i ($i = 1, \dots, N_0$), move with self-propelled speed v_0 in the unit direction $\hat{\mathbf{n}}(\theta_i) = (\cos \theta_i, \sin \theta_i)$ (the particle polarisation) and interact via harmonic repulsive forces, according to the following overdamped dynamics

$$\dot{\mathbf{r}}_i = v_0 \hat{\mathbf{n}}(\theta_i) + \beta \sum_j^{N_0} \mathbf{F}_{ij} \quad (7)$$

where β controls the intensity of harmonic repulsion (akin to a cell membrane tension), and where the repulsive force may be written as

$$\mathbf{F}_{ij} = \begin{cases} 0 & \text{if } r_{ij} \equiv |\mathbf{r}_i - \mathbf{r}_j| \geq (\sigma_i + \sigma_j) \\ [r_{ij} - (\sigma_i + \sigma_j)] \hat{\mathbf{r}}_{ij} & \text{if } r_{ij} < (\sigma_i + \sigma_j) \end{cases}, \quad (8)$$

with $\hat{\mathbf{r}}_{ij}$ being the unit vector from particle i to j . The soft disk radii can be polydisperse, and we choose them to be equally distributed in the interval $[1 - p/2, 1 + p/2]$, with p a polydispersity parameter.

In the simplest implementation of this model [51], the particle polarisation angles θ_i evolve randomly according to $\dot{\theta}_i = \xi_i$, where ξ_i is a zero-average and uniformly distributed white noise with standard deviation η ,

$$\begin{aligned} \langle \xi_i \rangle &= 0 \\ \langle \xi_i(t) \xi_j(t') \rangle &= \eta^2 \delta_{ij} \delta(t - t'). \end{aligned} \quad (9)$$

However, it is possible to obtain a transition towards collective motion by including the tendency of the cell polarization $\hat{\mathbf{n}}(\theta_i)$ to align towards its actual velocity $\mathbf{v}_i \equiv \dot{\mathbf{r}}_i = v_i (\cos \psi_i, \sin \psi_i)$, that is

$$\dot{\theta}_i = \frac{1}{\tau} (\theta_i - \psi_i) + \xi_i \quad (10)$$

where τ is the realignment timescale [44].

A final, relevant control parameter for the CVM is obviously the packing fraction

$$\phi = \rho\pi\langle\sigma_i^2\rangle \approx \rho\pi\left(1 + \frac{p^2}{12}\right). \quad (11)$$

where we have used the fact that for $N_0 \gg 1$ one has $\sigma_i^2 = (1 + p^2/12)$. We have chosen a length unit such that the mean disk radius $\bar{\sigma}$ is set to one. We also rescale time units to get rid of a second model parameter, setting the force constant $\beta = 1$. In the $\tau \rightarrow \infty$ limit, the CVM is a model for non-ordering active brownian particles and is characterised by a phase diagram showing a liquid phase with non-equilibrium clustering, phase separation and a completely jammed state [51]. For finite τ (and not too large noise values) the CVM is known to display a transition from a disordered gas-like state to a flocking state [44] (i.e, an ordered polar liquid) characterised by GNF [45] as expected by hydrodynamic theory. A completely jammed state, finally, is still present at high packing fraction and low self propulsion speed [45]. In this work, we will not explore this part of the phase diagram, limiting ourselves to regions where the system is in a gas-like or liquid-like state, both in the presence and in the absence of polar order.

Numerical simulations have been performed by a simple Euler-Maruyama method with time-steppings $\Delta t = 10^{-2} \sim 2 \cdot 10^{-3}\ddagger$ in a square domain with periodic boundary conditions and linear system size L .

For simplicity, we measure the instantaneous orientational order ψ making use, in Eq. (6), of the instantaneous particle velocities $\mathbf{v}_i(t)$ rather than working with spatially coarse-grained quantities as one is forced to do in experimental measures. On the other hand, to better mimic the experimental procedure, we find more appropriate to measure the particles mean squared velocity $v_{RMS}(t)$ measuring the actual displacement over a finite time t_D ($t_D = 50$ time units in our case) and plugging in Eq. (5) the discrete velocity $\Delta\mathbf{v}_i(t) = [\mathbf{r}_i(t + t_D) - \mathbf{r}_i(t)]/t_d$.

The static structure factor is measured through Eq. (2), where the coarse-grained density field $\rho(\mathbf{x})$ is computed over a fine grid (of linear size $\ell \approx 0.5$) in order to resolve the small scale details. The isotropically-averaged structure factor $S(q)$ is then further averaged by sampling the stationary state dynamics at periodic intervals (every t_D time units). The same stationary dynamics sampling is also used for number fluctuations, which are measured following the box counting procedure outlined in section 2.3. In this work we present results from simulations of the rescaled ($\bar{\sigma} = 1$, $\mu = 1$) CVM (7)-(10) in a square domain of side $L = 192$. To model a confluent tissue, we work with packing fractions $\phi > 1$, yielding typical particles numbers $N_0 \gtrsim 1.2 \cdot 10^4$. While this is at least 4 times larger than the number of cells imaged in the experimental FOV, it should be recalled that in the experiments the monolayer is much larger than a single FOV. Thus, our choice minimises boundary effects at the experimental FOV scales while keeping the computational effort largely manageable. Also, notice that while we aim at a multi-scale description of the epithelial tissue structure, in our numerical simulations we are more interested in the local and mesoscopic scales, rather than in the hydrodynamic behaviour. The latter, which can be accessed at large scales, does only depends on general symmetries and conservation laws. Therefore, it is generically shared by all polar flocking models in the Toner & Tu phase (*i.e.* the so-called Vicsek universality class). Larger scale simulations ($N_0 \sim 10^6$) will be reported elsewhere [52]. Based on visual inspection of cell samples we also introduce a 20 % polydispersity in the cell interaction radius ($p = 0.4$) and,

\ddagger Chosen appropriately according to the magnitude of the velocity and to mean particles radius ratio. We have also verified that our numerical results are unchanged (in the statistical sense) when a smaller step size is selected.

somehow arbitrarily, we fix the noise standard deviation at $\eta = 0.45$. This leaves us with only 3 free parameters for our model: the packing fraction ϕ , the velocity v_0 and the realignment timescale τ .

3. Results and discussion

3.1. Experiments

RAB5A expression did not alter significantly the migration of individual MCF-10A cells. In Ref.[42] the single-cell behaviour of both RAB5A overexpressing cell and control cells has been characterized in random migration assays in terms of mean velocity, directional persistence and extension of protrusions. For none of these indicators a significant difference was found between control and RAB5A cells. In particular, in both cases, a mean velocity of around 60 m/h was measured. Thus, RAB5A effects on motility are emergent properties of cell collectives that elicit reawakening of locomotion of kinetically arrested, jammed epithelia. A detailed description of the effects of the over-expression of RAB5 on the kinetics of the confluent MCF-10A monolayer can be found in [42]. As it can be appreciated from Fig. 2(a), as well as in supplementary movies M1_PC and M1_FL, the overall cellular motility in the control monolayer is substantially suppressed and no trace of cell-cell orientational order can be found ($\psi < 0.1$, Fig. 2(c)). On the contrary, in the same observation window, RAB5A over-expression induces a spectacular reawakening of the cellular motility leading to long range orientational order [42] and a 20-fold increase in the absolute velocity (see Fig. 2(b)). Moreover, the vectorial velocity field shows a high degree of polar order ($\psi \gtrsim 0.9$ in Fig. 2(d)), indicating a strong cooperativity in the cell migration pattern, typical of a flocking state. The decrease of motility observed in the last part of the experiment can be mainly ascribed to the exhaustion of the Epithelial growth factor (EGF) in the cell culture media. In fact, EGF is strictly necessary not only for the growth but also for the motility of MCF-10A cells [53, 42, 54, 55]

The structural signatures associated with this phenomenology are described in detailed in the following paragraphs.

Control cells The isotropically-averaged static structure factor $S_0(q)$ of the control monolayer is shown in Fig. 3(a) (blue squares), while a representative snapshot of the cell distribution is reported in Fig. 1. As it can be appreciated from Fig. 3(a), in the high- q regime (that is for $q \gtrsim d^{-1}$, where $d \simeq 12 \mu m$ is a typical intercellular distance) the presence of a well defined set of damped oscillations clearly indicates that a partially spatially ordered liquid-like structure exists, strongly reminiscent of that of an assembly of repulsive particles at intermediate density. For comparison, the continuous line in Fig. 3(a) represents the static structure factor calculated for an equilibrium ensemble of hard disks (HDs) with diameter $d_0 = 11.6 \mu m$ and surface fraction $\phi_0 = 0.45$ [56]. This curve is in very good agreement with the experimental data for large q . Interestingly enough, the associated number density $\rho_0 = \frac{4\phi_0}{\pi d_0^2} = 4.3 \cdot 10^{-3} \text{ cells}/\mu m^2$ matches almost perfectly the experimentally measured value: $\rho_{exp}^0 = (4.3 \pm 0.1) \cdot 10^{-3} \text{ cells}/\mu m^2$. Moreover, the HD diameter d_0 is fully compatible with the average size $d_{exp} \simeq 11 \pm 1 \mu m$ of the nuclei, as obtained from a direct measurement on the fluorescence images. Based solely on these results, it would be tempting to say that the local structure of the monolayer could be explained by assuming purely a steric, repulsive interactions between the nuclei.

However, for smaller wavenumbers, a strong deviation from the behaviour expected for equilibrium HDs is observed. As q is decreased, $S_0(q)$ is characterised by a pronounced dip, indicating the presence of mesoscopic-scale inhomogeneities in the cell distribution, or clustering

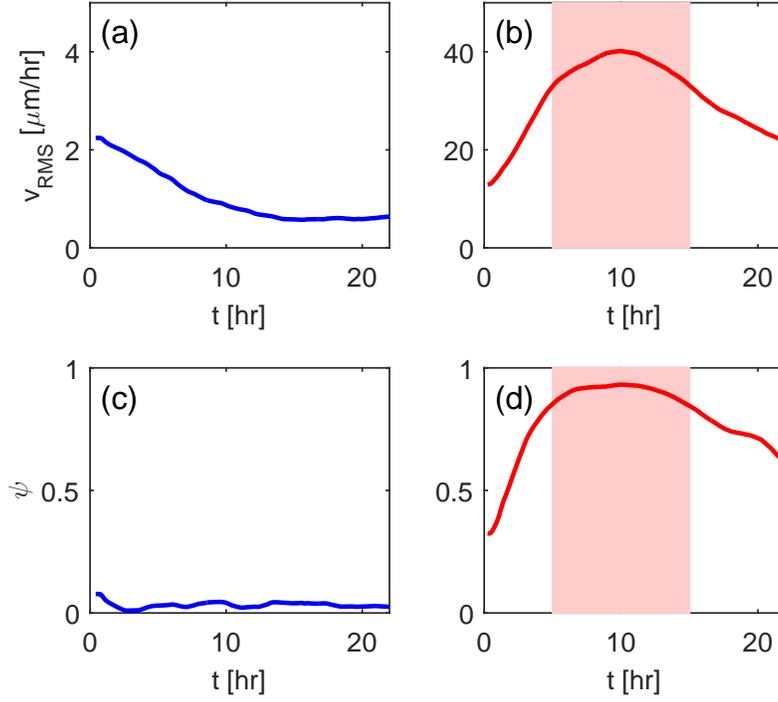


Figure 2. (a) Time evolution of the mean square velocity for the control monolayer during the experimental observation, showing a progressive slowing down of cellular motility. (b) Mean square velocity for RAB5A monolayer. The wide peak in v_{rms} (dashed area) coincides with a very large value assumed by the corresponding orientational order parameter ψ (panel (d)), indicating the RAB5A-induced reawakening of motility coincides with the emergence of a directed migratory pattern (flocking). The dashed area marks the time window 5-15 h over which stationary averages were taken. In stark contrast, no orientational order can be detected in the control monolayer (panel (c)), for which $\psi \simeq 0$.

[32, 36]. It is worth noting that this feature cannot appear in equilibrium systems without invoking attractive interactions and for this reason in active systems, one speaks of *non-equilibrium clustering*.

Finally, a plateau seems to set in for $q \rightarrow 0$, as predicted by the hydrodynamic theory. Correspondingly, the normalised fluctuations $\langle \delta N^2 \rangle / \langle N \rangle$ (blue squares in Fig.3(b)) show an anomalous increase at intermediate scales and eventually seem to approach a constant value for $\langle N \rangle \gtrsim 200$, signalling the onset of normal fluctuations at large length scales, unfortunately quite close to our experimentally accessible observation limit. An alternative estimate of the typical cluster size beyond which normal fluctuations sets in can be also obtained by considering the shoulder in $S_0(q)$, that occurs for $q_c \simeq 0.025 \mu m^{-1}$. The typical number of cells in a cluster can be estimated as $N_c \simeq \rho_{exp} (2\pi q_c^{-1})^2 \simeq 250$. This estimate is in agreement with the normalised mean squared number fluctuation of the control shown in Fig. 3(b), which saturates to a constant value for $\langle N \rangle \gtrsim 200$. It should be noted that, since the control monolayer is almost completely kinetically arrested, these fluctuations are not very dynamic in nature and appear to be substantially “frozen”

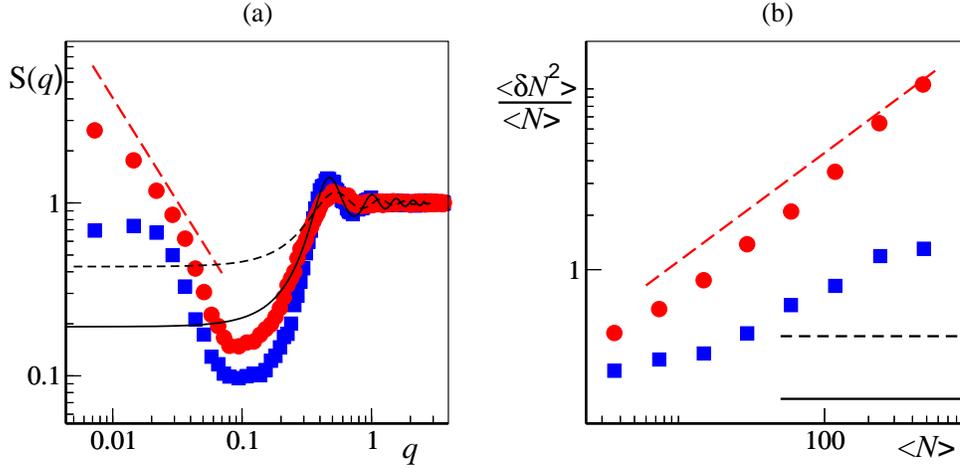


Figure 3. (a) Blue squares: static structure factor $S_0(q)$ of the control monolayer. The continuous line corresponds to the structure factor of an equilibrium ensemble of randomly scattered hard disks of diameter $d_0 = 11.6 \mu\text{m}$ and surface fraction $\phi_0 = 0.45$. This model is in fair agreement with the experimental data for large q but does not account for the behaviour of $S_0(q)$ in the lower- q regime, indicating the presence of non-equilibrium clusters. Red circles: structure factor $S_1(q)$ of the RAB5A monolayer. The dashed line corresponds to the structure factor for a monodisperse ensemble of hard disks of radius $d_0 = 10.0 \mu\text{m}$ and surface fraction $\phi = 0.25$. The marked divergence exhibited by $S_1(q)$ for $q \rightarrow 0$ is the signature of the presence of giant density fluctuations that are not compatible with equilibrium models. The dashed red line marks the theoretical prediction for flocking systems, $S_1(q) \sim q^{-6/5}$. (b) Normalised mean squared number fluctuations for the control (blue squares) and the RAB5A monolayer (red circles), respectively. The dashed red line marks the theoretical prediction (see text). The horizontal lines represent the asymptotic behaviours expected for large $\langle N \rangle$ by assuming the two hard disks models whose static structure factors are shown (with the same line types) in panel (a).

on the experimental timescales (see also Supplementary Movies M1_PC and M1_FL).

RAB5A cells The spatial distribution of RAB5A-overexpressing cells presents a number of peculiar features, as it can be appreciated by inspecting the representative snapshot shown in Fig. 4(b), the supplementary movies M2_PC and M2_FL, and Fig. 3(a) (red circles), where we show the experimentally measured structure factor $S_1(q)$. Similarly to the control monolayer, also in this case the high- q portion of the structure factor presents damped oscillations. Compared to controls, the peaks are however less pronounced and slightly shifted to the right, indicating a fluidisation of the monolayer. For reference, the dashed line in Fig. 3(a), which is in fair agreement with $S_1(q)$ for large q , represents the equilibrium structure factor of randomly scattered hard disks, with diameter $d_1 = 10.0 \mu\text{m}$ and surface fraction $\phi_1 = 0.25$. Compared with the one used for controls, this HD model is characterised by smaller particles, at a lower number density $\rho_1 = \frac{4\phi_1}{\pi d_1^2} = 3.2 \cdot 10^{-3} \text{ cells}/\mu\text{m}^2$.

It is worth noting that RAB5A over-expression has marginal effects on the rate of cell division of confluent monolayers [42] and, in fact, the measured cell number density $\rho_{exp}^{(1)} = (4.3 \pm 0.2) \cdot 10^{-3} \text{ cells}/\mu\text{m}^2$ is fully compatible with the one of control cells. Moreover, RAB5A has no effect on the size of the nuclei. These findings suggest that RAB5A expression significantly alter the effective

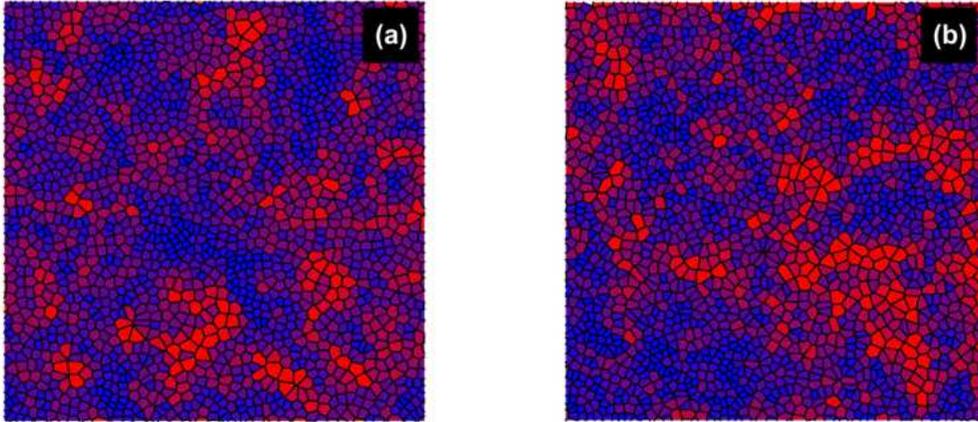


Figure 4. (a) Representative snapshot of the CVM in the ordered, high v_0 regime (see text) and (b) of the RAB5A monolayer. In both panels the Voronoi tessellation generated by the cell centres is shown. Voronoi cells are colour-coded according to their surface area. Each image contains around $N_0 \approx 2000$ cells for a linear size of about 1.0 mm in real space.

cell-cell interaction, namely, by decreasing the minimum distance between neighbouring cells.

At larger scales, a non-equilibrium clustering behaviour analogous to the one exhibited by the control is marked by a dip (albeit less pronounced) in the structure factor. However, in this case a marked divergence is observed for the smallest q , which is roughly compatible with the algebraic behaviour predicted by the hydrodynamic theory, $S(q) \sim q^{-6/5}$ in the hydrodynamic regime.

This behaviour is reflected by the marked CLT violation exhibited by the number fluctuations (red circles in Fig. 3(b)), showing, at least up to the larger investigated length scales, a clean algebraically scaling of the normalised number fluctuations $\langle \delta N^2 \rangle / \langle N \rangle \sim \langle N \rangle^\gamma$, with $\gamma = 0.66 \pm 0.04$. This value is quite close to the theoretical value predicted from the hydrodynamic theory for the flocking Toner & Tu Phase, $\gamma = \sigma/2 = 3/5$. To the best of our knowledge, this result represent the first clear measure of this anomalous fluctuations in a biological system showing flocking and long ranged polar order.

3.2. Active matter simulations

Disordered phase For reasons that will be clear in the following, we set the surface fraction to the value $\phi = 1.2$ and proceed to optimise the single cell velocity v_0 and the orientation time τ in order to reproduce the control cells experimental results. Good agreement is found for instance for $v_0 = 0.5$ and $\tau = 5.0$. For these values of the parameters, the systems lies in the disordered phase $\psi \approx 2 \cdot 10^{-2}$ (the blue line in Fig. 5(a)). Correspondingly, we show in Fig. 5(c) that the normalised number fluctuations (blue squares) converge to a constant plateau for sizes larger than the non-equilibrium clustering size, which indicates normal number fluctuations in agreement with the CLT.

A qualitative agreement can be found in the disordered, non completely jammed phase for a wide range of model parameters (see also [36], where the static structure factor for an ensemble of

non-equilibrium self-propelled hard disks is discussed). By contrast, a more quantitative agreement with the experimental static structure factor can be achieved by introducing a physical length scale. We do so by fitting the location of main high q peak of the numerically computed structure factor with the one of the experimental control $S_0(q)$. This provides us with a length scale of one simulation unit being approximately equal to $8.9\mu m$. In Fig. 5(d) (upper panel) the similarity between the disordered numerical static structure factor (full blue squares) and its experimental counterpart $S_0(q)$ (green open squares) can be fully appreciated over the full multiscale range. We even achieve an almost quantitative agreement ranging from the local, interaction range, structure to the non-equilibrium clustering mesoscopic range and up to the hydrodynamic range in which saturation to a constant is observed.

Moreover, once we express the mean disk interaction radius $\bar{\sigma}$ in the experimental length scale, Eq. (11) gives us a density $\rho \approx 4.7 \cdot 10^{-3} \text{ cells}/\mu m^2$, in reasonable agreement with the experimentally observed one, thus justifying our initial assumption $\phi = 1.2$. A typical sample of the CVM dynamics in the disordered phase can be found in supplementary movie M3.

Ordered phase For sufficiently low noise amplitudes, it is known that the CVM presents a transition to a polar flocking state as the density is increased or the particle polarity realignment timescale τ is decreased. As done in [42], we also chose to keep the noise amplitude and density constant ($\eta = 0.45$ and $\phi = 1.2$, respectively) and to change the re-orientation time. Keeping all the other parameters unchanged, we verify that the CVM undergoes a transition to collective motion for $\tau \lesssim 2$. We place ourself at a much lower value, namely $\tau = 0.45$, also to avoid the spontaneous segregation instability that is known to take place at the onset of collective motion, at least in low density systems [29].

Here the dynamics shows clear signs of collective motion, with an order parameter $\psi \approx 0.94$. However, the system shows spatial configurations characterised by a higher degree of positional order compared to the ones observed in the disordered phase. Correspondingly, the structure factor $S(q)$ shown in Fig. 5(d) (bottom panel, black line) is characterised by strong damped oscillations in the high q region. It is easy to understand this phenomenon: when the motion is disordered, the particles often push one against each other, resulting in a lower positional order. In the ordered regime, on the other hand, fluctuations around the mean velocity are smaller resulting in a higher positional order. This is actually the opposite of the fluidisation effects observed in the RAB5A experiments.

To reproduce the experimental data, we need to decrease the positional order without lowering drastically the orientational one. To this aim, it is necessary to increase the self-propulsion velocity v_0 . We do so keeping the noise amplitude constant, $\eta = 0.45$, and by finding an optimal parameter range for $v_0 \approx 2.0$. The equilibration kinetics and the static structure factor for these parameters are presented in Fig. 5(d), while a typical configuration is shown in Fig. 4(a). Also, a movie of the dynamics can be found as supplementary movie M4.

Starting from a fully disordered initial condition, the systems quickly aligns, as the order parameter reach its stationary value $\psi \approx 0.84$ (Fig. 5(a), red line). Correspondingly, the root mean squared velocity v_{rms} grows tenfold compared to the one measured in the disordered phase (Fig.5(b)), in good qualitative agreement with what observed comparing the root mean squared velocities of the RAB5A and the control experimental monolayers.

Due to the increased particle velocities (and the slightly decreased orientational order), positional order is significantly lowered, resulting in a notably less sharp peak structure in the high q structure factor, as shown in Fig. 5(d) (bottom panel, full red circles). In fact, one can verify

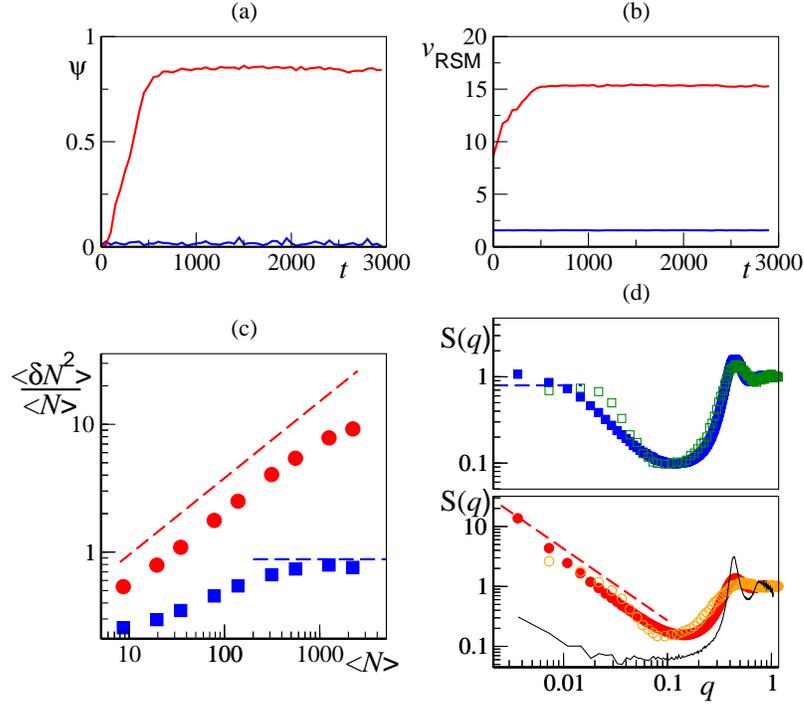


Figure 5. a) Orientational order parameter ψ vs. time (arbitrary units) for the ordered (red line) and disordered phases (see text for simulation parameters). b) Root mean squared velocity vs. time (both in arbitrary units) for the ordered (red line) and disordered phases. c) Normalised giant number fluctuations vs. the mean number of particles computed in sub-boxes of increasing linear size (see text). Red circles give the ordered phase results, while blue squares are for the disordered state. The dashed lines mark the expected hydrodynamic limit theoretical results for, respectively, the ordered phase (the red line is growing as $\langle N \rangle^{0.6}$) and disordered one (flat blue line). d) Isotropically averaged static structure factors vs. the wavenumber q (in μm^{-1} units): In the top panel, the numerical $S(q)$, computed in the disordered phase (full blue squares), is compared with the control experimental structure factor $S_0(q)$ (open green squares). In the bottom panel, the numerical $S(q)$ is computed in the ordered phase (full red circles) for $v_0 = 0.2$ and compared with the RAB5A experimental structure factor $S_1(q)$ (open orange circles). The black lines shows the numerical structure factor obtained in the ordered regime but at lower self-propulsion speed ($v_0 = 0.5$). The dashed lines mark the expected hydrodynamic limit theoretical results for the two phases (see text).

Simulations have been performed with timesteppings $\Delta t = 10^{-2}$ (disordered phase) and $\Delta t = 2 \cdot 10^{-3}$ (ordered phase). Static structure factors and GNF have been computed averaging over 1000 different configurations of the stationary dynamics (consecutive configurations are separated by $t_D = 50$ time units).

that the main peak is now lower than the numerical disordered data shown in the upper panel.

Overall, our numerical results suggests that in order to achieve both fluidisation and a transition to the flocking state, cells need to i) decrease their polarity reorientation timescale and ii) increase their self-propelled velocity. In our model, indeed, this is achieved by parameter changes of roughly one order of magnitude!

Good qualitative and quantitative agreement with the RAB5A monolayer structure is also

observed at smaller q , with the numerical $S(q)$ well reproducing the non-equilibrium clustering signatures at the mesoscale as well as the hydrodynamic scale divergence of the RAB5A structure factor (open orange circles in Fig. 5(d)).

The normalised number fluctuations are finally shown in Fig. 5(c) (red circles). Since we are able to access the hydrodynamic divergence at the relatively small scales of our present simulations (larger scale simulation will appear somewhere else [52]), we are also able to observe asymptotic density fluctuations compatible with theoretical predictions [29]. However, this is not the case for generic parameter values in the flocking phase, and we have verified that for different choices the hydrodynamic range gets pushed at larger length scales, above the one probed by the numerical simulations here reported.

In the above discussion the transition from a disordered state, representative of the control monolayer, to a polarly ordered, fluidised state is obtained by simultaneously decreasing the alignment time τ and increasing the self-propulsion speed v_0 , at constant membrane tension $\beta = 1$. Simple rescaling arguments show that changing β to values different from unit amounts to the following parameter shifts: $v_0 \rightarrow v_0/\beta$, $\tau \rightarrow \beta\tau$ and $\eta \rightarrow \eta/\sqrt{\beta}$. This suggests that the same transition can be also triggered by acting on different control parameters, for example by simultaneously decreasing τ and β , at constant the self-propulsion velocity. This route is similar to the one originally proposed in Ref. [42] in the context of a self-propelled Voronoi model [57] with alignment interactions [58], where the effect of RAB5A was interpreted in terms of faster cell-cell alignment (lower τ) combined with increased cell-cell adhesion. Beyond this qualitative correspondence, a detailed understanding of the differences (at small and intermediate scales) between the predictions of particle-based models, like the one used in this work, and vertex-like models [59, 57, 42, 58] is still lacking, and it is well beyond the scope of the present work.

4. Conclusions

We presented here a study of the structure of a confluent epithelial monolayer switching from an orientationally disordered state, close to kinetic arrest, to a fluidised, polarly ordered state showing directed collective migration. Strikingly, the effects of activity manifest themselves over a wide range of scales: at large scales, the flocking monolayer shows clear signs of anomalous number fluctuations in agreement with the prediction of the hydrodynamic theory of flocking. This is the first time these predictions have been unambiguously confirmed in biological active matter showing long ranged polar order. On the local scale, on the other hand, the reawakened cellular motility induces fluidisation effects with a clear structural signature. These findings are at variance with respect to previous studies where the transition from a fluid-like state to an arrested state (jamming) in a confluent epithelial monolayer was primarily described in terms of an increasing length scale associated with the heterogenous dynamics, with marginal or no effects on the static structure [32].

To gain further insights on the nature of this transition to fluidised collective motion, we studied numerically a simple mechanical model of soft self propelled disks, first introduced in [44, 45], kept at high packing fraction. This simple model, lacking adhesion forces, is able to reproduce qualitatively and even quantitatively the monolayer structure of both the control and the reawakened RAB5A on a wide range of scales, ranging from the single cell scale to the hydrodynamic collective range.

While in the context of the model, the transition at constant cell density is mainly controlled by the efficiency of the polarity alignment mechanism, the concurring experimental fluidisation is reproduced by a large increase of the particles self propelling speed. Experimentally, we showed that RAB5A expression does not alter the migration speed of sparsely seeded cells, which is

comparable, in absolute terms, to the one characterizing the collective migration of RAB5A cells in a confluent monolayer. Altogether, these results suggest that another effect of the RAB5A induced transition to flocking could be to reduce the mechanically induced feedbacks that tend to suppress cellular motility in the disordered state, such the ones driving the control monolayer close to a kinetically arrested state [46]. This intriguing hypothesis will be tested in future experiments, aimed at dissecting the detailed biochemical and biomechanical mechanisms connecting the overexpression of RAB5A with the emergence of the reported collective behaviour.

In the future, it would be also very interesting to investigate the relationship between the structural effects observed here and the monolayer dynamics at the mesoscopic, biologically relevant scale, such as for instance the effect of the spatial structure on the relaxation of density fluctuations or the coupling between high density elastic modes and velocity fluctuations, or the role played the intracellular adhesion in stabilizing the local structure.

Acknowledgments

We thank S. Henkes for useful discussions. FGia and RC acknowledge funding from the Italian Ministry of University and Scientific Research (MIUR) under the program *Futuro in Ricerca* - Project ANISOFT (RBFR125H0M) and from Regione Lombardia and CARIPLO foundation under the joint action *Avviso congiunto per l'incremento dell'attrattività del sistema ricerca lombardo e della competitività dei ricercatori candidati su strumenti ERC* - Project Light4Life. CM, SC and GS acknowledge funding from Associazione Italiana per la Ricerca sul Cancro (AIRC 10168 and 18621), MIUR, the Italian Ministry of Health, Ricerca Finalizzata (RF0235844), Worldwide Cancer Research (AICR-14-0335), and the European Research Council (Advanced-ERC-268836). CM was also supported by Fondazione Umberto Veronesi and SC by an AIRC fellowship. FGin acknowledges support from the Marie Curie Career Integration Grant (CIG) PCIG13-GA-2013-618399, and wish to thank the University of Milan and LibrOsteria for their hospitality while this work was underway.

-
- [1] Domb C 1996 *The Critical Point* (Taylor & Francis, London)
 - [2] Berne B J and Pecora R 1976 *Dynamic light scattering: with applications to chemistry, biology, and physics* (Courier Corporation)
 - [3] Sengers J V and Shanks J G 2009 *J. Stat. Phys.* **137** 857
 - [4] Giavazzi F and Cerbino R 2014 *J. Opt.* **16** 083001
 - [5] Giavazzi F, Fornasieri A, Vailati A and Cerbino R 2016 *The European Physical Journal E* **39** 103 ISSN 1292-895X URL <http://dx.doi.org/10.1140/epje/i2016-16103-9>
 - [6] Ortiz De Zarate J M and Sengers J V 2006 *Hydrodynamic fluctuations in fluids and fluid mixtures* (Elsevier)
 - [7] Giavazzi F, Savorana G, Vailati A and Cerbino R 2016 *Soft Matter* **12** 6588
 - [8] Ramaswamy S 2010 *Annual Review of Condensed Matter Physics* **1** 323–345 (Preprint <http://dx.doi.org/10.1146/annurev-conmatphys-070909-104101>) URL <http://dx.doi.org/10.1146/annurev-conmatphys-070909-104101>
 - [9] Marchetti M C, Joanny J F, Ramaswamy S, Liverpool T B, Prost J, Rao M and Simha R A 2013 *Rev. Mod. Phys.* **85**(3) 1143–1189 URL <https://link.aps.org/doi/10.1103/RevModPhys.85.1143>
 - [10] Schaller V 2010 *Nature* **467** 73
 - [11] Schaller V, Weber C A, Hammerich B, Frey E and Bausch A R 2011 *Proceedings of the National Academy of Sciences* **108** 19183–19188 (Preprint <http://www.pnas.org/content/108/48/19183.full.pdf>) URL <http://www.pnas.org/content/108/48/19183.abstract>
 - [12] Sumino Y, Nagai K H, Shitaka Y, Tanaka D, Yoshikawa K, Chaté H and Oiwa K 2012 *Nature* **483** 448
 - [13] Sanchez T, Chen D T N, DeCamp S J, Heymann M and Dogic Z 2012 *Nature* **491** 431

- [14] Zhang H P, Beer A, Florin E L and Swinney H L 2010 *Proceedings of the National Academy of Sciences* **107** 13626–13630 (Preprint <http://www.pnas.org/content/107/31/13626.full.pdf>) URL <http://www.pnas.org/content/107/31/13626.abstract>
- [15] Peruani F, Starruß J, Jakovljevic V, Sogaard-Andersen L, Deutsch A and B'ar M 2012 *Phys. Rev. Lett.* **108**(9) 098102 URL <https://link.aps.org/doi/10.1103/PhysRevLett.108.098102>
- [16] Attanasi A, Cavagna A, Castello L D, Giardina I, Melillo S, Parisi L, Pohl O, Rossaro B, Shen E, Silvestri E and Viale M 2014 *PLoS Computational Biology*. **10** e1003697
- [17] Ballerini M, Cabibbo N, Candelier R, Cavagna A, Cisbani E, Giardina I, Lecomte V, Orlandi A, Parisi G, Procaccini A, Viale M and Zdravkovic V 2008 *Proceedings of the National Academy of Sciences* **105** 1232–1237 (Preprint <http://www.pnas.org/content/105/4/1232.full.pdf>) URL <http://www.pnas.org/content/105/4/1232.abstract>
- [18] Cavagna A, Cimarelli A, Giardina I, Parisi G, Santagati R, Stefanini F and Viale M 2010 *Proceedings of the National Academy of Sciences* **107** 11865–11870 (Preprint <http://www.pnas.org/content/107/26/11865.full.pdf>) URL <http://www.pnas.org/content/107/26/11865.abstract>
- [19] Mora T, Walczak A M, Castello L D, Ginelli F, Melillo S, Parisi L, Viale M, Cavagna A and Giardina I 2016 *Nature Physics* **12** 1153
- [20] Ginelli F, Peruani F, Pillot M H, Chat   H, Theraulaz G and Bon R 2015 *Proceedings of the National Academy of Sciences* **112** 12729–12734 (Preprint <http://www.pnas.org/content/112/41/12729.full.pdf>) URL <http://www.pnas.org/content/112/41/12729.abstract>
- [21] Gabrielli A, Sylos Labini F, Joyce M and Pietronero L 2005 *Statistical Physics for Cosmic Structures* (Springer Berlin)
- [22] Deseigne J, Dauchot O and Chat   H 2010 *Phys. Rev. Lett.* **105**(9) 098001 URL <https://link.aps.org/doi/10.1103/PhysRevLett.105.098001>
- [23] Garcia S, Hannezo E, Elgeti J, Joanny J F, Silberzan P and Gov N S 2015 *Proceedings of the National Academy of Sciences* **112** 15314–15319 (Preprint <http://www.pnas.org/content/112/50/15314.full.pdf>) URL <http://www.pnas.org/content/112/50/15314.abstract>
- [24] Zehnder S, Suaris M, Bellaire M and Angelini T 2015 *Biophysical Journal* **108** 247–250 ISSN 0006-3495 URL <http://dx.doi.org/10.1016/j.bpj.2014.11.1856>
- [25] Vicsek T, Czir  k A, Ben-Jacob E, Cohen I and Shochet O 1995 *Phys. Rev. Lett.* **75**(6) 1226–1229 URL <http://link.aps.org/doi/10.1103/PhysRevLett.75.1226>
- [26] Toner J and Tu Y 1995 *Phys. Rev. Lett.* **75**(23) 4326–4329 URL <https://link.aps.org/doi/10.1103/PhysRevLett.75.4326>
- [27] Toner J and Tu Y 1998 *Phys. Rev. E* **58**(4) 4828–4858 URL <https://link.aps.org/doi/10.1103/PhysRevE.58.4828>
- [28] Chat   H, Ginelli F, Gr  goire G and Raynaud F 2008 *Phys. Rev. E* **77**(4) 046113 URL <https://link.aps.org/doi/10.1103/PhysRevE.77.046113>
- [29] Ginelli F 2016 *The European Physical Journal Special Topics* **225** 2099–2117 ISSN 1951-6401 URL <http://dx.doi.org/10.1140/epjst/e2016-60066-8>
- [30] Nishiguchi D, Nagal K, Chat   H and Sano M 2017 *Phys. Rev. E* **95** 020601(R) URL <https://doi.org/10.1103/PhysRevE.95.020601>
- [31] Gr  goire G and Chat   H 2004 *Phys. Rev. Lett.* **92**(2) 025702 URL <https://link.aps.org/doi/10.1103/PhysRevLett.92.025702>
- [32] Angelini T E, Hannezo E, Trepat X, Marquez M, Fredberg J J and Weitz D A 2011 *Proceedings of the National Academy of Sciences* **108** 4714–4719 (Preprint <http://www.pnas.org/content/108/12/4714.full.pdf>) URL <http://www.pnas.org/content/108/12/4714.abstract>
- [33] Zehnder S M, Wiatt M K, Uruena J M, Dunn A C, Sawyer W G and Angelini T E 2015 *Phys. Rev. E* **92**(3) 032729 URL <http://link.aps.org/doi/10.1103/PhysRevE.92.032729>
- [34] Tjhung E and Berthier L 2016 *ArXiv e-prints* (Preprint 1607.01734)
- [35] Cates M E and Tailleur J 2015 *Annual Review of Condensed Matter Physics* **6** 219
- [36] Levis D and Berthier L 2014 *Phys. Rev. E* **89**(6) 062301 URL <https://link.aps.org/doi/10.1103/PhysRevE.89.062301>
- [37] Farrell F D C, Marchetti M C, Marenduzzo D and Tailleur J 2012 *Phys. Rev. Lett.* **108**(24) 248101 URL <https://link.aps.org/doi/10.1103/PhysRevLett.108.248101>
- [38] Frittoli E, Palamidessi A, Marighetti P, Confalonieri S, Bianchi F, Malinverno C, Mazzarol G, Viale G, Martin-Padura I, Garre M, Parazzoli D, Mattei V, Cortellino S, Bertalot G, Fiore P D and Scita G 2014 *Journal of Cell Biology* **206** 307
- [39] Mendoza P, Diaz J, Silva and Torres V 2014 *Small GTPases* **5** e28195
- [40] Zhao Z, Liu X, Wu H, Zou S, Wang J, Ni P, Chen X and Fan Q 2010 *Cancer Science* **101** 1454
- [41] Sadati M, Nourhani A, Fredberg J J and Taheri Qazvini N 2014 *Wiley Interdisciplinary Reviews: Systems*

- Biology and Medicine* **6** 137–149 ISSN 1939-005X URL <http://dx.doi.org/10.1002/wsbm.1258>
- [42] Malinverno C, Corallino S, Giavazzi F, Bergert M, Li Q, Leoni M, Disanza A, Frittoli E, Oldani A, Martini E, Lendenmann T, Deflorian G, Beznoussenko G V, Poulidakos D, Ong K H, Uroz M, Trepas X, Parazzoli D, Maiuri P, Yu W, Ferrari A, Cerbino R and Scita G 2017 *Nat Mater* **advance online publication** – ISSN 1476-4660 URL <http://dx.doi.org/10.1038/nmat4848>
- [43] Vicsek T and Zafeiris A 2012 *Physics Reports* **517** 71 – 140 ISSN 0370-1573 collective motion URL <http://www.sciencedirect.com/science/article/pii/S0370157312000968>
- [44] Szabó B, Szóllósi G J, Gónci B, Jurányi Z, Selmeczi D and Vicsek T 2006 *Phys. Rev. E* **74**(6) 061908 URL <http://link.aps.org/doi/10.1103/PhysRevE.74.061908>
- [45] Henkes S, Fily Y and Marchetti M C 2011 *Phys. Rev. E* **84**(4) 040301 URL <https://link.aps.org/doi/10.1103/PhysRevE.84.040301>
- [46] Abercrombie M and Heaysman J E 1953 *Experimental cell research* **5** 111–131
- [47] Xiao X, Geyer V F, Bowne-Anderson H, Howard J and Sbalzarini I F 2016 *Medical Image Analysis* **32** 157 – 172 ISSN 1361-8415 URL <http://www.sciencedirect.com/science/article/pii/S1361841516300068>
- [48] Ramaswamy, S, Aditi Simha, R and Toner, J 2003 *Europhys. Lett.* **62** 196–202 URL <https://doi.org/10.1209/epl/i2003-00346-7>
- [49] Nambu Y 1960 *Phys. Rev.* **117**(3) 648–663 URL <https://link.aps.org/doi/10.1103/PhysRev.117.648>
- [50] Goldstone J 1961 *Il Nuovo Cimento (1955-1965)* **19** 154–164 ISSN 1827-6121 URL <http://dx.doi.org/10.1007/BF02812722>
- [51] Fily Y, Henkes S and Marchetti M C 2014 *Soft Matter* **10**(13) 2132–2140 URL <http://dx.doi.org/10.1039/C3SM52469H>
- [52] Ginelli F and Henkes S 2017 *In preparation*
- [53] Debnath J, Muthuswamy S and Brugge J 2003 *Methods* **30** 256
- [54] Muthuswamy S, Li D, Lelievre S, Bissell M and Brugge J 2001 *Nature Cell Biology* **3** 785
- [55] Seton-Rogers S, Lu Y, Hines L, Koundinya M, LaBaer J, Muthuswamy S and Brugge J 2004 *Proceedings of the National Academy of Sciences of the United States of America* **101** 1257
- [56] Adda-Bedia M, Katzav E and Vella D 2008 *The Journal of Chemical Physics* **129** 144506 (*Preprint* <http://dx.doi.org/10.1063/1.2991338>) URL <http://dx.doi.org/10.1063/1.2991338>
- [57] Bi D, Yang X, Marchetti M C and Manning M L 2016 *Phys. Rev. X* **6**(2) 021011 URL <https://link.aps.org/doi/10.1103/PhysRevX.6.021011>
- [58] Giavazzi F, Paoluzzi M, Macchi M, Manning L, Scita G, Cerbino R and Marchetti C *In preparation*
- [59] Bi D, Lopez J H, Schwarz J M and Manning M L 2015 *Nature Physics* **11** 1074