A least microenvironmental uncertainty principle (LEUP) as a generative model of collective cell migration mechanisms.

Arnab Barua*, Josue M. Nava-Sedeño*, and Haralampos Hatzikirou

August 30, 2018

Abstract

Collective migration is commonly observed in groups of migrating cells, in the form of swarming or aggregation patterns. Mathematical models have been proven very useful in understanding the dynamics of collective cell migration. Such models leading to collective migration depend on how cells decide over their velocity upon their microenvironmental information. This cell decision-making is mediated by a variety of biophysical mechanisms. The latter can be either too complex or largely unknown. To circumvent the problem of knowing all the details of the underlying mechanisms, we need to identify a principle that serves as a generative model of collective migration independent from the biophysical details of the system. Employing the least microenvironmental uncertainty principle (LEUP), we construct a model where cells sense the microenvironmental velocity distribution and reorient towards the least microenvironmentally uncertain direction. Depending on the choice of parameters, for instance interaction strength and size of the local interaction neighborhood, a multitude of collective migration patterns can emerge such stable global polar aligned state, partial nematically-ordered state or even vortices. Finally, we apply this model to the study of the collective migration of spherical *Serratia marcescens* bacteria. We show that our mode vl satisfactory reproduces the experimentally observed collective vortical behavior of spherical bacteria.

1 Introduction

Collective movement of populations is observed in several biological systems at different scales, from massive migration of mammals (Bennett and Tang, 2007) to cells during embryogenesis (Sato et al. 2010). In these systems, cells which are able to propel themselves independently, start moving in a coordinated fashion once enough similar cells are brought together. Due to the relevance of many of these processes to human activity, as well as their pervasiveness, several models have been proposed for studying collective migration. Mechanistic models, in particular, incorporate the driving interactions between individuals

*Contributed equally

in the specific system modeled. It is clear that different types of individuals, especially across different spatial scales, synchronize their movements through different mechanisms. This results in a variety of models specific to a certain individual species (Bennett and Tang, 2007; Hernandez-Ortiz et al. 2005; Ginelli et al. 2015). However, it is not uncommon the partial knowledge of the exact biophysical details of these mechanisms.

To address the above, phenomenological models of collective migration have been proposed. Typically, these mathematical models introduce a phenomenological short-range bias every individual feels. In one of the most influencial collective migration models (Vicsek et al. 1995), the direction of movement of individuals changes towards the mean velocity of individuals in a local neighborhood, inducing long-range swarming at the population level. Such models can be further refined into "physical" models, where individual individual dynamics are dictated by a system of Langevin equations. In Langevin equation models, the reorientation of individual velocities is brought about by the existence of a local interaction potential, which is determined by neighboring cells. Collective migration can be achieved, for example, through a ferromagnetic-like interaction potential, which locally aligns individual velocities polarly, or a liquid-crystal-like interaction potential, promotes nematic alignment (Peruani et al. 2008).

Here, we view migration as an active *decision-making process* of single cell velocity, in terms of speed and direction. In the specific case of biological cells, velocity decision-making involves intracellular processes such as actin polymerization, receptor recruitment, or in bacteria flagellar motor reversal mechanisms, to name a few. It is generally accepted that many kinds of cells base their phenotypic decisions according to their microenvironment (Kobayashi 2010; Libby et al. 2007; Andrews and Iglesias, 2007; Perkins and Swain 2009 for a review). More formally, the probability of cells to display certain velocity is dependent on the probability of assessing their local microenvironment. It has been argued that an energcelly-efficient manner for cells to encode their microenvironmental information is by following microenvironmental entropy gradient-based decisions (Hatzikirou 2018). In other words, a cell may react to its environment by either conforming to the information present in its environment, or acting against it. This entropy-driven decision making principle has been termed least microenvironmental uncertainty principle (LEUP). If the microenvironmental information includes information about the states of other cells, this cell-decision making process can be considered as the result of "social pressure" (Castellano et al. 2009).

In this work, we present a Langevin model of swarming where cells can sense the velocity orientations of other cells in their surroundings. Cells act as Bayesian inferrers and change their own orientation to optimize their prior, according to environmental orientation information. Under these assumptions, cells reorient according to the entropy gradient of the environmental information. A parameter, named the sensitivity, controls the strength and directionality of the reorientation in relation to the local entropy gradient. We find that the system adopts a steady, polar-ordered state for negative values of the sensitivity, when the sensitivity is negative. Conversely, the system remains out of equilibtium, but partially nematic-ordered when the sensitivity is positive. Furthermore, we find that the qualitative behavior of the model depends on the values of the cell density, noise strength, sensitivity, and size of the interac-

tion neighborhood. Finally, we show that our model replicates the experimental behavior of a system of interacting spherical bacteria.

2 Cell velocity dynamics

Moving and interacting biological cells are modeled by a two-dimensional self-propelled particle model (SPP). In this model, $N \in \mathbb{N}$ cells move on a two-dimensional plane. The n-th cell is characterized by its position, $\vec{r}_n \in \mathbb{R}^2$, and an orientation $\theta_n \in [0, 2\pi) \subset \mathbb{R}$. Due to the small size of cells, it is assumed that viscous forces dominate. Cells are also assumed to move with a constant speed $v_n \in \mathbb{R}^+$, which may vary among cells. Changes in velocity are therefore only due to a change in direction. Changes in direction (reorientations) result from a local potential $U(\vec{r}_m, \theta_m) : \mathbb{R}^2 \times [0, 2\pi) \mapsto \mathbb{R}$ which depends on the positions and orientations of cells within a radius $R \in \mathbb{R}^+$. The bias of the cell to follow the potential gradient is regulated by a parameter $\beta \in \mathbb{R}$, called the sensitivity. Additionally, angular fluctuations occur due to a stochastic noise term $\xi_n(t) \in [0, 2\pi)$, where $t \in \mathbb{R}^+$ denotes time. The noise will be assumed to be a zero-mean, white noise term, which has the statistical properties $\langle \xi_n(t) \rangle = 0$ and $\langle \xi_n(t_1)\, \xi_m(t_2) \rangle = 2D\delta(t_1 - t_2)\, \delta_{nm}$, where t_1 and t_2 are two time points, $D \in \mathbb{R}^+$ is the angular diffusion, $\delta(t)$ is the Dirac delta, and δ_{nm} is the Kronecker delta. Taking everything into account, the stochastic equations of motion of the n-th cell read (Peruani et al. 2008)

$$\frac{\mathrm{d}}{\mathrm{d}t}\vec{r}_n = v_n \vec{v}(\theta_n) \tag{1a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\theta_n = -\beta \frac{\partial}{\partial \theta_n} U(\vec{r}_m, \theta_m) + \xi_n(t), \tag{1b}$$

where $\vec{v}(\theta_n)$ is the normalized velocity of the cell. A representation of the SPP model is shown in Fig. 1.

The interaction potential $U(\vec{r}_m, \theta_m)$, which dictates the reorientation dynamics of cells, needs to be specified. Biophysically, the potential should encompass steric effects, hydrodynamic interactions, chemotactic effects, and terms arising from internal cellular processes, for example, flagellar motor dynamics, actin polymerization, receptor dynamics, etc. Finding such a potential is a formidable task since not all of the mechanisms and interactions involved are known. To circumvent this problem, we associate velocity cell decision-making to a variational principle of particular information theoretic measures (Bialek, 2012), known as the least microenvironmental uncertainty principle (LEUP) as introduced in (Hatzikirou 2018).

2.1 Cell decision-making: Least microenvironmental uncertainty principle

The state of the n-th cell in this case is its orientation θ_n. The set of intrinsic states of other cells within its radius of interaction is given by $\Theta_n = \{\theta_m : \|\vec{r}_n - \vec{r}_m\| \leq R\}$. The cell reacts to the environmental information, Θ_n, by changing its own state, θ_n. The cell then acts as a Bayesian decision-maker, such that

$$P(\theta_n \mid \Theta_n) = \frac{P(\Theta_n \mid \theta_n)\, P(\theta_n)}{P(\Theta_n)},$$

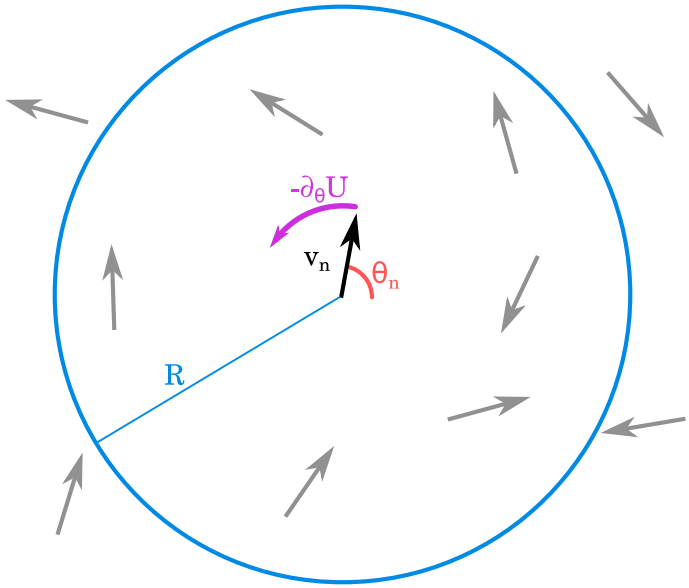


Figure 1: Sketch of the SPP model dynamics. The n-th cell is represented by a point particle with constant speed v_n and orientation θ_n. Depending on the form of the interaction potential, the cell may feel a reorientation force $-\partial_\theta U$ due to interaction with other cells inside a neighborhood with radius R.

where $P(\Theta_n \mid \theta_n)$ can be interpreted as the probability that the cell perceives all other cells in its surroundings, and $P(\theta_n)$ is the probability distribution of the cell's orientation (or prior). However, sensing other cells and evaluating $P(\Theta_n \mid \theta_n)$ entails an energy cost. It is reasonable to assume that the cell will try to optimize its prior $P(\theta_n)$ for the sake of energetic frugality.

Using statistical mechanical arguments, the problem is equivalent to finding the prior that minimizes the entropy of the cell with its surroundings. Let $S(\theta_n, \Theta_n)$ be the entropy of the cell-surroundings system, $S(\theta_n)$ the internal entropy of the cell, and $S(\Theta_n \mid \theta_n)$ the entropy of the information sensed by the cell. The entropies are connected by the relation $S(\theta_n, \Theta_n) = S(\theta_n) + S(\Theta_n \mid \theta_n)$. The optimization problem is finding $P(\theta_n)$ that minimizes $S(\theta_n, \Theta_n)$, while making sure that $P(\theta_n)$ is normalized; in other words

$$\frac{\delta}{\delta P(\theta_n)} \left\{ S(\theta_n, \Theta_n) - \tilde{\beta} \left[\int P(\theta_n = \vartheta)\, S(\Theta_n \mid \theta_n = \vartheta)\, \mathrm{d}\vartheta - \bar{S}(\Theta_n \mid \theta_n) \right] - \lambda \left[\int P(\theta_n = \vartheta)\, \mathrm{d}\vartheta - 1 \right] \right\} = 0, \tag{2}$$

where $\frac{\delta}{\delta P(\theta_n)}$ is the functional derivative, $\bar{S}(\Theta_n \mid \theta_n)$ is the expected ensamble statistics, and λ and $\tilde{\beta}$ are Lagrange multipliers. Taking into account the relations among entropies, Eq. (2) yields

$$P(\theta_n = \vartheta) = \frac{e^{-\tilde{\beta} S(\Theta_n \mid \theta_n = \vartheta)}}{Z}, \tag{3}$$

where $Z = \int e^{-\tilde{\beta} S(\Theta_n \mid \theta_n = \vartheta)} \mathrm{d}\vartheta$ is a normalization constant, and $\tilde{\beta}$ is the respon-

siveness of the cell. Using Eq. (3), the internal entropy of the cell, defined as $S(\theta_n) = \int P(\theta_n = \vartheta) \ln P(\theta_n = \vartheta)\, \mathrm{d}\vartheta$, is given by

$$S(\theta_n) = -\tilde{\beta} S(\Theta_n \mid \theta_n) + \ln Z. \tag{4}$$

Using the relation between thermodynamic potentials, it is evident that the internal energy is given by

$$U(\Theta_n, \theta_n) = -S(\Theta_n \mid \theta_n), \tag{5}$$

and the Helmholtz free energy is

$$F = -\frac{1}{\tilde{\beta}} \ln Z. \tag{6}$$

The internal energy depends on the orientation of the cell, as well as the orientations of other cells in the surroundings. Thus, it plays the role of the required interaction potential in the equations of motion. In this case, the potential depends mainly on the cells' intrinsic states, i.e. on their orientations. By doing so, it is evident that the responsiveness of the cells to LEUP, and the sensitivity in the equations of motion are the same, hence $\beta = \tilde{\beta}$. By substituting Eq. (5) into Eq. (1b), one obtains the equations of motion of the model

$$\frac{\mathrm{d}}{\mathrm{d}t}\vec{r}_n = v_n \vec{v}(\theta_n) \tag{7a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\theta_n = \beta \frac{\partial}{\partial \theta_n} S(\Theta_n \mid \theta_n) + \xi_n(t). \tag{7b}$$

Without loss of generality, it will be assumed that the orientations of cells within the interaction neighborhood are distributed according to

$$P(\vartheta \in \Theta_n \mid \theta_n) = \frac{\sinh \gamma}{2\pi \left[\cosh(\gamma) - \cos(\vartheta - \mu)\right]}, \tag{8}$$

where μ is the mean of the distribution and γ is a parameter related to the variance of neighbouring cell velocities. This is a wrapped Cauchy distribution, periodic over the interval $[-\pi, \pi]$. Please note that the results are independent of this choice (see SI K) According to this distribution, the entropy is

$$S(\Theta_n \mid \theta_n) = \ln(2\pi) + \ln\left(1 - e^{-2\gamma}\right). \tag{9}$$

The parameter γ depends on the local polar order (i.e. the degree of parallel alignment) of cell velocities in the neighborhood. Before defining γ, we will first define the observables characterizing the order of the velocity field.

2.2 Collective migration observables

Let us define the normalized complex velocity of the n-th cell, $z_n \in \mathbb{C}$ as $z_n = e^{i\theta_n}$, where i is the imaginary unit. The k-th moment of the velocity over an area A is given by $\langle z^k \rangle_A = \frac{1}{N_A} \sum_{m \in A} z_m^k$, where the sum is over all cells in area A, and N_A is the total number of cells in A. The polar order parameter in the area A is given by

$$S_A^1 = |\langle z \rangle_A|, \tag{10}$$

which is the modulus of the first moment of the complex velocity in A, while the nematic order parameter in the area A is given by

$$S_A^2 = \left|\left\langle z^2 \right\rangle_A\right|, \tag{11}$$

which is the modulus of the second moment of the complex velocity in A. The order parameters are bounded, i.e.

$$0 \leq S_A^1, S_A^2 \leq 1, \tag{12}$$

due to the complex velocities z_n being normalized. The parameter γ for the distribution of orientations in the neighborhood of the n-th cell is given by

$$\gamma = -\ln\left(S^1_{C_{R,n}}\right), \tag{13}$$

where the subindices $C_{R,n}$ indicate a circular area of radius R centered at $\vec{r}_n$.

While global polar and/or nematic order are characteristic of steady flows, rotating flow fields are commonly observed in out-of-equilibrium systems. The vorticity is an observable which is equal to twice the local angular velocity, and is thus a measure of the local strength and direction of rotation of the field. The vorticity ω is defined as

$$\omega\left(\vec{r}\right) = \left[\nabla \times \vec{v}_{\text{mean}}\left(\vec{r}\right)\right] \cdot \vec{k}, \tag{14}$$

where $\vec{v}_{\text{mean}}\left(\vec{r}\right)$ is the mean velcity field at point $\vec{r}$, and $\vec{k}$ is vector normal to the plane where cells move.

3 Ascending $(\beta > 0)$ or descending $(\beta < 0)$ the entropy gradient

By using the LEUP, we have modeled interaction as a reorientation dictated by the local entropy gradient. However, the specific response of cells to the entropy gradient is mediated by the parameter β. The absolute value $|\beta|$ is proportional to the likelihood of the cell to reorient according to a given entropy gradient. More importantly, the sign of β determines the specific reaction of cells towards the entropy gradient. If $\beta < 0$, cells tend to go against the entropy gradient towards the entropy minimum, restricting the distribution of angles to a narrow section. Conversely, $\beta > 0$ forces cells to follow the entropy gradient towards the entropy maximum, broadening the angle distribution.

To evaluate the effect of these two opposite migration strategies, we analyze the steady states in the two parameter regimes. By expanding $S^1_{C_{R,n}}$ using Eq. (10), defining the components of the mean neighborhood velocity as $\bar{v}_{y,n} = \sum_{C_{R,n} \ni m \neq n} \sin\theta_m$ and $\bar{v}_{x,n} = \sum_{C_{R,n} \ni m \neq n} \cos\theta_m$, and differentiating Eq. (9), we find that the orientation of θ_n at the entropy extrema must be such that (see Supporting information)

$$\tan\theta_n = \frac{\bar{v}_{y,n}}{\bar{v}_{x,n}},$$

but $\frac{\bar{v}_{y,n}}{\bar{v}_{x,n}} = \tan\bar{\theta}$, the tangent of the mean orientation of the neighbors, excluding the n-th cell. This results in two extremum points $\theta_n = \bar{\theta}$ and $\theta_n = \bar{\theta} + \pi$,

one where the velocity of the n-th cell is parallel to the average velocity of its neighbors, and one when it is antiparallel. In the first case

$$\sin\theta_n \propto \bar{v}_{y,n} \text{ and} \tag{15a}$$
$$\cos\theta_n \propto \bar{v}_{x,n}, \tag{15b}$$

while in the second case

$$\sin\theta_n \propto -\bar{v}_{y,n} \text{ and} \tag{16a}$$
$$\cos\theta_n \propto -\bar{v}_{x,n}. \tag{16b}$$

It can be shown (see Supporting information) that $\theta_n = \bar{\theta}$ corresponds to an entropy minimum, while $\theta_n = \bar{\theta} + \pi$ corresponds to an entropy maximum. Consequently, the behavior of **the regime $\beta < 0$ is analogous to that of the Vicsek model** (Vicsek et al. 1995). Conversely, the **regime $\beta > 0$ corresponds to a nematic analog of the Vicsek model**.

Next, let us assume that the model has a steady state, where the Helmholtz free energy per cell is given by Eq. (6). Due to its extensivity, the Helmholtz free energy of complete, non-interacting, steady state system is

$$F_T \approx -\frac{1}{\beta}\sum_{n=1}^{N} \ln Z_n = -\frac{1}{\beta}\ln\left(\prod_{n=1}^{N} Z_n\right),$$

where Z_n is the normalization constant of Eq. (3) for the n-th cell. The effective normalization constant $Z_T \coloneqq \prod_{n=1}^{N} Z_n$ is given by

$$Z_T = \int e^{-\beta\sum_{n=1}^{N}\left[\ln(2\pi)+\ln\left(1-e^{-2\gamma_n}\right)\right]} \mathrm{d}\vartheta_n. \tag{17}$$

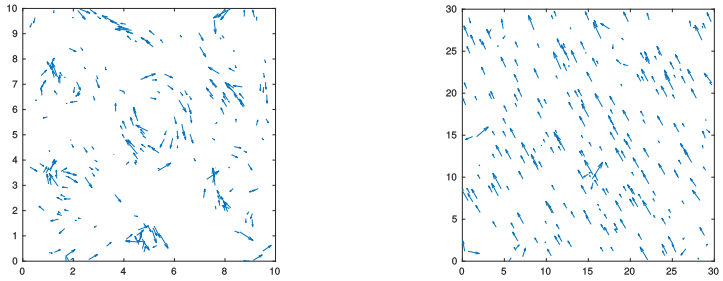
Integrating and susbtituting the resulting Z_T into Eq. (6) (see Supporting information), yields the Helmoltz free energy

$$F = N\left[\left(1-\frac{1}{\beta}\right)\ln(\gamma_n) + \ln(4\pi) + \frac{\ln(1-\beta)}{\beta}\right]. \tag{18}$$

Eq. (18) is well-defined only for $\beta < 1$. This indicates that no steady state exists for $\beta \geq 1$, hinting at an out-of-equilibrium regime (Rebenshtok et al. 2014). The present model belongs to the class of models with logarithmic potentials (see Eqs. (5) and (9)). The existence of a non-normalizable state in certain parameter regimes is a staple of systems with logarithmic potentials (Kessler and Barkai, 2010).

4 Spectrum of collective cell migration patterns for different parametric regimes

The model was implemented computationally to characterize the model and the effects of the different parameters on the resulting macroscopic behavior. The general qualitative behavior of the model can be observed in Fig. 2. In the regime $\beta < 0$, cells tend to travel in a single direction after some time has elapsed, similar to the Vicsek model. Conversely, in the $\beta > 0$ regime,



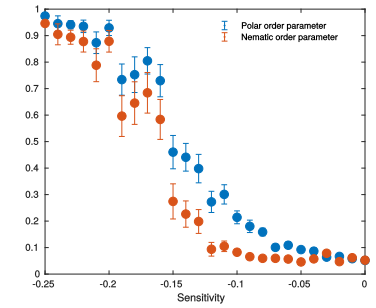
(a) $\beta > 0$ (b) $\beta < 0$

Figure 2: Simulation snapshots of the velocity field at long times. Arrows show the direction and magnitude of the velocity field. The snapshots were taken after 200 time steps. 300 cells were simulated, with an interaction radius to 3, and noise standard deviation equal to 0.1. Speeds of all cells were Rayleigh distributed with parameter $\sigma = 16.6$. In (a) the value of the sensitivity was 60 while in (b) the sensitivity was -2. Periodic boundaries were employed.

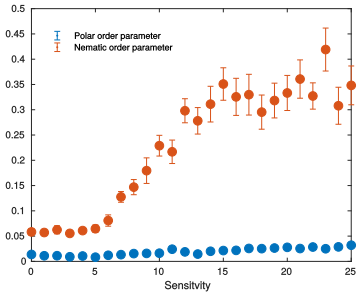
cells are seen to move collectively in transient vortice-like structures, even after long times have elapsed. Qualitatively, the patterns resulting from different parameter combinations are summarized in Table 1. Analyzing simualtions, two important phenomena are observed. First, there is a critical value of the interaction radius, R_C, which separates two regimes with qualitatively different behavior. Specifically, when $\beta > 0$, *no structures are formed for values of the interaction radius below* R_C. This indicates that medium-to-long range spread of information is necessary for ordering in this regime. Secondly, *patterns are also dependent on the choice of speed distribution* among cells. Heuristically, a broad speed distribution divides the population into fast and slow cells. While fast cells are useful for spreading information (and therefore, increasing the effective interaction range), slow cells are necessary for maintaining local ordering.

Furthermore, we quantitatively characterized global ordering at long times. The global polar order parameter, given by Eq. (10), for the complete simulation domain, measures the global degree of polar alignment, or polarization. The global nematic order parameter, given by Eq. (11) for the complete simulation domain, measures the tendency of all cells to align nematically, or along a single axis. These order parameters take a value of one when there is global order, while taking a value of zero when the system is completely disordered. It should be noted that polar order implies nematic order, but the reverse is not true.

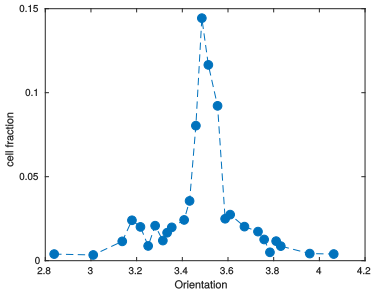
Similarly to other velocity alignment models (Peruani et al. 2008), the model shows an order-disorder transition with increasing noise amplitude and decreasing density (supplemental figure Fig. S1). More importantly, we observe that in the regime $\beta < 0$ the system also undergoes a transition towards polar order with decreasing β. After the transition, most cells have a similar orientation (Figs. 3a and 3c). In the regime $\beta > 0$, a phase transition is also observed towards nematic order with increasing β. In this case, however, the nematic ordering is not perfect, as evidenced by the nematic order parameter reaching values of around 0.25 after transition (Fig. 3b) compared to the value of 0.9 of the polar order parameter after transition in the $\beta < 0$ regime. This is further



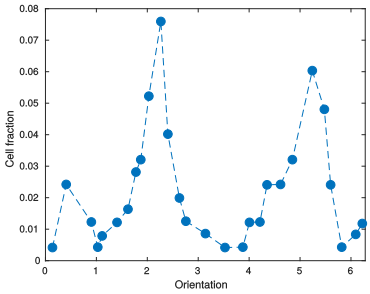
(a) $\beta < 0$



(b) $\beta > 0$



(c) $\beta < 0$



(d) $\beta > 0$

Figure 3: Order-disorder phase transitions and orientation distributions in two parameter regimes. (a) In the regime $\beta < 0$, a phase transition towards polar order occurs at a critical value of the sensitivity. (c) After the phase transition, polar order arises, and all cells have roughly the same orientation. (b) In the regime $\beta > 0$, the phase transition towards nematic order occurs at critical value of the sensitivity. (d) There is partial nematic order after the phase transition. Accordingly, several cells have opposite orientations. (a) and (b) The number of cells was fixed at 250, noise standard deviation equal to 0.05, and interaction radius at 3. Values of the order parameters were averaged over 20 realizations after 1000 time steps. (c) and (d) The number of cells was fixed at 250, noise standard deviation equal to 0.01 , and interaction radius at 4. The histogram was created with data from 20 realizations after 1000 time steps.

Table 1: Qualitative description of the observed patterns for different sensitivity and interaction radius regimes, as well as speed distributions.

$\beta < 0$	Speed distribution	$R < R_C$	$R > R_C$
	Delta distribution	Polar aligned streets of cells	Scattered polar aligned cells
	Uniform distribution	Compact polar aligned cluster	Compact polar aligned cluster
	Rayleigh distribution	Compact polar aligned cluster	Compact polar aligned cluster
$\beta > 0$	Speed distribution	$R < R_C$	$R > R_C$
	Delta distribution	No order or patterns	Nematic streaming
	Uniform distribution	No order or patterns	Nematic streaming and vorticules
	Rayleigh distribution	No order or patterns	Vortices

evidenced by the bimodal distribution of orientations with peak separation of approximately π radians (Fig. 3d). These simulation results further corroborate the theoretical results from Section 3. Finally, in the Supplementary Information we show that the phase transitions are robust under system variations.

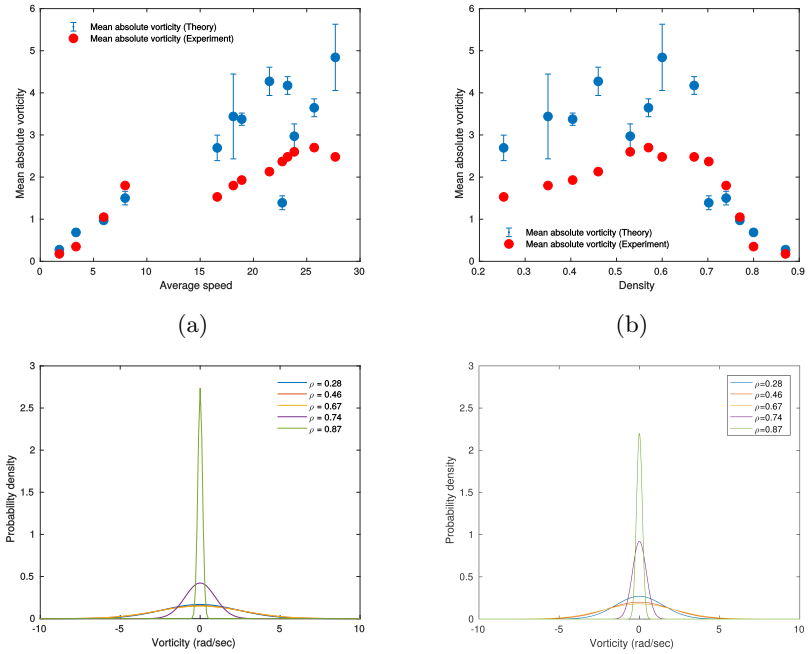
5 Application: collective behavior of spherical bacteria

Collective motion of bacteria has been extensively studied in vitro and in slilico. Most studies have focused on the collective properties of *S. enterica*, *E. coli*, and *M. xanthus*. These species of bacteria are similar since they exhibit a high aspect ratio. It has been shown that volume exclusion, coupled with a high aspect ratio, is sufficient to induce velocity alignment in the system (Peruani et al. 2006), and accordingly, ordered clusters of bacteria are observed at high densities. On the other hand, for round bacteria was mathematically shown that do not have the capacity to induce collective migration.

However, it has been recently shown (Rabani et al. 2013) that also spherical bacteria do display collective migration. The biophysical mechanism whereby spherical bacteria interact with one another must be different from the high body ratio/volume exclusion mechanism proposed for other bacterial species.

An important aspect to consider is the bacterial speed v_n. It was found experimentally (Rabani et al. 2013) that bacterial speed follows a Rayleigh distribution, dependent on bacterial density. Accordingly, the speed v_n for every cell in our model was Rayleigh-distributed matching the reported experimental distributions. Simulation snapshots at long times were characterized by their vorticity. The vortical behavior of our model was compared to the experimental vortical behavior reported in Rabani et al. 2013.

As can be appreciated in Fig. 4, our model qualitatively and quantitatively



(a) (b)

(c) Vorticity distribution (simulations) (d) Vorticity distribution (experiment)

Figure 4: Comparison between vorticity trends in experiments and in simulations. (a) Relation between spatially averaged speed and spatially averaged absolute value of the vorticity. The simulation values shown are the mean values averaged over three realizations. (b) Dependence of the spatially averaged absolute value of vorticity on the density. The simulation values shown are the mean values averaged over three realizations (c) Probability density function of the vorticity obtained from simulations for various densities. (d) Probability densitiy function of the vorticity as obtained from experiments for various densities. Experimental values were taken from Rabani et al. (2013). cell speed v_n was stochastic and Rayleigh distributed with distribution parameter γ depending on density (see Supplementary material), following the distributions reported in Rabani et al. (2013). Throughout all simulations, noise was set to $\langle \xi_n(t)^2 \rangle = 0.1$, interaction radius at $R = 10$, and sensitivity at $\beta = 60$. Data was obtained after 500 time steps.

satisfactory reproduces the vorticty behavior of the experimental system. Interestingly, the behavior of the experimental system was replicated for high values of the sensitivity β, and large interaction radii R. Such values of the sensitivity and interaction radii indicate far-reaching, strong tendencies of bacteria to reorient and migrate differently from their neighbours. Considering that *S. marcescens* is an example of a spherical, rear-propelled particle (Ishikawa 2009), our results agree with previous findings indicating that *S. marcescens* interacts through long-range hydrodynamics (Steager et al. 2008). Spherical, rear-propelled cells have been shown to destroy polar order as a result of hydrodynamic interactions (Evans et al. 2011), similarly to our model. It should be noted that our model tends to overestimate the value of the vorticity in certain regimes. This could be an artifact of the constant speed of cell cells in our model. While in the real system strong hydrodynamic interactions could result in significant deacceleartion of bacteria, in our model an analogous situation would result in sharp turns.

6 Discussion

In this work, we have introduced an off-lattice model of LEUP-induced collective migration, based on the self-propelled particles modeling framework. It was assumed that cells moved with constant speed, and changes in velocity were only due to changes in their orientation. Reorientation is governed by a stochastic differential equation depending on a white noise term and a force arising from an interaction potential.

The exact form of the interaction potential is very complex, and its specific form is dependent on particular details of the modeled system. While it has been shown that, in general, interactions between cells can effectively drive the entropy of the entire system towards an extremum point (Crosato et al. 2018; Großman et al. 2013), here we do the opposite. Instead of modeling the interaction potential biophysically, it was assumed that cells followed the LEUP, which dictates that cells change their internal states in order to minimize the uncertainty of the internal states of cells in their surroundings. In this model, orientations were considered the only internal states. Under this assumption, cells reorient either towards or against the gradient of entropy of the orientational distribution of cells in their neighbourhood, depending on the sign of the sensitivity parameter, which also dictates the strength of the interaction. The orientational distribution in the neighbourhood was assumed to be wrapped Cauchy distributed. Such a distribution facilitates the mathematical analysis of the model. However, the usage of other wrapped distributions does not qualitatively change the general behavior of the model (see Supporting information).

We show that, when the sensitivity parameter is negative, the model produces steady-state polar alignment patterns. Interestingly, we showed that the famous Vicsek model is a special case of LEUP. Conversely, when the parameter is positive, cells tend to reorient against the mean velocity of their neighbourhood. In this regime, the free energy diverges, indicating an out-of-equlibrium parameter regime. This kind of parameter-dependent dichotomy is similarly observed in systems with logarithmic potentials (Dechant et al. 2011), involved in processes such as long range-interacting gases (Bouchet and Dauxois, 2005), optical lattices (Lutz 2004), and DNA denaturation (Bar et al. 2008). The di-

chotomy arises from the logarithmic form of the entropy driving interaction in our model. It has been shown that, due to the non-normalizability of the steady state solution, such systems require a time-dependent expression for their analysis (Kessler and Barkai, 2010). Therefore, an in-depth theoretical analysis of our model would require a similar multicellular, time-dependent expression of the angular probability densities.

Another interesting observation is that in the LEUP regime the emergence of alignment patterns is almost independent of the interaction radius. On the contrary, when $\beta > 0$ there exists a minimum interaction radius that allows patterns to be formed. Thus, in this out-of-equilibrium regime pattern formation requires from the agents/cells a great amount of local microenvironmental sampling, i.e. requires significant energy expenditures. Therefore, this should be related to cells that either live in energy/nutrient rich environments or microenvironmental sampling is extremely cheap. Definitely, this is a point that worths more attention.

As a proof of principle, we show that our model replicates the collective vortical behavior of spherical motile cells. Recently, the collective behavior of spherical bacteria been modelled as a combination of steric repulsion and hydrodynamic interactions (Lushi et al. 2014). Our study has shown that hydrodynamics and steric interactions induce long-range microenvironmental entropy maximization, which coincides with the $\beta > 0$ LEUP regime. This generalises the type of biophysical mechanisms requires to produce vortical patterns and the above are just a realization.

As already mentioned, we have made a series of assumptions which helped to simplify the model. Our model assumes a Gaussian, white noise term in the orientational SDE. This results in normal diffusive behavior in the absence of interactions. It has been observed experimentally, however, that in some conditions, cells perform Lévy walks resulting in superdiffusive behavior (Matthäus et al. 2009). By changing the distribution or time correlations of the noise (Chechkin and Klages, 2009; Nava-Sedeño et al. 2017), it would be possible to both replicate the non-Gaussian dynamics of single cells, and investigate the effect of single anomalous dynamics on the collective behavior.

We have also assumed that cell velocities are the only internal states relevant for reorientation, for simplicity and as a proof of concept of the LEUP principle. However, it is reasonable to think that other states, such as relative position, speed, or even the interaction radius, may be relevant to include when modeling specific systems. It remains to be seen how additional states may impact the order and emerging patterns of the system.

As stated above, LEUP circumvents the biophysical details of cell migration and allows to reproduce a plethora of collective migration patterns. For instance, we have analytically derived the polar and apolar alignment Vicsek models for LEUP arguments. In this sense, LEUP acts as a generative model for collective migration mechanisms. In particular, it acts as a probability generating function for the single cell velocity (here only angle) equilibrium distribution. This is particularly useful upon limited knowledge of such mechanisms, a problem called structural model uncertainty. Another advantage of LEUP is the mapping of biophysical mechanism combination to the $\beta > 0$ or $\beta < 0$ regimes. This allows for unifying the model analysis but for a better classification of migration mechanisms. Finally, known mechanisms or data could be easily integrated to our proposed framework by further constraining the LEUP dynamics.

References

[1] Andrews BW, Iglesias PA (2007) An information-theoretic characterization of the optimal gradient sensing response of cells. *PLOS Comput Biol*, 3(8), e153.

[2] Bar A, Kafri Y, Mukamel D (2008) Dynamics of DNA melting. *J Phys-Condens Mat*, 21(3), 034110.

[3] Bennett DA, Tang W (2006) Modelling adaptive, spatially aware, and mobile agents: Elk migration in Yellowstone. *Int J Geogr Inf Sci*, 20(9), 1039-1066

[4] Bialek, W. (2012) Biophysics: searching for principles. Princeton University Press.

[5] Bouchet F, Dauxois T (2005) Kinetics of anomalous transport and algebraic correlations in a long-range interacting system. *J Phys Conf Ser* 7, 34-47

[6] Castellano C, Fortunato S, Loreto V. (2009) Statistical physics of social dynamics. *Rev Mod Phys*, 81(2), 591.

[7] Chechkin AV, Klages R (2009) Fluctuation relations for anomalous dynamics. *J Stat Mech-Theory E*, 2009(03), L03002.

[8] Crosato E, Spinney RE, Nigmatullin R, Lizier JT, Prokopenko M (2018) Thermodynamics and computation during collective motion near criticality. *Phys Rev E*, 97(1), 012120.

[9] Dechant A, Lutz E, Barkai E, Kessler DA (2011) Solution of the Fokker-Planck equation with a logarithmic potential. *J Stat Phys*, 145(6), 1524-1545.

[10] Evans AA, Ishikawa T, Yamaguchi T, Lauga E (2011). Orientational order in concentrated suspensions of spherical microswimmers. *Phys Fluids*, 23(11), 111702.

[11] Ginelli F, Peruani F, Pillot MH, Chaté H, Theraulaz G, Bon R (2015) Intermittent collective dynamics emerge from conflicting imperatives in sheep herds. *P Natl Acad Sci USA*, 112(41), 12729-12734.

[12] Großmann R, Schimansky-Geier L, Romanczuk P (2013) Self-propelled particles with selective attractionrepulsion interaction: from microscopic dynamics to coarse-grained theories. *New J Phys*, 15(8), 085014.

[13] Hatzikirou H (2018) Statistical mechanics of cell decision-making: the cell migration force distribution. *J Mech Behav Mater*, 27(1-2)

[14] Hernandez-Ortiz JP, Stoltz CG, Graham MD (2005) Transport and collective dynamics in suspensions of confined swimming particles. *Phys Rev Lett*, 95(20), 204501.

[15] Ishikawa T (2009) Suspension biomechanics of swimming microbes. *J R Soc Interface*, rsif20090223.

[16] Kessler DA, Barkai E (2010) Infinite covariant density for diffusion in logarithmic potentials and optical lattices. *Phys Rev Lett*, 105(12), 120602.

[17] Kobayashi TJ (2010) Implementation of dynamic Bayesian decision making by intracellular kinetics. *Phys Rev Lett*, 104(22), 228104.

[18] Libby E, Perkins TJ, Swain PS (2007) Noisy information processing through transcriptional regulation. *P Natl Acad Sci USA*, 104(17), 7151-7156.

[19] Lushi E, Wioland H, Goldstein RE (2014) Fluid flows created by swimming bacteria drive self-organization in confined suspensions. *P Natl Acad Sci USA*, 201405698.

[20] Lutz E (2004) Power-law tail distributions and nonergodicity. *Phys Rev Lett*, 93(19), 190602.

[21] Matthäus F, Jagodič M, Dobnikar J (2009) E. coli superdiffusion and chemotaxis-search strategy, precision, and motility. *Biophys J*, 97(4), 946-957.

[22] Nava-Sedeño JM, Hatzikirou H, Klages R, Deutsch A (2017) Cellular automaton models for time-correlated random walks: derivation and analysis. *Sci Rep*, 7(1), 16952.

[23] Perkins TJ, Swain PS (2009) Strategies for cellular decisionmaking. *Mol Syst Biol*, 5(1), 326.

[24] Peruani F, Deutsch A, Bär M (2008) A mean-field theory for self-propelled particles interacting by velocity alignment mechanisms. *Eur Phys J-Spec Top* 157(1):111-122

[25] Rabani A, Ariel G, Be'er A (2013) Collective motion of spherical bacteria. *PLOS One*, 8(12), e83760.

[26] Rebenshtok A, Denisov S, Hänggi P, Barkai E (2014) Non-normalizable densities in strong anomalous diffusion: Beyond the central limit theorem. *Phys Rev Lett*, 112(11), 110601.

[27] Sato Y, *et al.* (2010) Dynamic analysis of vascular morphogenesis using transgenic quail embryos. *PLOS One*, 5(9), e12674.

[28] Steager EB, Kim CB, Kim MJ (2008) Dynamics of pattern formation in bacterial swarms. *Phys Fluids*, 20(7), 073601.

[29] Vicsek T, Czirók A, Ben-Jacob E, Cohen I, Shochet O (1995) Novel type of phase transition in a system of self-driven particles. *Phys Rev Lett* 75(6):1226

Supplementary information

A Calculation of the curvature of micro-environmental entropy and Helmholtz free energy

Initially, we expand $S^1_{C_{R,n}}$ using Eq. (10) and the definition of the complex velocity z_n

$$S^1_{C_{R,n}} = \frac{1}{N_{C_{R,n}}} \sqrt{\left(\sum_{m \in C_{R,n}} \sin\theta_m \right)^2 + \left(\sum_{m \in C_{R,n}} \cos\theta_m \right)^2}.$$

We define the quantities $\bar{v}_{y,n} = \sum_{C_{R,n} \ni m \neq n} \sin\theta_m$ and $\bar{v}_{x,n} = \sum_{C_{R,n} \ni m \neq n} \cos\theta_m$. Using these definitions, and expanding the square terms, the polar order parameter can be written as

$$S^1_{C_{R,n}} = \frac{1}{N_{C_{R,n}}} \sqrt{\bar{v}^2_{y,n} + \bar{v}^2_{x,n} + 2\left(\bar{v}_{y,n}\sin\theta_n + \bar{v}_{x,n}\cos\theta_n\right)}. \tag{S1}$$

Next, using Eq. (13), we substitute Eq. (S1) into Eq. (9) to obtain

$$\begin{aligned} S\left(\Theta_n \mid \theta_n\right) &= \ln\left(2\pi\right) \\ &+ \ln\left\{1 - \frac{1}{N^2_{C_{R,n}}}\left[\bar{v}^2_{y,n} + \bar{v}^2_{x,n} + 2\left(\bar{v}_{y,n}\sin\theta_n + \bar{v}_{x,n}\cos\theta_n\right)\right]\right\}. \end{aligned} \tag{S2}$$

As mentioned before, the cell will reorient towards the minimum of $S\left(\Theta_n \mid \theta_n\right)$ when $\beta < 0$, while reorienting towards the maximum when $\beta > 0$. To find the entropy maxima and minima, we start by finding the entropy extrema, where $\frac{\partial}{\partial\theta_n} S\left(\Theta_n \mid \theta_n\right) = 0$. Differentiating Eq. (S2) we have that, in an extremum point

$$\frac{-\frac{2}{N^2_{C_{R,n}}}\left(\bar{v}_{y,n}\cos\theta_n - \bar{v}_{x,n}\sin\theta_n\right)}{1 - \frac{1}{N^2_{C_{R,n}}}\left[\bar{v}^2_{y,n} + \bar{v}^2_{x,n} + 2\left(\bar{v}_{y,n}\sin\theta_n + \bar{v}_{x,n}\cos\theta_n\right)\right]} = 0,$$

which reduces to the condition

$$\bar{v}_{y,n}\cos\theta_n = \bar{v}_{x,n}\sin\theta_n. \tag{S3}$$

From here, it is evident that the orientation of θ_n at the entropy extrema must be such that

$$\tan\theta_n = \frac{\bar{v}_{y,n}}{\bar{v}_{x,n}},$$

but $\frac{\bar{v}_{y,n}}{\bar{v}_{x,n}} = \tan\bar{\theta}$, the tangent of the mean orientation of the neighbors, excluding the n-th cell. This results in two extremum points $\theta_n = \bar{\theta}$ and $\theta_n = \bar{\theta} + \pi$, one where the velocity of the n-th cell is parallel to the average velocity of its neighbors, and one when it is antiparallel. In the first case

$$\sin\theta_n \propto \bar{v}_{y,n} \text{ and} \tag{S4a}$$

$$\cos\theta_n \propto \bar{v}_{x,n}, \tag{S4b}$$

while in the second case

$$\sin\theta_n \propto -\bar{v}_{y,n} \text{ and} \tag{S5a}$$

$$\cos\theta_n \propto -\bar{v}_{x,n}, \tag{S5b}$$

To discriminate the maximum from the minimum point, the value of $\frac{\partial^2}{\partial\theta_n^2} S\left(\Theta_n \mid \theta_n\right)$ must be evaluated at each extremum point. The second derivative is given by

$$\frac{\partial^2}{\partial\theta_n^2} S\left(\Theta_n \mid \theta_n\right) = \frac{\frac{2}{N_{C_{R,n}}^2}\left(\bar{v}_{y,n}\sin\theta_n + \bar{v}_{x,n}\cos\theta_n\right)}{1 - \frac{1}{N_{C_{R,n}}^2}\left[\bar{v}_{y,n}^2 + \bar{v}_{x,n}^2 + 2\left(\bar{v}_{y,n}\sin\theta_n + \bar{v}_{x,n}\cos\theta_n\right)\right]} - \left\{\frac{\frac{2}{N_{C_{R,n}}^2}\left(\bar{v}_{y,n}\cos\theta_n - \bar{v}_{x,n}\sin\theta_n\right)}{1 - \frac{1}{N_{C_{R,n}}^2}\left[\bar{v}_{y,n}^2 + \bar{v}_{x,n}^2 + 2\left(\bar{v}_{y,n}\sin\theta_n + \bar{v}_{x,n}\cos\theta_n\right)\right]}\right\}^2.$$

In each extremum point, Eq. (S3) must hold, therefore

$$\left.\frac{\partial^2}{\partial\theta_n^2} S\left(\Theta_n \mid \theta_n\right)\right|_{\text{ext}} = \frac{\frac{2}{N_{C_{R,n}}^2}\left(\bar{v}_{y,n}\sin\theta_n + \bar{v}_{x,n}\cos\theta_n\right)}{1 - \frac{1}{N_{C_{R,n}}^2}\left[\bar{v}_{y,n}^2 + \bar{v}_{x,n}^2 + 2\left(\bar{v}_{y,n}\sin\theta_n + \bar{v}_{x,n}\cos\theta_n\right)\right]}. \tag{S6}$$

Defining κ as the proportionality constant relating $\sin\theta_n$ and $\cos\theta_n$ with $\bar{v}_{y,n}$ and $\bar{v}_{x,n}$, respectively, the second derivative evaluated at $\theta_n = \bar{\theta}$ is

$$\left.\frac{\partial^2}{\partial\theta_n^2} S\left(\Theta_n \mid \theta_n\right)\right|_{\theta_n=\bar{\theta}} = \frac{\frac{2\kappa}{N_{C_{R,n}}^2}\left(\bar{v}_{y,n}^2 + \bar{v}_{x,n}^2\right)}{1 - \left(S_{C_{R,n}}^1\right)^2} > 0, \tag{S7}$$

because the numerator is positive definite, and the denominator is positive following Eq. (12). The extremum point $\theta_n = \bar{\theta}$ therefore corresponds to an entropy minimum. Consequently, the behavior of the regime $\beta < 0$ is analogous to that of the Vicsek model (Vicsek et al. 1995). Conversely, at $\theta_n = \bar{\theta} + \pi$ we find that

$$\left.\frac{\partial^2}{\partial\theta_n^2} S\left(\Theta_n \mid \theta_n\right)\right|_{\theta_n=\bar{\theta}+\pi} = \frac{\frac{-2\kappa}{N_{C_{R,n}}^2}\left(\bar{v}_{y,n}^2 + \bar{v}_{x,n}^2\right)}{1 - \left(S_{C_{R,n}}^1\right)^2} < 0, \tag{S8}$$

uisng the same arguments as for the $\theta_n = \bar{\theta}$ point. Therefore, the point $\theta_n = \bar{\theta}+\pi$ corresponds to the entropy maximum. Then, the regime $\beta > 0$ corresponds to a nematic analog of the Vicsek model.

Next, let us assume that the model has a steady state, where the Helmholtz free energy per bacterium is given by $F = -\frac{1}{\beta}\ln Z$. Due to its extensivity, the Helmholtz free energy of the complete system is

$$F_T = -\frac{1}{\beta}\sum_{n=1}^{N}\ln Z_n = -\frac{1}{\beta}\ln\left(\prod_{n=1}^{N} Z_n\right),$$

where Z_n is the normalization constant of n-th cell(see Eq.(3) in paper).

The effective normalization constant $Z_T \coloneqq \prod_{n=1}^{N} Z_n$ is given by

$$Z_T = \int e^{-\beta\sum_{n=1}^{N}\left[\ln(2\pi)+\ln\left(1-e^{-2\gamma_n}\right)\right]}\mathrm{d}\vartheta_n. \tag{S9}$$

The integration is performed over the orientations of all cells in the system. Moreover, the dependency of each γ_n on all angles θ_n is complex and makes integration challenging. However, variation of θ_n for all n translates into a variation in all γ_n. Therefore, Eq. S9 is equivalent to

$$Z_T = \int e^{-\beta \sum_{n=1}^{N} \left[\ln(2\pi) + \ln\left(1 - e^{-2\gamma_n}\right)\right]} \mathrm{d}\gamma_n. \tag{S10}$$

Expanding up to linear terms around $\gamma_n = 0$ yields

$$Z_T = \int e^{-\beta \sum_{n=1}^{N} \left[\ln(2\pi) + \ln(2\gamma_n)\right]} \mathrm{d}\gamma_n,$$

which after rearranging terms and integrating reduces to

$$Z_T = \left[\frac{1}{(4\pi)^{\beta}} \frac{\gamma_n^{1-\beta}}{1-\beta}\right]^N. \tag{S11}$$

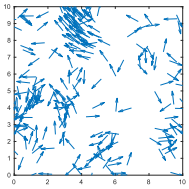
Substituting Eq. S11 into Eq. 6, and rearranging terms, yields the Helmholtz free energy

$$F = N\left[\left(1 - \frac{1}{\beta}\right)\ln\left(\gamma_n\right) + \ln(4\pi) + \frac{\ln\left(1-\beta\right)}{\beta}\right]. \tag{S12}$$

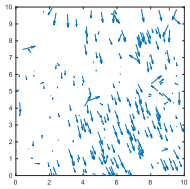
Eq. S12 is well-defined only for $\beta < 1$. This indicates that no steady state exists for $\beta \geq 1$, hinting at an out-of-equilibrium regime (Rebenshtok et al. 2014). The present model belongs to the class of models with logarithmic potentials. The existence of a non-normalizable state in certain parameter regimes is a staple of systems with logarithmic potentials (Kessler and Barkai, 2010).

B Pattern formation in different β regime (see Table 1)

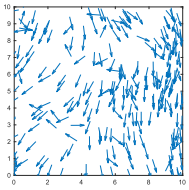
Here, we present a simulation snapshot for the different parameter constellations as defined in Table 1. To give a quantitative impression of each simulation, we define the range of the polar order parameter as P, the nematic order parameter as N and the mean absolute vorticity as V.



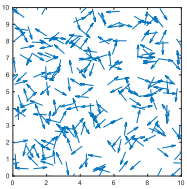
(a) Polar aligned street cells (P=0.6 to 0.7, N=0.5 to 0.7, V=0 to 0.04)



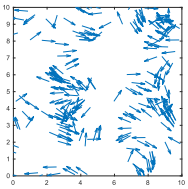
(b) Compact polar aligned cells (P=0.8 to 1.0, N=0.7 to 1.0, V=0 to 0.065)



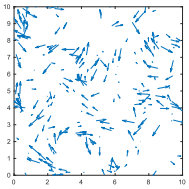
(c) Scattered polar aligned cells (P=0.7 to 1.0, N=0.7 to 1.0, V=0 to 0.05)



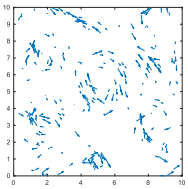
(d) No order or patterns (P=0 to 0.09, N=0 to 0.07, V=0 to 0.065)



(e) Nematic streaming (P=0 to 0.05, N=0.3 to 0.5, V=0 to 0.065)



(f) Nematic streaming and vorticules (P=0 to 0.03, N=0.2 to 0.4, V=0 to 0.04)



(g) Vortices (P=0 to 0.03, N=0.2 to 0.4, V=0.1 to 0.35)

C Finite size effects in polar order parameter versus sensitivity ($\beta < 0$)

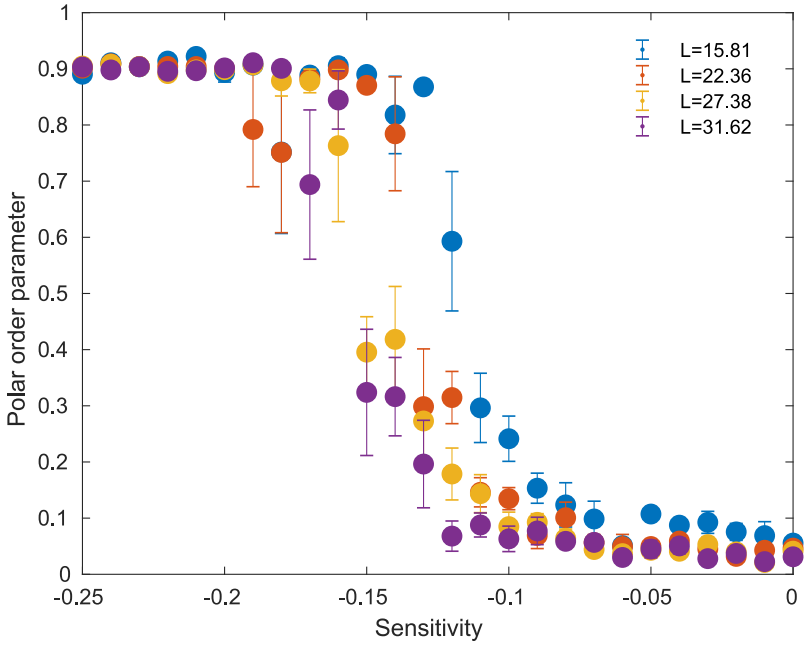


Figure S2: Polar order parameter vs. sensitivity graph with interaction radius is at 3, and standard deviation of noise is fixed at 0.1. Density is fixed at 1.0 and the system size is varied. The micro-environmental distribution has been taken from a wrapped Cauchy distribution. All order parameters were averaged over 5 realizations after 10^3 time steps.

D Finite size effects in polar order parameter vs. noise graph

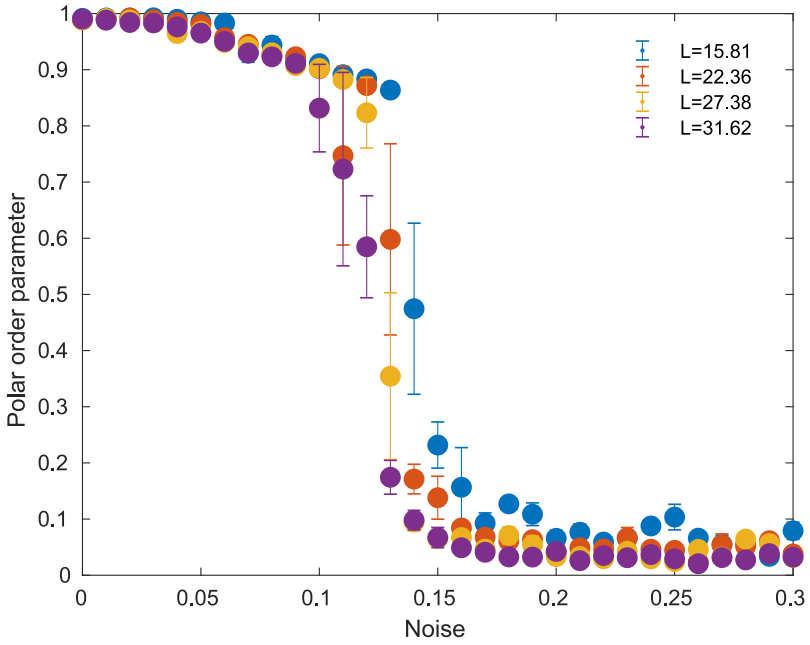


Figure S3: Polar order parameter vs. noise graph with interaction radius is at 3 and sensitivity is at -0.2. Density is fixed at 1.0 and the system size is varied. The micro-environmental distribution has been taken from a wrapped Cauchy distribution. All order parameters were averaged over 5 realizations after 10^3 time steps.

E Finite size effects in nematic order parameter for the positive sensitivity regime ($\beta > 0$)

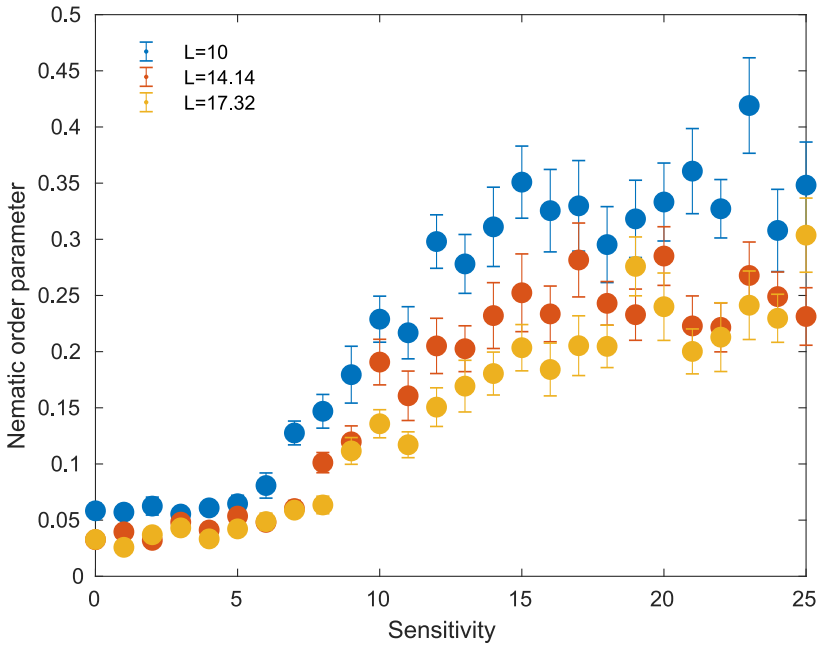


Figure S4: Nematic order parameter vs. sensitivity graph with interaction radius is at 3 and standard deviation of noise is at 0.05. Density is fixed at 2.5 and the system size is varied. The micro-environmental distribution has been taken from a wrapped Cauchy distribution. All order parameters were averaged over 20 realizations after 10^3 time steps.

F Finite size effects in nematic order parameter vs. noise graph

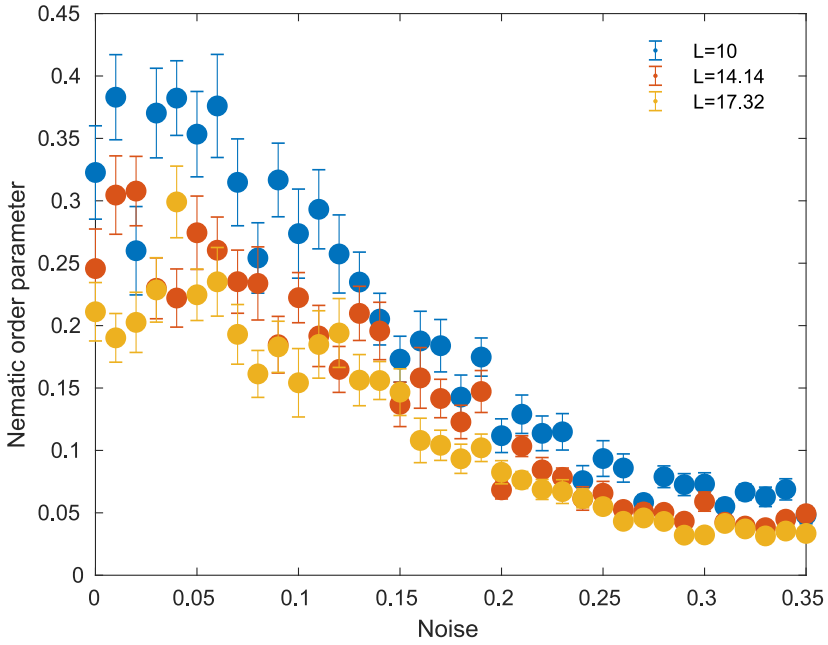


Figure S5: Nematic order parameter vs. noise graph with interaction radius is at 3 and sensitivity is at 20. Density is fixed at 2.5 and the system size is varied. The micro-environmental distribution has been taken from a wrapped Cauchy distribution. All order parameters were averaged over 20 realizations after 10^3 time steps.

G Polar and nematic order parameters vs. density graph

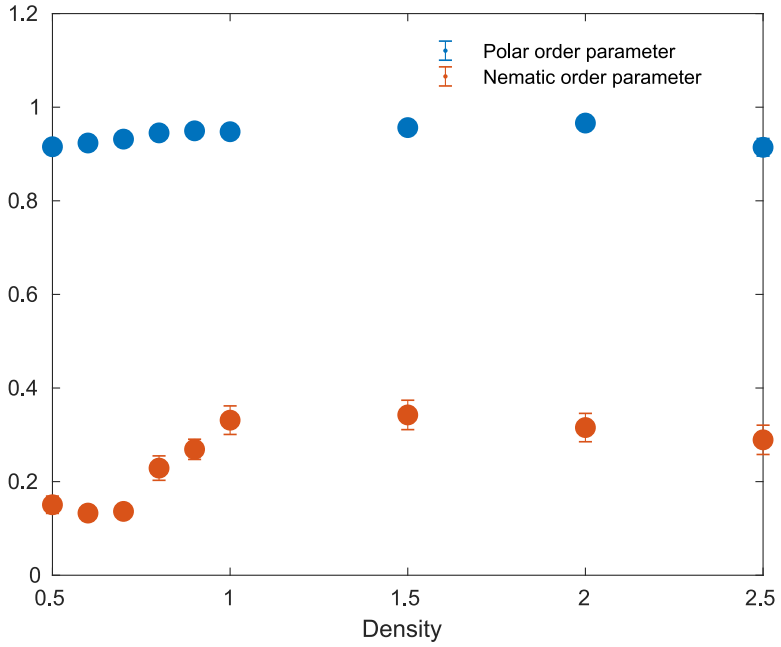


Figure S6: Polar order parameter and nematic order parameter vs. density graph with interaction radius is at 3 and sensitivity is fixed at 15 (for nematic order parameter) and -0.2 (for polar order parameter). The micro-environmental distribution has been taken from a wrapped Cauchy distribution. All order parameters were averaged over 20 realizations after 10^3 time steps.

H Finite size effects in mean absolute vorticity for the positive sensitivity regime ($\beta > 0$)

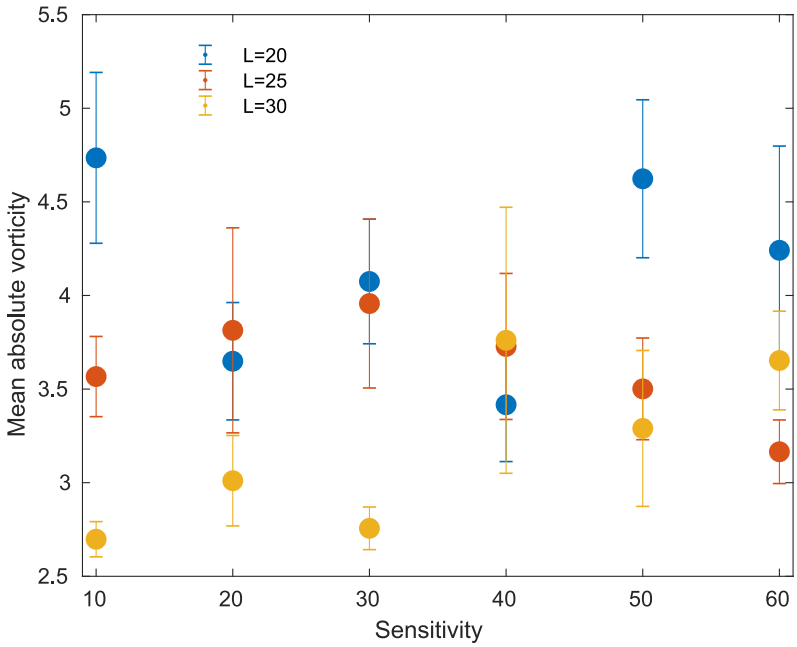


Figure S7: Mean absolute vorticity vs sensitivity graph. Density is fixed at 0.28 and the system size is varied. Interaction radius is fixed at 10, the variance of noise was 0.1. Here micro-environmental distribution has been taken from a wrapped Cauchy distribution. Values of the mean absolute vorticity were averaged over 15 realizations after 500 time steps.

I Finite size effects in mean absolute vorticity vs. interaction radius graph

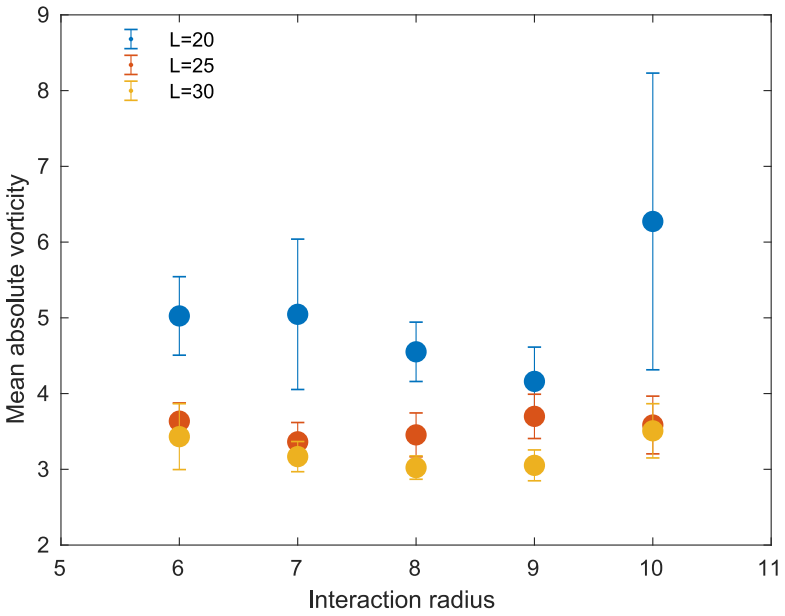


Figure S8: Mean absolute vorticity vs interaction radius graph. Density is fixed at 0.28 and the system size is varied. Sensitivity is fixed at 10 and the variance of noise at 0.1. Values of the mean absolute vorticity were averaged over 15 realizations after 500 time steps.

J Average speed vs density graph of spherical bacteria

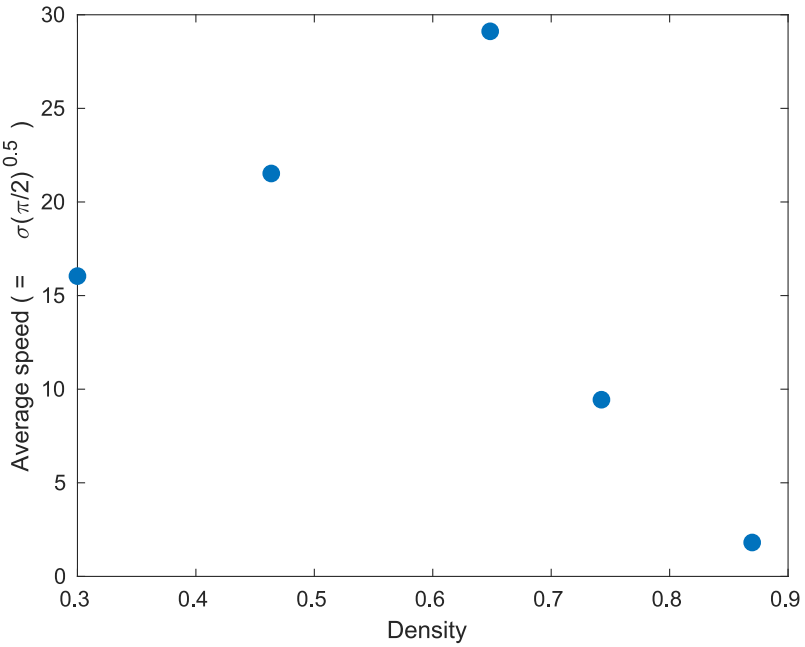


Figure S9: The average speed of spherical bacteria vs. density, where σ is the Rayleigh parameter. The graph was reproduced from the corresponding figure in Rabani et al. 2013.

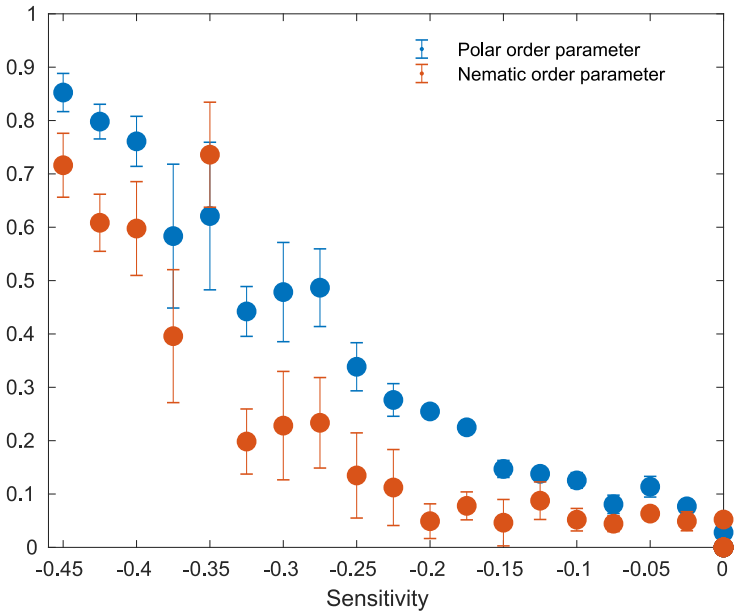


Figure S10: Polar and nematic order parameter vs. sensitivity ($\beta < 0$) graph. Interaction radius is fixed at 3, and standard deviation of noise at 0.05. Density is fixed at 2.5. Here micro-environmental entropy has been taken from a wrapped exponential distribution. All order parameters were averaged over 5 realizations after 250 time steps.

K Polar and nematic order parameter versus sensitivity for wrapped exponential distribution

The qualitative behavior of the phase transitions is independent to the choice of the microenvironmental distribution. To showcase this, we use the wrapped exponential distribution. Although minot quantitative changes are observed, the phase transitions for positive and negative sensitivities remain qualitatively invariant (see Figs S10 and S11).

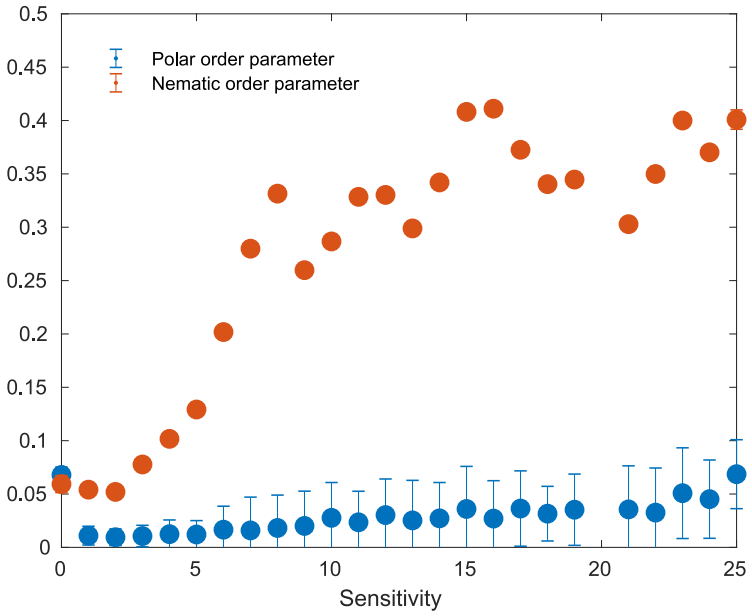


Figure S11: Polar order parameter and nematic order parameter vs. sensitivity ($\beta > 0$) graph. Interaction radius is fixed at 3, and standard deviation of noise at 0.05. Density is fixed at 2.5. Here micro-environmental entropy has been taken from a wrapped exponential distribution. All order parameters were averaged over 20 realizations after 10^3 time steps.