Entropy-driven cell decision-making predicts "fluid-to-solid" transition in multicellular systems

Arnab Barua[a,b,*], Simon Syga[b,*], Pietro Mascheroni[a], Nikos Kavallaris[c], Michael Meyer-Hermann[a,d], Andreas Deutsch[b], and Haralampos Hatzikirou[a,b,e]

[a] *Department of Systems Immunology and Braunschweig Integrated Centre of Systems Biology, Helmholtz Centre for Infection Research, Rebenring 56, 38106 Braunschweig, Germany*
[b] *Technische Univesität Dresden, Center for Information Services and High Performance Computing, Nöthnitzer Straße 46, 01062, Dresden, Germany*
[c] *Department of Mathematical and Physical Sciences, University of Chester, Chester*
[d] *Institute for Biochemistry, Biotechnology and Bioinformatics,Technische Universität Braunschweig, Braunschweig, Germany*
[e] *Mathematics Department, Khalifa University, P.O. Box: 127788, Abu Dhabi, UAE*

Cellular decision making allows cells to assume functionally different phenotypes in response to microenvironmental cues, with or without genetic change. It is an open question, how individual cell decisions influence the dynamics at the tissue level. Here, we study spatio-temporal pattern formation in a population of cells exhibiting phenotypic plasticity, which is a paradigm of cell decision making. We focus on the migration/resting and the migration/proliferation plasticity which underly the epithelial-mesenchymal transition (EMT) and the go or grow dichotomy. We assume that cells change their phenotype in order to minimize their microenvironmental entropy following the LEUP (Least microEnvironmental Uncertainty Principle) hypothesis. In turn, we study the impact of the LEUP-driven migration/resting and migration/proliferation plasticity on the corresponding multicellular spatio-temporal dynamics with a stochastic cell-based mathematical model for the spatio-temporal dynamics of the cell phenotypes. In the case of the go or rest plasticity, a corresponding mean-field approximation allows to identify a bistable switching mechanism between a diffusive (fluid) and an epithelial

*Contributed equally

(solid) tissue phase which depends on the sensitivity of the phenotypes to the environment. For the go or grow plasticity, we show the possibility of Turing pattern formation for the "solid" tissue phase and its relation with the parameters of the LEUP-driven cell decisions.

Keywords— Cell-decision making; Phenotypic plasticity; Langevin equations; Mean-field theory; Least microEnvironmental Uncertainty Principle (LEUP); Fluid-to-solid transition; Spatio-temporal patterns

1 Introduction

Cellular decision making allows cells to respond to microenvironmental cues [1]. For instance, if an external ligand from the microenvironment binds to the cell membrane, it may trigger a chain of biochemical reactions inside the cell [2]. After the microenvironmental information has been processed, the cell may respond by changing its properties implying cell differentiation, proliferation, migration, apoptosis etc. Thus, cellular decision making depends on intrinsic signal transduction pathways, the genetic cell network [3], extrinsic stimuli, and molecular noise [4]. Cell-decision making plays a key role in cell fate determination and the maintenance of phenotypic plasticity. In the former, cells irreversibly acquire new fates by following a hierarchical lineage, where pluripotent stem cells in a proper microenvironment (stem cell niche) differentiate into e. g. bone, muscle, epithelial and further specialised cells. On the other hand, phenotypic plasticity allows a reversible adaptation to different microenvironmental stimuli [5]. While cellular decision making is well-studied for single cells, it is an open question how individual cell decisions influence the dynamics at the tissue level.

In this paper, we analyse the implications of phenotypic plasticity on spatio-temporal pattern formation in the cell population and focus on the migration/resting and migration/proliferation plasticity. The migration/resting type of plasticity is observed during the epithelial-mesenchymal transition (EMT/MET)[6–8] while the migration/proliferation plasticity characterizes the go-or-grow (GoG) dichotomy found in different types of cancers, such as glioblastoma or breast cancer, but also during normal development [9–12]. Epithelial cells are immotile, connected by strong adhesion bonds, exhibit regular shapes and show strong apico-basal cell polarity. Epithelial and mesenchymal phenotypes can be interpreted as representing two migration modes, resting and motile, respectively. On the other hand, GoG refers to the mutual exclusion of cell migration and proliferation, leading to the existence of proliferative/immotile and nonproliferative/mobile cell phenotypes.

We assume that a Least microEnvironmental Uncertainty Principle (LEUP) dictates the phenotypic switching in both types of the considered phenotypic plastic behaviours. The Least microEnvironmental Uncertainty Principle idea is based on the idea of energy minimization in the context of cell decision making. In order to make phenotypic decisions, cells sense their surrounding by employing energetically expensive processes, such as polymerizing pseudopodia,

translocating receptor molecules or modifying the cytoskeleton according to mechanical signals [13–15]. However, the cost can be minimized if cells build informative priors about their microenvironment, i.e. encoding persistent microenvironemntal information [16]. At the same time, in many biological processes, such as development and wound healing, cell phenotypic decisions effectively promote the co-evolution of the cellular microenvironment into more organized states, leading to a decrease of the microenviromental entropy from high to low. The assumption of microenvironmenal entropy minimization is exploited by LEUP to define the dynamics of cell phenotypic decisions. The idea of viewing cells as Bayesian decision-makers under energetic constraints can be translated into a free-energy principle, implying a statistical mechanics theory for cell decision making [17]. According to this theory, cells try to optimize their phenotypic prior to reduce the joint entropy of their own phenotype and the microenvironmental state. Similar ideas using information theoretic measures have been proposed also in seminal works by W. Bialek [14]. Our LEUP-based low-dimensional mathematical model can be used to analyze collective effects of cell decision making which is the main goal of our paper. At this point, it is important to mention that LEUP is still a hypothesis which currently undergoes experimental validation.

The manuscript is structured as follows: In the next section, we define a LEUP-driven cell decision model based on the migration/resting plasticity. In Section 3, we develop a stochastic individual-based model (IBM) for moving and resting cells where cells can change their phenotype according to the LEUP-driven cell decision model and show the formation of aggregates of resting cells in corresponding simulations. In section 4, we derive a macroscopic mean-field approximation ofour microscopic model which allows to explain the formation of aggregates in the microscopic model. In section 5, we analyze the stability of the steady states and the pattern formation potential of the macroscopic go-or grow model. Finally, in section 6, we discuss the biological implications of our results in terms of multicellular growth and pattern control.

2 A minimal model of LEUP-driven cell migration plasticity

Here, we define a mathematical model for LEUP-based switching between moving and resting phenotypes. We assume that cells can sense their microenvironment and can change their own phenotype X_i on a domain $\mathcal{L} \subset \mathbb{R}^2$ accordingly, where $X_i = 0, 1$ corresponds to the resting and migrating state, respectively. We define the microenvironment of cell i as the number N_i^0 of cells having phenotype $X_i = 0$ and by the number N_i^1 of cells having phenotype $X_i = 1$. By assuming a maximum cell capacity N of the microenvironment , we can define

$$N = N_i^1 + N_i^0 + N_i^\phi, \tag{1}$$

where N_i^ϕ are the free spaces/slots. The total number of cells N_T is defined as the sum of cells having phenotype $(X_i = 1)$ and cells having phenotype $(X_i = 0)$. So,

$$N_T = N_i^1 + N_i^0. \tag{2}$$

Cells will select their phenotype in a Bayesian fashion, i.e.

$$P\left(X_i \mid Y_i\right) = \frac{P\left(Y_i \mid X_i\right) P\left(X_i\right)}{P\left(Y_i\right)}, \tag{3}$$

where $P\left(Y_i \mid X_i\right)$ can be interpreted as the probability that the cell perceives all other cells in its surroundings, and $P\left(X_i\right)$ is the prior probability distribution of the cell's phenotypes. The former models the microenvironmental sensing the intrinsic programming of cellular phenotype. However, sensing other cells and evaluating $P\left(Y_i \mid X_i\right)$ entails an energy cost. It is reasonable to assume that the cell will try to optimize its prior $P\left(X_i\right)$ for the sake of energetic frugality.

Here we define Y_i as the extrinsic random variable of the i-th cell, where $Y_i = N_i^{X_i} \mid N_T, X_i = 0, 1$ represents the number of cells of phenotype X_i given the total local number of cells N_T. The conditional probability of having N_i^1 number of cells present in the microenvironment follows a binomial distribution ($\boldsymbol{B}$) (for the derivation details see S.I.):

$$P\left(N_i^1 \mid N_T\right) = P\left(Y_i = N_i^1 \mid N_T\right) = \boldsymbol{B}\left(N_T, p_1\right), \tag{4}$$

where p_1 is the probability of N_i^1 number of cells having phenotype $X_i = 1$ out of N_T cells:

$$p_1 = \frac{N_i^1}{N_T}. \tag{5}$$

Using statistical mechanical arguments [17], the problem of finding the probability of the cell's phenotype is equivalent to finding the prior that minimizes the entropy of a cell's phenotype and its surroundings. Let $S\left(X_i, Y_i\right)$ be the entropy of the cell-surroundings, $S\left(X_i\right)$ the internal entropy of the cell, and $S\left(Y_i \mid X_i\right)$ the entropy of the information sensed by the cell. Then, the entropies are connected by the relation $S\left(X_i, Y_i\right) = S\left(X_i\right) + S\left(Y_i \mid X_i\right)$. The optimization problem is finding $P\left(X_i\right)$ that minimizes $S\left(X_i, Y_i\right)$, while making sure that $P\left(X_i\right)$ is normalized; in other words we impose

$$\begin{aligned} \frac{\delta}{\delta P\left(X_i\right)} \Big\{ S\left(X_i\right) + \beta \Big[\sum P\left(X_i\right) S\left(Y_i \mid X_i\right) - \bar{S}\left(Y_i \mid X_i\right) \Big] \\ -\lambda \Big[\sum P\left(X_i\right) - 1 \Big] \Big\} = 0, \end{aligned} \tag{6}$$

where $\frac{\delta}{\delta P(X_i)}$ is the functional derivative, $\bar{S}\left(Y_i \mid X_i\right)$ is the expected ensemble statistics [17], and λ and β are Lagrange multipliers. Taking into account the relations among entropies, eqn. (6) yields

$$P\left(X_i\right) = \frac{e^{\beta S(Y_i|X_i)}}{Z} = \frac{1}{1 + e^{\beta \Delta S}}, \tag{7}$$

where Z is the normalization factor

$$Z = \sum_{\alpha=0,1} e^{\beta S(Y_i|X_i=\alpha)} = e^{\beta S(Y_i|X_i=0)} + e^{\beta S(Y_i|X_i=1)}. \tag{8}$$

Please note that the parameter β has a biological interpretation, since it quantifies the intensity with which a cell senses and complies to the microenvironment. To facilitate the evaluation of the entropy distribution of the microenvironment, we have used a Gaussian distribution, which approximates the binomial distribution (for $N_T > 5$), i.e.

$$P(N_i^1, p_1) \to \boldsymbol{N}(N_T p_1, N_T p_1 (1 - p_1)), \tag{9}$$

for the microenvironment [18]. Now we can evaluate the equilibrium distribution $P(X_i = 0)$ as

$$P\left(X_i = 0\right) = \frac{e^{\beta S\left(Y_i = [N_i^1 | N_T] | X_i = 0\right)}}{Z} = \frac{1}{1 + e^{\beta \Delta S}} = \\ = \frac{1}{1 + \left[\frac{N_i^0\left(N_i^1 - 1\right)}{N_i^1\left(N_i^0 - 1\right)}\right]^{\frac{\beta}{2}}} = \frac{1}{1 + \left[\frac{\rho_i^0\left(\rho_i^1 - \frac{1}{V}\right)}{\rho_i^1\left(\rho_i^0 - \frac{1}{V}\right)}\right]^{\frac{\beta}{2}}}, \tag{10}$$

with the cell densities $\rho_i^{0,1} := \frac{N_i^{0,1}}{V}$ (in the SI Fig. (12), we show the dependency of $P\left(X_i = 0\right)$ on resting density ρ_0 and β). Accordingly, the probability of a i-th cell's phenotype $P(X_i = 1)$ reads

$$P\left(X_i = 1\right) = \frac{e^{\beta S\left(Y_i = [N_i^1 | N_T] | X_i = 1\right)}}{Z} = 1 - \frac{1}{1 + e^{\beta \Delta S}} = \\ = 1 - \frac{1}{1 + \left[\frac{N_i^0\left(N_i^1 - 1\right)}{N_i^1\left(N_i^0 - 1\right)}\right]^{\frac{\beta}{2}}} = \frac{\left[\frac{\rho_i^0\left(\rho_i^1 - \frac{1}{V}\right)}{\rho_i^1\left(\rho_i^0 - \frac{1}{V}\right)}\right]^{\frac{\beta}{2}}}{1 + \left[\frac{\rho_i^0\left(\rho_i^1 - \frac{1}{V}\right)}{\rho_i^1\left(\rho_i^0 - \frac{1}{V}\right)}\right]^{\frac{\beta}{2}}}. \tag{11}$$

Please note that, the difference between the microenvironmental entropy of the i-th cell is

$$\Delta S = S(Y_i = [N_i^1 \mid N_T] | X_i = 1) - S(Y_i = [N_i^1 \mid N_T] | X_i = 0) = \frac{1}{2} \ln \left[\frac{N_i^0\left(N_i^1 - 1\right)}{N_i^1\left(N_i^0 - 1\right)}\right], \tag{12}$$

N_i^0 and N_i^1 are the number of cell of phenotype $(X_i = 0)$ and phenotype $(X_i = 1)$ in the microenvironment of i-th cell, respectively (See S.I. for details).

Fig. (1a), shows a sketch of our model. In short, we make the following assumptions
$(A1)$ The microenvironment is defined by the numbers of two phenotypes only.
$(A2)$ We have assumed a binomial distribution for the occurrence of the two phenotypes.
$(A3)$ Cells are making decisions at a fast time scale. This justifies to assume an equilibrium distribution for the different phenotypes.

In this section, we have defined a minimal LEUP-driven cell decision making model and derived the corresponding phenotypic steady states. Based on this, we subsequently develop a cell-based model to understand the resulting multicelluar spatiotemporal dynamics.

3 An individual-based model (IBM) for LEUP-driven cell migration

We define a discrete stochastic, spatio-temporal IBM that incorporates the phenotypic switch dynamics according to LEUP. To this end, we model the movement of single cells with Langevin equations. Langevin equations are well-suited to model cell-cell interactions and cell migration [19, 20].

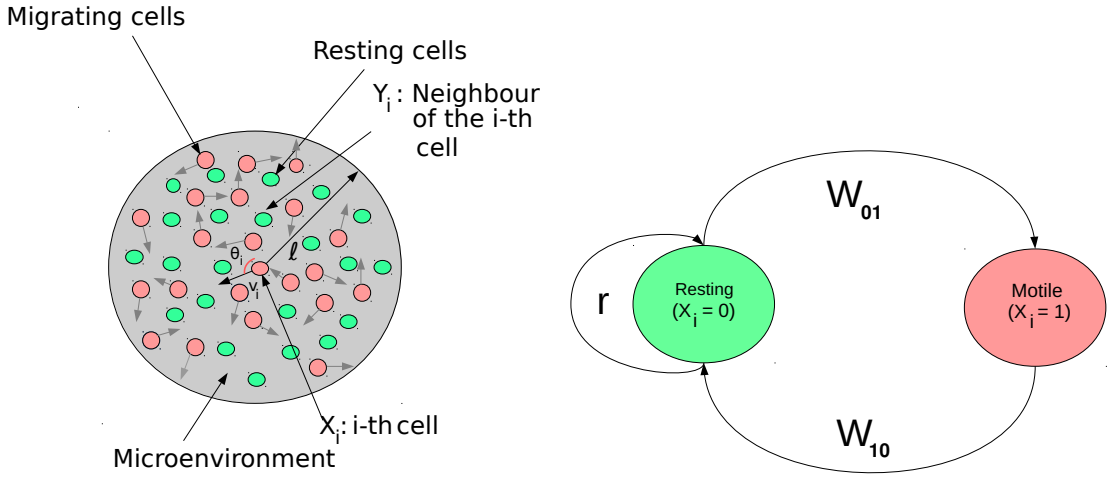


Figure 1. Schematic diagram of (**a**) the microenvironment of a motile cell, where ℓ is the sensing radius. The sensed number of cells is proportional to the volume of the sensed microenvironment, i.e. $N \propto \ell^d$, where d is the dimension. (**b**) Transitions in the "Go or Grow" model [9]. The switching between motile ($X_i = 1$) and resting ($X_i = 0$) phenotype is shown, where the transtion probabilities are defined as W_{10} and W_{01}. The proliferation rate is defined by r.

Our Langevin's equation is defined on a domain $\mathcal{L} \subset \mathbb{R}^2$ with periodic boundary conditions. We define an interaction radius $\ell \in \mathbb{R}$ around the i-th cell at position $\mathbf{x_i} \in \mathcal{L}$. The expected interaction volume is $V \propto \ell^2$. The time evolution of our model is defined by the following rules:

(R1) Cells change their phenotype by sensing their microenvironment within the interaction radius ℓ according to LEUP.

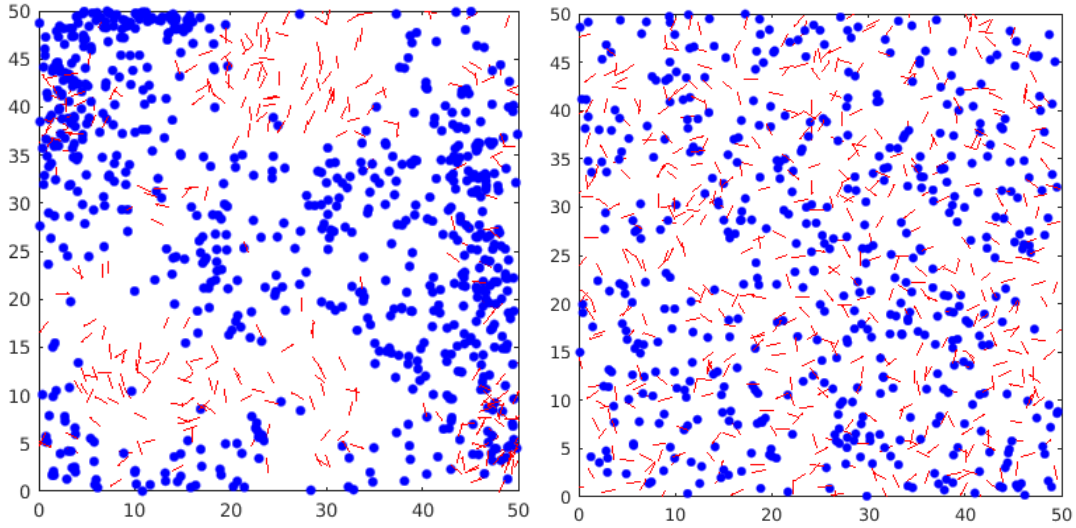
(R2) Moving cells change their orientations randomly (random walk).

(R3) Once cells become migratory they move with a constant speed $\bar{v}$.

The above are translated to the following Langevin's equations

$$\begin{aligned}
\frac{\mathrm{d}\mathbf{x_i}}{\mathrm{d}t} &= \bar{v}\mathbf{v_i}(\theta_i), \\
\frac{\mathrm{d}\theta_i}{\mathrm{d}t} &= \frac{1}{\bar{v}}\xi_i^{\theta}(t), \\
\frac{\mathrm{d}p_i}{\mathrm{d}t} &= -\frac{1}{\tau}(p_i - p_i^{eq}), \\
v_i &= \begin{cases} \bar{v}, & \text{if} \quad X_i = 1, \\ 0, & \text{if} \quad X_i = 0, \end{cases}
\end{aligned} \tag{13}$$

where $\mathbf{v_i} = (\cos\theta_i, \sin\theta_i)^T$ is the direction of movement of cell i. For the temporal evolution of the probability p_i of the motile state, we use the BGK (Bhatnagar–Gross–Krook) operator



(a) **(b)**

Figure 2. Clustering of resting cells in the IBM. (**a**) Snapshot of an example realization at $t = 100$, where the total number of cells is fixed at 1000 and $\beta = 40$. The spatial configuration shows the clustering of resting cells and motile cells moving in the free space. (**b**) Snapshot of the realization at $t = 100$, where the total number of cells is fixed at 1000 and $\beta = 5$. Here we observe a completely random spatial configuration. Resting cells are shown in blue and migrating cells are marked by red arrows indicating their direction of movement.

technique [21]. We assume that the probability p_i evolves weakly out of its equilibrium probability p_i^{eq}, which is the LEUP steady state probability $P(X_i = 1)$ (see eq. (11)). In turn, the parameter τ is the relaxation time towards the corresponding probability distribution p_i^{eq}. Here noise is assumed to have a zero-mean, white noise term, which has the statistical properties $\left\langle \xi_i^{\theta}(t) \right\rangle = 0$ and $\left\langle \xi_i^{\theta}(t_1)\, \xi_j^{\theta}(t_2) \right\rangle = 2D_{\theta}\delta(t_1 - t_2)\,\delta_{ij}$, where t_1 and t_2 are two time points, $D_{\theta} \in \mathbb{R}_+$ is the angular diffusion coefficient, $\delta(t)$ is the Dirac delta, and δ_{ij} is the Kronecker delta. Parameter $\bar{v}$ is the constant speed of every motile cell.

We simulate the Langevin model, on a two-dimensional domain, with a varying interaction radius ℓ, mean cell density and sensitivity β. We assume an initial state that is approximately homogeneous in space and that each of the initial 1000 cells is in the resting/migratory state with equal probability. In Figures (2a) and (2b), we showcase simulations that exhibit clustering and random spatial configurations, respectively. We consider a density regime which is compatible with typical situations during cancer invasion. In particular, the densities in Fig. 2 correspond to 1000 glioma cancer cells in a $2.5 \times 10^4 \mu m^2$ domain, which agrees with the typical glioma cell density $4.5 \times 10^{-3}\, \mu m^{-2}$ [22].

To quantify cell clustering, we calculate the radial distribution function for different sensitivities and different interaction radii as shown in Figures (3a) and (3b). The key observations from the simulation study are:

1. There is a critical threshold for the parameter $\beta > 0$, where clustering of resting cells

occurs (Fig. 3a). This is rather expected since a sufficient sensing of the microenvironment is required.

2. For an intermediate interaction radius ℓ, we observe cell clusters as shown in Fig. 3b. Very high ℓ corresponds to a large number of sensed cells. This leads to also equal steady state probabilities, as the entropy difference in the microenvironment, associated with the single cell states becomes negligible. On the other hand, the lower bound of ℓ is expected, since enough sampling size of cellular microenvironment is required to induce aggregation patterns.

3. In Fig. 3d, we show the phase diagram of the system, when parameter β and the average density are varied. In particular, we observe that there is a parametric regime where clustering behavior is emerging. Interestingly, high densities do not always imply patterning and require higher values of β to support cluster emergence (see Fig. 3c). Finally, as expected, low densities also reduce the area of the patterning regime for low β.

Interestingly, existing radial distribution function data from patient-derived glioma tumors show also a "flat" behavior, i.e. they do not exhibit any clustering [22]. In particular, the configuration in Figure (2b) corresponds to such a "flat" radial distribution function, for $\beta = 5$, at densities relevant to real glioma tumors. However, there are some extra cell-cell interactions mechanisms acting in glioma cells, i.e. a repulsive force conferring a large volume exclusion effect. Studying effects of these additional mechanisms is the topic of current research.

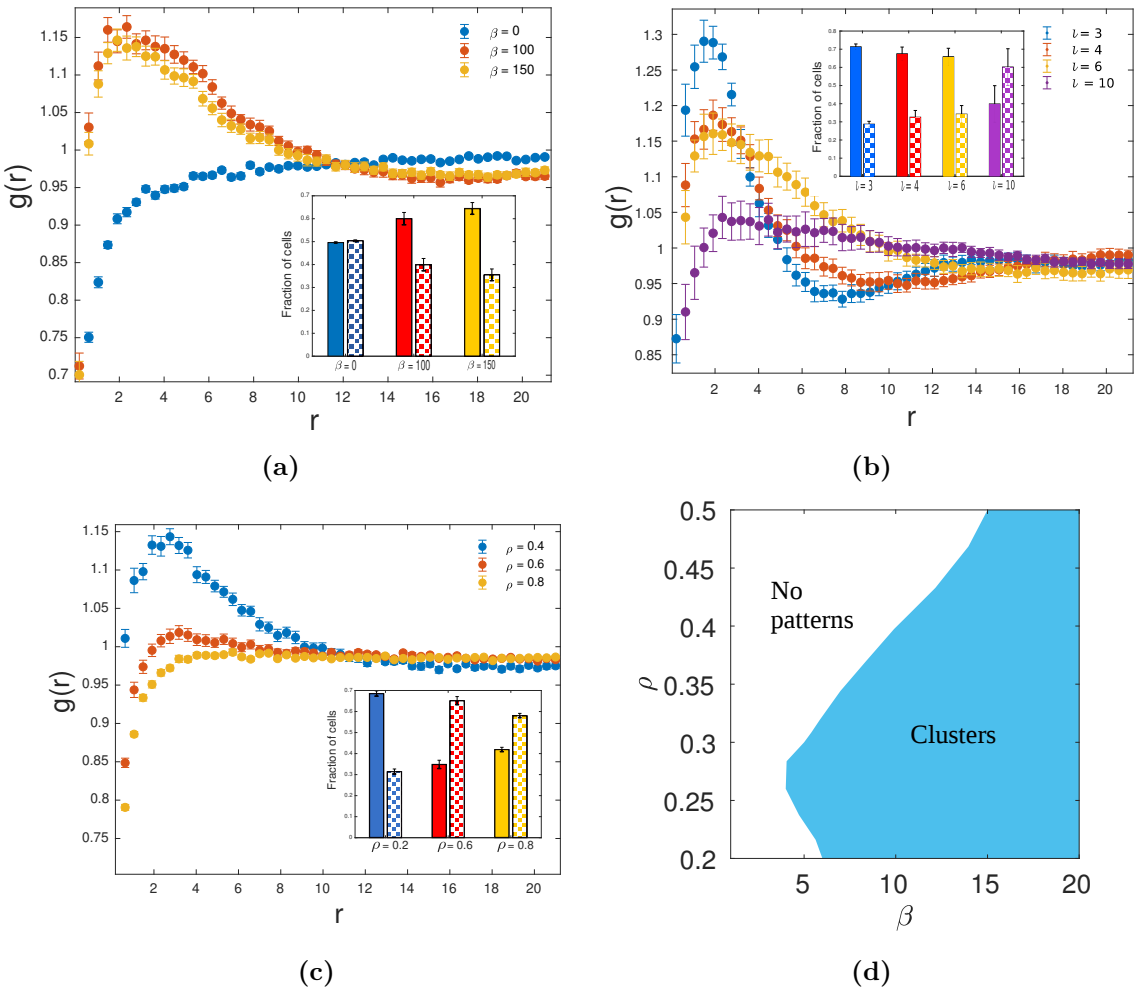
4 Mean-field approximation

In this section, we derive a continuous approximation of the aforementioned discrete model using a mean-field approach. The goal is to shed light of the pattern formation mechanisms, i.e. cell clustering, as observed in the microscopic simulations. The main idea of the mean-field approximation is to replace the description of many-particle interactions by a single particle description based on an average or effective interaction. Thereby, any multi-particle problem can be replaced by an effective description, that can be stated in the form of an ordinary (ODE) or partial differential equation (PDE). In order to proceed, we will first treat the switch dynamics and the migration process separately.

4.1 Mean-field description of the phenotypic switch dynamics

Let's consider cell i at position $\mathbf{x_i}$ with states $X_i = 0, 1$. The cell changes its state in dependence on the local microenvironment according to the LEUP with rates W_{01}, W_{10} and the probabilities to be in either states are denoted by $p_{0,1}$. We assume that the system is always close to the steady state, so the master equation reads

$$\frac{\mathrm{d}}{\mathrm{d}t} p_0 = W_{10} p_1 - W_{01} p_0 \approx 0, \tag{14}$$



(a) **(b)** **(c)** **(d)**

Figure 3. Quantification of cell clustering. (**a**) Radial distribution function with standard deviation for different values of β, for a fixed interaction radius $\ell = 6$. The number of cells was fixed at 1000, corresponding to a mean cell density of 0.4. (**b**) Radial distribution function with standard deviation for varying interaction radius ℓ at fixed sensitivity $\beta = 100$. There is an optimal sensing radius for a given mean cell density, so that interaction radii that are too large or too low do not lead to aggregation, indicated by an almost flat radial distribution, see purple and blue lines.. (**c**) Radial distribution function with standard deviation for different mean densities, at $\beta = 20$ and $\ell = 4$. When the mean density becomes too large, no clustering can be observed. The inlets in (**a - c**) show the corresponding fractions of motile (checkerboard bars) and resting (filled bars) cells. (**d**) Phase space diagram for cluster pattern formation. Here we define clustering when $\max_r (g(r) > 1.09)$ (see text for explanation). The interaction radius is fixed at $\ell = 4$. All data points in (**a - d**) correspond to the mean of 20 independent simulations at $t = 100$.

and $p_{0,1}$ are given by eq. (7). By rearranging the terms, we obtain for the switching rates

$$\frac{W_{10}}{W_{01}} = \frac{p_0}{p_1}. \tag{15}$$

For simplicity, we set $W_{01} = p_1$ where the transition probability towards motile phenotypes equals to the moving steady state probability p_1. This coincides with the "detailed balance condition". The steady state probabilities $p_{0,1}$ depend on the number of cells $N_i^{0,1}$ in the respective phenotypes in the microenvironment. Cell i at position $\mathbf{x_i}$ senses

$$N_i^X(\mathbf{x_1}, \ldots, \mathbf{x_i}, \ldots, \mathbf{x_N}, t) := \sum_{j=1}^{N} \xi\big(\mathbf{x_j}(t) - \mathbf{x_i}(t), t\big)\delta_{X_j,X}\,, \quad X = 0, 1. \tag{16}$$

Here, we sum over all cells j, and $\xi\big(\mathbf{x_j}(t) - \mathbf{x_i}(t), t\big)$ is a Boolean stochastic variable that serves as the sensing function of a cell at $\mathbf{x_i}$. It depends on the distance between cells and time, and $\xi = 1$ indicates that the cell j at $\mathbf{x_j}$ is sensed by cell i at position $\mathbf{x_i}$. To match the IBM, we assume that

$$\xi(\mathbf{x_j}(t) - \mathbf{x_i}(t), t) = \Theta(\ell - |\mathbf{x}_j(t) - \mathbf{x_i}(t)|), \quad \ell > 0, \tag{17}$$

with the Heaviside step function $\Theta(x)$, so that all cells in a ball of radius ℓ around $\mathbf{x_i}$ are sensed.

To proceed, we apply a mean-field approximation to calculate the expected switching rate $\langle W_{10}(N_i^0, N_i^1)\rangle \approx W_{01}(\langle N_i^0\rangle, \langle N_i^1\rangle)$. Note that we dropped the dependence on space and time for better readability. Let $\phi_i(X, \mathbf{x}, t)$ denote the probability of finding cell i in a small volume $\mathrm{d}V$ around $\mathbf{x}$ at time t with the phenotype $X = 0, 1$. Formally we have

$$\phi_i(X, \mathbf{x}, t) = \langle \delta(\mathbf{x} - \mathbf{x_i}(t))\delta_{X,X_i}\rangle\,, \tag{18}$$

where the average is the ensemble average, and the total densities of resting/migrating cells are

$$\Phi(\mathbf{x}, t) := \sum_i \phi_i(0, \mathbf{x}, t), \tag{19}$$

$$\Psi(\mathbf{x}, t) := \sum_i \phi_i(1, \mathbf{x}, t). \tag{20}$$

For the expected density of sensed cells in the microenvironment around $\mathbf{x}$ we obtain

$$\rho_0(\mathbf{x}, t) := \frac{\langle N_i^0(\mathbf{x}, t)\rangle}{V} = \frac{1}{V}\int_{\mathcal{B}_\ell(\mathbf{x})} \Phi(\mathbf{y}, t)\,\mathrm{d}V, \tag{21}$$

$$\rho_1(\mathbf{x}, t) := \frac{\langle N_i^1(\mathbf{x}, t)\rangle}{V} = \frac{1}{V}\int_{\mathcal{B}_\ell(\mathbf{x})} \Psi(\mathbf{y}, t)\,\mathrm{d}V, \tag{22}$$

where $V \propto \ell^d$, and where we have already used eq. (17). Consequently, we obtain the approximate switching rate

$$\bar{W}_{10} = p_0(\rho_0, \rho_1), \tag{23}$$

and the switch dynamics is (dropping dependencies on space and time for simplicity)

$$\frac{\partial}{\partial t}\phi_i(0) = p_0\phi_i(1) - p_1\phi_i(0). \tag{24}$$

Summing over all cells, we obtain for the total density of resting cells

$$\frac{\partial}{\partial t}\Phi = p_0\Psi - p_1\Phi. \tag{25}$$

Note that this is a non-local, non-linear set of PDEs, which are difficult to treat analytically. However, we can further simplify our analysis by making another approximation, assuming that the sensing radius ℓ is much smaller than the total domain size. Then, we can replace the non-local sensing function $\xi(\mathbf{x_j}(t) - \mathbf{x_i}(t), t)$ by a local delta distribution

$$\xi(\mathbf{x_j}(t) - \mathbf{x_i}(t), t) \to \delta(\mathbf{x_j}(t) - \mathbf{x_i}(t)). \tag{26}$$

In this case, the expected number of sensed cells in the microenvironment around $\mathbf{x}$ simply becomes the local density of the respective phenotype

$$\rho_0(\mathbf{x}, t) := \left\langle N_i^0(\mathbf{x}, t) \right\rangle = \frac{\Phi(\mathbf{x}, t)}{\ell^d}, \tag{27}$$

$$\rho_1(\mathbf{x}, t) := \left\langle N_i^1(\mathbf{x}, t) \right\rangle = \frac{\Psi(\mathbf{x}, t)}{\ell^d}. \tag{28}$$

Finally, the rate $\bar{W}_{10}$ reduces to

$$\bar{W}_{10} = p_0(\rho_0, \rho_1) = \frac{1}{1 + e^{\beta \Delta S}} = \frac{1}{1 + \left[\frac{\rho_0\left(\rho_1 - \frac{1}{V}\right)}{\rho_1\left(\rho_0 - \frac{1}{V}\right)}\right]^{\frac{\beta}{2}}}, \tag{29}$$

and the reversed transition probability $\bar{W}_{01}$ is

$$\bar{W}_{01} = p_1(\rho_0, \rho_1) = 1 - \frac{1}{1 + e^{\beta \Delta S}} = 1 - \frac{1}{1 + \left[\frac{\rho_0\left(\rho_1 - \frac{1}{V}\right)}{\rho_1\left(\rho_0 - \frac{1}{V}\right)}\right]^{\frac{\beta}{2}}}, \tag{30}$$

where the volume is defined as $V = \ell^d$.

4.2 Mean-field description of the cell migration process

In this section, we derive the macroscopic equation in two spatial dimensions for the motile cell population. As before we can write the corresponding stochastic Langevin's equations as

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x_i} &= \bar{v}\mathbf{v_i}(\theta_i), \\ \frac{\mathrm{d}}{\mathrm{d}t}\theta_i &= \frac{1}{\bar{v}}\xi_i^{\theta}(t). \end{aligned} \tag{31}$$

This process can be considered as a special kind of active Brownian motion. In this kind of Langevin's equation, the stochastic force creates variations of orientation. According to [23] we can derive the corresponding Fokker-Planck equation for migrating cells using adiabatic elimination and averaging it over different noise realizations obtaining the following diffusion equation

$$\frac{\partial \rho_1(x, t)}{\partial t} = \frac{\bar{v}^4}{2D_\theta}\nabla^2\rho_1(x, t). \tag{32}$$

Here we define D_1 as diffusion coefficient which is $\frac{\bar{v}^4}{2D_\theta}$.

4.3 Coupling migration and switching dynamics

Combining our results in the previous sections, we can easily formulate a system of PDEs

$$\partial_t \rho_0(x,t) = \nu E(\rho_0,\rho_1) \tag{33}$$
$$\partial_t \rho_1(x,t) = D_1 \nabla^2 \rho_1 - \nu E(\rho_0,\rho_1) \tag{34}$$
$$E(\rho_0,\rho_1) = \bar{W}_{10}(\rho_0,\rho_1)\rho_1 - \bar{W}_{01}(\rho_0,\rho_1)\rho_0 \tag{35}$$

where $E(\rho_0,\rho_1)$ is the phenotypic exchange term and ν is the corresponding timescale ratio of the switching and diffusion processes, i.e. $\nu = \frac{\tau_S}{\tau_D}$. To ensure the numerical consistency of the above system, we assume that resting cells diffuse in a very slow manner i.e. $D_0 \ll D_1$, which results in the following reaction-diffusion system of equations

$$\partial_t \rho_0(\mathbf{x},t) = D_0 \nabla^2 \rho_0 + \nu E(\rho_0,\rho_1), \tag{36}$$
$$\partial_t \rho_1(\mathbf{x},t) = D_1 \nabla^2 \rho_1 - \nu E(\rho_0,\rho_1). \tag{37}$$

5 Spatio-temporal dynamics of the LEUP-driven migration/proliferation plasticity

In this section, we study the mean-field approximation of the aforementioned stochastic plasticity dynamics for different regimes of the sensitivity β and the interaction radius ℓ. In particular, we initially study the system dynamics, when cells have a very large interaction radius ($\ell \gg 1$) and then a finite one. Finally, we study also the special case $\beta = 0$, i.e. cells decide independently of their microenvironment.

5.1 Large interaction radius case: existence of a critical sensitivity

Here, we focus on the very large interaction radius system ($\ell \gg 1 \implies V \gg 1$). Although, this parameter regime is not biologically relevant, it is very instructive since it allows us to derive analytical estimates for our systems dynamics. Here, we generalize the macroscopic system by adding proliferation dynamics (logistic growth). Our phenotypic switch dynamics are recapitulated by setting the proliferation rate to zero, i.e. $r = 0$. The full system reads

$$\begin{aligned} \frac{\partial \rho_0}{\partial t} &= D_0 \nabla^2 \rho_0 + \nu E_0\left(\rho_0,\rho_1\right) + r\rho_0\left(1-\rho_1-\rho_0\right), \\ \frac{\partial \rho_1}{\partial t} &= D_1 \nabla^2 \rho_1 - \nu E_0\left(\rho_0,\rho_1\right). \end{aligned} \tag{38}$$

In turn, we conduct a non-dimensionalization of eqns. (38) to identify the variables. Moreover, this helps us to gain a knowledge about the relationships between the different model parameters.

By assuming that the system size is fixed at L the non-dimensional quantities read

$$\begin{aligned} x^* &= \frac{x}{L} \implies \frac{\partial^2}{\partial x^{*2}} = L^2 \frac{\partial^2}{\partial x^2}, \\ t^* &= \frac{D_0 t}{L^2} \implies \frac{\partial}{\partial t^*} = \frac{L^2}{D_0} \frac{\partial}{\partial t}, \\ \gamma &= \frac{L^2 \nu}{D_0}, \\ D &= \frac{D_1}{D_0}, \\ r^{'} &= \frac{r}{\nu}. \end{aligned} \tag{39}$$

In the limit $V \gg 1$, the eqn. (38) can also be written as

$$\begin{aligned} \frac{\partial \rho_0}{\partial t^*} &= \nabla^{*2} \rho_0 + \gamma \left((\rho_1 - \rho_0) \left(\frac{1}{2} - \tilde{\beta} \frac{(\rho_1 + \rho_0)}{\rho_1 \rho_0} \right) + r^{'} \rho_0 \left(1 - \rho_1 - \rho_0 \right) \right), \\ \frac{\partial \rho_1}{\partial t^*} &= D \nabla^{*2} \rho_1 - \gamma \left((\rho_1 - \rho_0) \left(\frac{1}{2} - \tilde{\beta} \frac{(\rho_1 + \rho_0)}{\rho_1 \rho_0} \right) \right). \end{aligned} \tag{40}$$

To understand the behaviour of the system at long times, we conduct a fixed point analysis. Initially we assume a well-stirred system, i.e. no spatial interactions. Then, eqs. (40) can be written as coupled non-linear ODEs which have three fixed points

$$\begin{aligned} \left(\rho_0^{A*}, \rho_1^{A*} \right) &= \left(\frac{1}{2}, \frac{1}{2} \right), \\ \left(\rho_0^{B*}, \rho_1^{B*} \right) &= \left(\frac{1}{2} \left(1 + \sqrt{1 - 8\tilde{\beta}} \right), \frac{1}{2} \left(1 - \sqrt{1 - 8\tilde{\beta}} \right) \right), \\ \left(\rho_0^{C*}, \rho_1^{C*} \right) &= \left(\frac{1}{2} \left(1 - \sqrt{1 - 8\tilde{\beta}} \right), \frac{1}{2} \left(1 + \sqrt{1 - 8\tilde{\beta}} \right) \right). \end{aligned} \tag{41}$$

The above imply a pitchfork bifurcation for the $\tilde{\beta} = \frac{\beta}{8V}$ parameter, i.e. there exists a critical value $\tilde{\beta}_c$ that introduces the bistable state. From eqs.(41), we can easily deduce that the critical sensitivity value

$$1 - 8\tilde{\beta}_c = 0 \iff \beta_c = V. \tag{42}$$

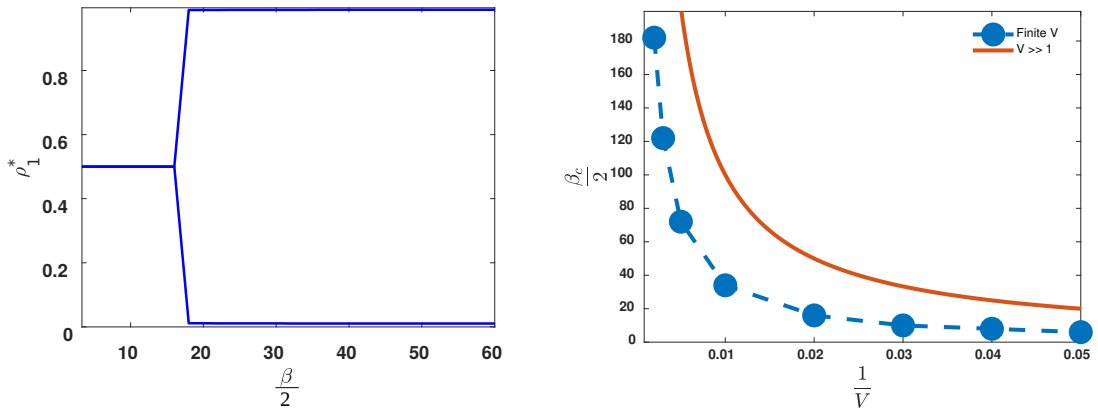
This is an acceptable approximation even for the finite interaction radius system (see next section). In the following, we analyze if the system is able to produce spatial patterns. Applying linear stability analysis for the spatially resolved system eqs. (40), we can deduce that no pattern formation is possible (for details check next section). Any perturbations to the homogeneous state lead always to a spatially homogeneous steady state. When $r = 0$ this result is consistent with our findings in discrete IBM simulations where very large interaction radii do not confer any clustering, as shown in Fig. 3b.

5.2 The finite interaction radius case: emergence of a bistable switch between "fluid" and "solid" tissue phases and pattern formation

Now we turn to the full system for intermediate interaction radius also (assuming proliferation). This implies a finite interaction volume V and for analytical feasibility we are interested in the Gaussian approximation of the switching probabilities and their corresponding mean-field terms. The full system of PDEs assuming also proliferation reads

$$\begin{aligned}
\frac{\partial \rho_0}{\partial t^*} &= \nabla^{*2}\rho_0 + \gamma \left(\frac{\rho_1 - \rho_0 \left(\frac{\rho_0\left(\rho_1 - \frac{1}{V}\right)}{\rho_1\left(\rho_0 - \frac{1}{V}\right)} \right)^{\frac{\beta}{2}}}{1 + \left(\frac{\rho_0\left(\rho_1 - \frac{1}{V}\right)}{\rho_1\left(\rho_0 - \frac{1}{V}\right)} \right)^{\frac{\beta}{2}}} + r^{'} \rho_0 \left(1 - \rho_1 - \rho_0\right) \right), \\
\frac{\partial \rho_1}{\partial t^*} &= D\nabla^{*2}\rho_1 - \gamma \frac{\rho_1 - \rho_0 \left(\frac{\rho_0\left(\rho_1 - \frac{1}{V}\right)}{\rho_1\left(\rho_0 - \frac{1}{V}\right)} \right)^{\frac{\beta}{2}}}{1 + \left(\frac{\rho_0\left(\rho_1 - \frac{1}{V}\right)}{\rho_1\left(\rho_0 - \frac{1}{V}\right)} \right)^{\frac{\beta}{2}}}.
\end{aligned} \tag{43}$$

Since closed expression of steady states are not analytically feasible, we obtain the bifurcation diagram numerically, see Fig (4a). We observe the existence of a supercritical pitchfork



(a) **(b)**

Figure 4. (**a**) Bifurcation diagram for moving cells in the finite interaction radius limit with respect to β, where proliferation rate r is fixed to 0 and V^{-1} at 0.02. The bifurcation diagram is symmetric for resting cells i.e. ρ_0^*. (**b**) The critical point value of the bifurcation diagram (where the solution splits in two branches) is plotted against the inverse volume $\frac{1}{V}$.

bifurcation and the existence of a critical β_c. For $\beta \geq \beta_c$, the systems depart from balanced state ($\rho_0 = \rho_1 = \frac{1}{2}$) to the coexistence of a "fluid" (most cells migrate) and a "solid" (most cells are resting) phase. The switch is controlled by the perturbation on the ratio of migratory and resting cells. Interestingly, we can compare the analytic estimate of $\beta_c = V$, from the $V \gg 1$ case, with the one calculated for the finite V system. Fig. (4b) shows that the infinite system

approximation provides an upper bound for β_c, which is not too far from the real value of finite systems. The existence of a critical β_c with respect to the two phases is evident in the IBM simulations as well. In particular if we quantify the ratio of stationary over motile cells, we observe a similar behavior of the critical sensitivity for increasing V, since it decreases (see Fig.13 in SI). However, we cannot conduct a strict quantitative comparison since our bifurcation analysis does not involve any diffusion as opposed to the IBM simulation.

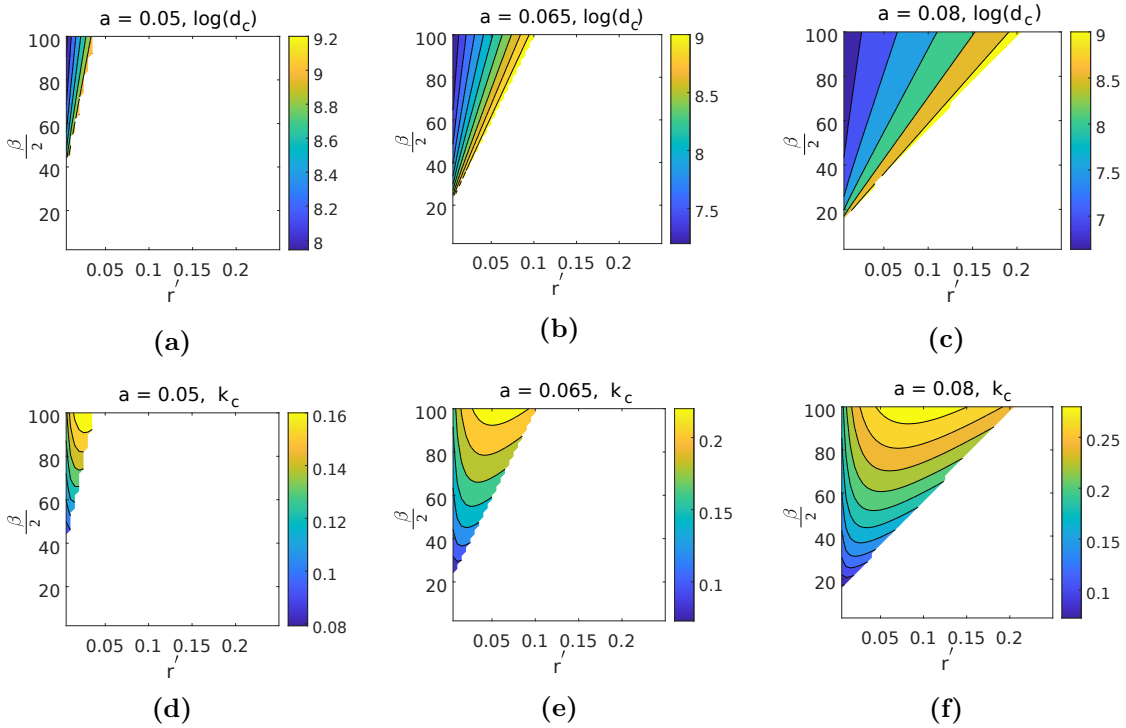


Figure 5. Spatio-temporal pattern formation in dependence of sensitivity $\beta/2$, proliferation rate r' and inverse sensing volume $a = \frac{1}{V}$. **(a - c)** The critical diffusion coefficient $d_c = D_1/D_0$ decreases for increasing sensitivity or increasing proliferation rate. **(d - f)** The critical wave number k_c of the observed pattern increases for an increasing sensitivity.

In turn, we apply linear stability analysis to identify parameter regimes that promote diffusion driven pattern formation or Turing instability. To analyze the Turing instability [24] we have to find the system's steady state, (i.e. when diffusion is not present in the systems of eqn.(38)). It has been shown that $d = \frac{D_1}{D_0} \gg 1$ is a necessary condition for the emergence of spatially heterogeneous solution i.e. patterns. Now, we can write the system of PDEs (i.e eqn.(38)) in a

generalized matrix form

$$\begin{aligned}
&\frac{\partial \boldsymbol{\rho}}{\partial t} = \boldsymbol{D}\nabla^2\boldsymbol{\rho} + \gamma \boldsymbol{R}\boldsymbol{\rho}, \quad \boldsymbol{D} = \begin{pmatrix} 1 & 0 \\ 0 & d \end{pmatrix}, \\
&\boldsymbol{R} = \begin{pmatrix} \frac{\partial E_0(\rho_0,\rho_1)}{\partial \rho_0} + r^{'}(1-\rho_1-\rho_0) - r^{'}\rho_0 & \frac{\partial E_0(\rho_0,\rho_1)}{\partial \rho_1} - r^{'}\rho_0 \\ -\frac{\partial E_0(\rho_0,\rho_1)}{\partial \rho_0} & -\frac{\partial E_0(\rho_0,\rho_1)}{\partial \rho_1} \end{pmatrix}_{(\rho_0^*,\rho_1^*)}, \\
&\quad = \begin{pmatrix} f_u & f_v \\ g_u & g_v \end{pmatrix}_{(\rho_0^*,\rho_1^*)},
\end{aligned} \tag{44}$$

where $\boldsymbol{\rho}$ is defined as $(\rho_0 - \rho_0^*, \rho_1 - \rho_1^*)^T$ and $\boldsymbol{R}$ is the Jacobian at $\rho^* = (\rho_0^*, \rho_1^*)$. Using the Turing conditions of instability [24], we found patterns when N is finite in zero-flux boundary conditions. We have checked all the Turing instability conditions:

$$\begin{aligned}
&f_u + g_v < 0, \qquad f_u g_v - f_v g_u > 0, \\
&Df_u + g_v > 0, \qquad (Df_u + g_v)^2 - 4D(f_u g_v - f_v g_u) > 0.
\end{aligned} \tag{45}$$

Interestingly, only when the density of resting cells is larger than that of the moving cells, patterns are formed under Turing instability conditions. In order to investigate the system's potential to exhibit pattern formation, we check the range of validity of the Turing instability criteria (45). Diffusion-driven instability conditions are satisfied for a large portion of the parameter space. In turn for these parameters, we calculate the critical diffusion coefficient d_c. For values $d > d_c$ we are able to observe patterns. This condition is generally fulfilled as the resting cells are only moving passively due to external noise and cell-cell interactions like cell-cell adhesion, while the migratory cells also move actively. The existence of d_c is associated with the existence of a critical wavenumber k_c [24].

$$\begin{aligned}
d_c &= \frac{-2(2f_v g_u - f_u g_v) \pm \sqrt{(2(2f_v g_u - f_u g_v))^2 - 4f_u^2 g_v^2}}{2f_u^2}, \\
k_c &= \gamma \left[\frac{Det(A)}{d_c}\right]^{\frac{1}{2}}.
\end{aligned} \tag{46}$$

In Fig. (5) we identify the parameter regime that allows us to observe patterns and then corresponding d_c and k_c.

By fixing the initial conditions to a cosine wave, we observe in one dimension the existence of regular spikes as shown in Fig (6). Please note that the exact pattern is sensitive to the initial conditions. In turn, we simulate our system in 2D and for the same parameters and we observe patterning in the form of dots. Finally, we investigate if the type of patterns changes for variations in parameters β and $r^{'}$ (see Fig. (7)). If $d > d_c$ then we need a large domain or size of the system to observe the patterns. In the 2D case, we have observed the discoidal patterns of resting phenotype which resemble the 1D case. So, the radius of the circles of the patterns are increased if we fix the domain size. Moreover, we observe that in both dimensions (i.e. 1D and 2D) the critical spatial frequency increases with decreasing $r^{'}$. In comparison with the discrete IBM simulations, we have to state that the simulation clusters can be identified by the discoidal mean-field patterns. Under this statement, we find that:

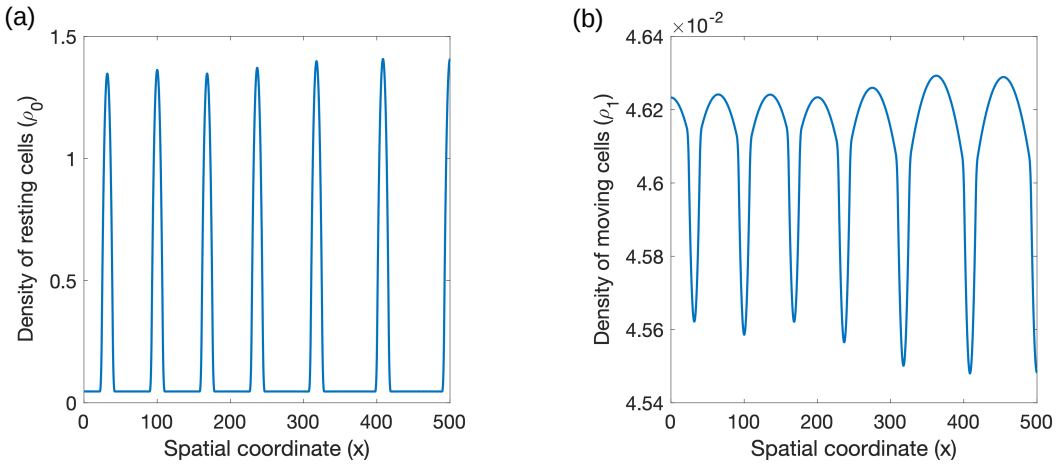


Figure 6. Pattern formation in the mean-field model. Densities of proliferating (**a**) and migrating (**b**) cells at $t = 50000$. Starting from an almost homogeneous initial state with a sinusoidal perturbation, the resting cells form a periodic pattern of high-density spikes, while the density of moving cells has minima at the position of the high-density spikes of the resting cells.

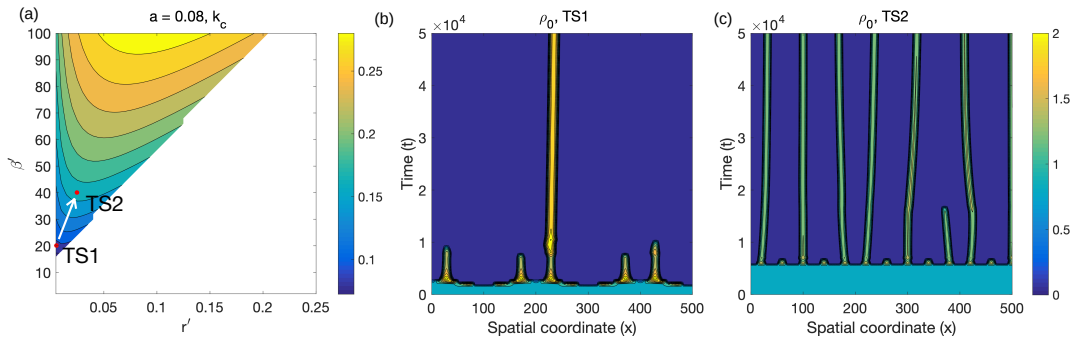


Figure 7. Saptio-temporal pattern formation of the mean-field model. (**a**) Turing space and critical wave number. **(b - c)** Density of resting cells over time. (**b**) A single high-density spike forms in agreement with the prediction of a small critical wave length for the given parameters (red dot TS1 in (**a**)). (**c**) For a higher sensitivity and proliferation rate a periodic pattern with smaller wave length emerges, corresponding to the prediction of a larger critical wave number (red dot TS2 in (**a**)). Here $\beta' = \frac{\beta}{2}$ and $a = \frac{1}{V}$.

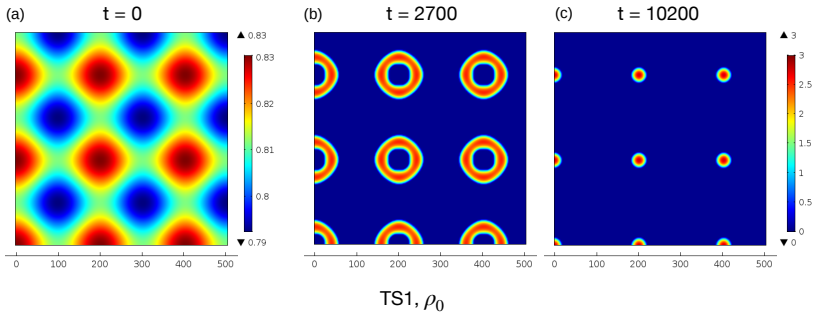


Figure 8. Two-dimensional simulation of the mean-field approximation with finite interaction radius. (a) Inital state, (b) $t = 2700$, (c) $t = 10200$ corresponding to a stable steady state. Parameters refer to point (TS1) in the Turing space, as displayed in Fig. (7).

1. There exists a critical β that allows for the emergence of patterns as in the IBM simulations.

2. For very low and very high interaction radius ℓ, we observe no patterns, which is consistent with our IBM results (see Fig.3b).

5.3 Microenvironment independent phenotypic switching leads to uncontrolled growth dynamics

Finally we investigate the case $\beta = 0$ where cells do not sense their microenvironment and so we end up with the following system,

$$\begin{aligned} \frac{\partial \rho_0}{\partial t^*} &= \nabla^{*2}\rho_0 + \gamma\left(\frac{(\rho_1 - \rho_0)}{2} + r^{'}\rho_0\left(1 - \rho_1 - \rho_0\right)\right), \\ \frac{\partial \rho_1}{\partial t^*} &= D\nabla^{*2}\rho_1 - \gamma\frac{(\rho_1 - \rho_0)}{2}. \end{aligned} \tag{47}$$

We can do a fixed point analysis similar to equation (40). Assuming no spatial dynamics, we can find two fixed points i.e.

$$\begin{aligned} \left(\rho_0^{P*}, \rho_1^{P*}\right) &= \left(\frac{1}{2}, \frac{1}{2}\right), \\ \left(\rho_0^{Q*}, \rho_1^{Q*}\right) &= (0, 0). \end{aligned} \tag{48}$$

For further details see Fig. 9 in S.I. We can clearly show that the fixed point (0,0) is unstable and $\left(\frac{1}{2}, \frac{1}{2}\right)$ is a saddle point. We can also write the coupled PDE equations in a single PDE as well, if we consider $\rho_0 = \rho_1 = \tilde{\rho}$. Then we obtain

$$\frac{\partial \tilde{\rho}}{\partial t^*} = \frac{(1 + D)}{2}\nabla^{*2}\tilde{\rho} + \frac{r}{2}\tilde{\rho}\left(1 - 2\tilde{\rho}\right). \tag{49}$$

Now we can clearly see that the eqn.(49) is similar to the Fisher-Kolmogorov equation [24]. It is known that the Fisher-Kolmogorov equation does not exhibit any pattern formation instabilities.

The latter observation is consistent with our discrete IBM simulations, where no clustering is observed.

6 Discussion

Recently, the Least microEnvironmental Uncertainty Principle (LEUP) has been proposed as an organization principle for cell decision-making in multicellular systems. In this paper, we apply this principle to shed light on the effects of phenotypic plasticity for tissue dynamics with a mathematical model. We focus on two types of plasticity: a go or rest, and a go or grow phenotypic dichotomy, which play key roles in important processes in biological development and pathological situations as cancer invasion. We assume that in any given spatial mesoscopic sample the presence of cell phenotypes follows a binomial distribution. Using this assumption, we are able to calculate the microenvironmental entropy and the LEUP-driven probability distribution of each phenotype. On the basis of this distribution, we defined an appropriate microscopic stochastic model for the spatio-temporal dynamics of both phenotypes, and in turn derived the corresponding mean-field description resulting in a system of coupled reaction-diffusion equations. The main results of our study are: (i) in the case of go or rest plasticity, there exists a supercritical pitchfork bifurcation that defines a switch between a "fluid" and a "solid" phase, and (ii) in the case of go or grow plasticity, for the "solid" phase, we can derive conditions for the emergence of Turing patterns. Interestingly, the existence of Turing patterns requires a critical LEUP sensitivity to the microenvironment and a minimum interaction range. Note that in this work we refer to a clustered, immobile configuration as "solid" and the homogeneous mobile configuration as "fluid". In Physics, solid states are characterized by deep "valleys" in their radial distribution function, separated by characteristic distances controlled by the underlying repulsive and adhesive forces. Here, the radial distribution function of our "solid" state does not like physical analogs because we do not include cell-cell adhesion and volume exclusion in our cell dynamics. We expect that adding such interaction forces to our IBM would result in classical solid and fluid radial distribution functions, which is subject to future work.

Our model assumes binary transitions between two discrete phenotypes, which is an extreme simplification of biological reality. Certainly, one could include a continuous spectrum of motile/proliferative phenotypes that is expected to imply even richer spatio-temporal dynamics. Indeed, assuming a continuous state (velocity) space would potentially lead to further interesting bifurcations, such as the metastable EMT state as found in [25, 26], or complex spatiotemporal patterns as indicated in [27]. But the very abstracted case chosen here is still instructive to indicate how entropy-driven phenotypic decisions can lead to particular types of spatio-temporal pattern formation.

Interestingly, our LEUP-driven IBM is an extension of a Vicsek-type model [28], formulated in the context of self-propelled particles [29]. Our model exhibits a novel collective behavior when compared to the past published results from Vicsek-type of models. In particular, to our knowledge it is the first time to produce with such models Turing patterns, i.e. dynamics clusters of non-motile cells of specific characteristic wavelength. Typically, in Viscek-type

models we may observe moving clusters of swirling cells (e.g. the milling Viscek model) but never static ones.

At this point, let us focus on the biological assumptions and implications of our study. The molecular regulatory mechanisms involved in EMT or GoG remain largely unknown, where the latter can be viewed as an EMT with proliferation constrained to the epithelial/resting phase. Here, we assume that the phenotypic regulation of both mechanisms is based on the minimization of microenvironmental entropy in physiological tissues, which allows us to predict the multicellular spatiotemporal dynamics. This assumption is supported by the fact that healthy physiological processes, biological development or processes like wound healing, where EMT/MET or GoG are present, typically lead to an ordered (low entropy) tissue from a disordered initial condition. On the other hand, deregulation of EMT and GoG have been already identified as pivotal elements in invading cancers [9, 30], where genetic and phenotypic heterogeneity, characterized by high entropy, is a key characteristic. Assuming a LEUP-driven migration/proliferation phenotypic regulation allows us to understand how cells control multicellular dynamics in terms of growth and patterning.

Implications in multicellular growth control: The central finding of our study is associated with the bifurcation diagram of the LEUP sensitivity parameter β in Fig. 4a. As stated before, the parameter β quantifies how prone cells are in sensing and responding to their microenvironmental stimuli. When $\beta = 0$ cells migrate and proliferate independent of their microenvironmental cues, which corresponds to one of the cancer hallmarks [31]. In this case, the systems grows uncontrollably, resembling a cancerous tissue [31]. The resulting Fisher-Kolmogorov macroscopic behavior has been prototypically used to model invading tumours [32]. Moreover, for low β our theory predicts no spatial order, in terms of clustering implying "flat" radial distribution functions (see Figure (2b)). This finding is confirmed by the spatial analysis of patient glioma biopsy samples in Jiao et al. [22]. On the other hand, by adding a death process for any motility state in the GoG model, we can recapitulate the Allee effect (bistability between extinction and growth) as found in Boettger et al [33]. A cell population at the “fluid” state will go extinct (motile cells do not proliferate but still have a probability to die,) where as systems in the “solid” state will always grow until carrying capacity. Therefore, in the bistability regime cell sensing properties lead to multicellular growth control.

Implications in multicellular pattern control: Increased cell sensing β represents the physiological tissue dynamics, since it allows the system to control its behavior. By tuning β and the ratio of motile/resting cells, the system exhibits a bistable behavior between a "fluid"/mesenchymal-type and a “solid”/epithelial-type tissue phase. This kind of tissue level switch is of utmost importance in physiological processes such as wound healing or embryogenesis [34]. After a tissue injury, the healing is characterized by a "fluid" diffusive expansion of fibroblast cells, that adopt a migratory phenotype via EMT [35]. After covering the wound, the "solid" phase emerges as cells stop the migration program and proliferate to finalize tissue repair. In the abscence of proliferation, the bistable switch from the "solid" to the "fluid" state could potentially explain the jamming phase transition observed in epithelial colonies, under EGF modulation and Rab5a knock-out [36]. Typically, the "solid" multicellular phase is prone to the emergence of pattern formation, which frequently occurs in physiological epithelial tissues [35, 37]. When EMT is combined with a Notch-Delta cell-cell communication then epithelial/immotile cell clusters emerge [26], as observed in our IBM and mean-field simulations. Adding proliferation

in the GoG model, the type and the size of such emerging patterns require a tight regulation of microenvironmental sensing β and proliferation rate r, as indicated by Fig. 11a, since the critical wavelength k_c depends on the ratio $\frac{\beta}{r}$. Such Turing patterns are in agreement with previous GoG studies [38], where Turing patterns where emerging.

At this point, we would like to outline some ideas on the experimental validation of LEUP. An ideal experimental system would involve a cell population exhibiting either EMT/MET or migration/proliferation plasticity. As an example of EMT/MET plasticity, we can choose an epithelial cell colony. The main premise of LEUP is that cells equip Bayesian inference to decide over their phenotypes, expressed as a combination of a sensed microenvironmental distribution of external variables and a phenotypic prior (cell intrinsic variable). As internal variables we can choose the cell motility and proliferation rates. The former can be measured by particle image velocitometry (PIV) and the latter by live imaging or Ki67 (or BrdU) staining. Also we could pharmacologically control cell proliferation. As extrinsic variables we can assume the local cell density and in particular the corresponding fluctuations. The second moment of the local cell density depends on the average local cell density (first moment) and can be manipulated by localized cell ablation or by mixing with non-cellular material, such as collagen, thus creating different spatial cell distributions. After splitting the experiments in training and test data, one needs to fit the LEUP parameters, i.e. the sensitivity to the microenvironment and the interaction radius, according to training data in order to predict the experimentally measured test data. A good basis for our approach are experiments by Puliafito [39], which show that increasing density in a colony of epithelial kidney cells may induce transitions from the mesenchymal to the epithelial phenotype (MET). In the future, one could also investigate the inverse transition and control the changes in the local cell density distribution. Such an experimental setting would allow us to support the LEUP hypothesis.

The focus of this paper is the application of the LEUP idea to study collective effects of different types of phenotypic plasticity allowing reversible adaptations to changing environmental conditions. However, the LEUP idea can also be applied to analyze collective effects of irreversible cell fate determination playing a key role in cell differentiation and development. In particular, early embryonic tissue consists of a mix of proliferating pluripotent progenitor cells [40].]. In this case the cellular microenvironment possesses a high phenotypic entropy since all possible phenotypes are still equiprobable. Subsequently, spatial mechanochemical perturbations produce appropriate microenvironmental changes leading to so-called stem cell niches that allow for cell differentiation [41]. Pluripotent cells sense a broad spectrum of microenvironmental cues aiming to find an appropriate differentiation niche. As pluripotent cells further specialize, they irreversibly fine-tune their sensing machinery to adapt to the corresponding microenvironment which further differentiates to a specific tissue (implying minimization of microenvironmental entropy). A particularly important cell fate specification mechanism is the Notch-Delta interaction [42], where neighboring cells use lateral inhibition to adapt distinct fates (high Delta low Notch and vice versa). The collective effect of this mechanism is the evolution of the system to low entropy organized patterns [43]. Currently, we try to understand cell fate determination and tissue development under the prism of LEUP. Our vision is that describing phenotypic plasticity and cell fate determination within the LEUP framework allows to develop a unified theory for both types of cell decision-making and their corresponding dynamics.

In conclusion, our study shows how individual LEUP-driven cell decisions the dynamics at the tissue level and how knowledge of collective cell decision-making can be used to control of growth and pattern.

Acknowledgment

AB thanks the International Graduate School of HZI, Braunschweig. HH, MMH and PM gratefully acknowledge the funding support of the Helmholtz Association of German Research Centers—Initiative and Networking Fund for the project on Reduced Complexity Models (ZT-I-0010). HH and PM acknowledge the funding support of MicMode-I2T (01ZX1710B) by the Federal Ministry of Education and Research (BMBF). HH is supported by MulticellML (01ZX1707C) of the Federal Ministry of Education and Research (BMBF) and the Volkswagen-stiftung for the its support within the "Life?" programm (96732). SS acknowledges financial support co-financed by the European Social Fund (ESF) and tax funds in accordance with the budget adopted by the members of the Saxon State Parliament. Part of the current work was inspired and initiated when NK was visiting the Helmholtz Centre for Infection Research and he would like to express his gratitude for the warm hospitality of the institute. NK would also like to acknowledge financial support from the Faculty of Science and Engineering of University of Chester. The authors thank the Centre for Information Services and High Performance Computing (ZIH) at TU Dresden for providing an excellent infrastructure.

References

1. Bowsher, C. G. & Swain, P. S. Environmental sensing, information transfer, and cellular decision-making. *Curr. Opin. Biotechnol.* **28,** 149–155. doi:10.1016/j.copbio.2014.04.010 (2014).
2. Alberts, B. *Molecular Biology of the Cell* (New York: W. W, 2015).
3. Prochazka, L., Benenson, Y. & Zandstra, P. W. Synthetic gene circuits and cellular decision-making in human pluripotent stem cells. *Curr. Opin. Syst. Biol.* **5,** 93–103. doi:10.1016/j.coisb.2017.09.003 (2017).
4. Balázsi, G., van Oudenaarden, A. & Collins, J. J. Cellular decision making and biological noise: from microbes to mammals. *Cell* **144,** 910–925. doi:10.1016/j.cell.2011.01.030 (2011).
5. Price, T. D., Qvarnström, A. & Irwin, D. E. The role of phenotypic plasticity in driving genetic evolution. *Proc. R. Soc. London. Ser. B Biol. Sci.* **270,** 1433–1440. doi:10.1098/rspb.2003.2372 (2003).
6. Kalluri, R. & Weinberg, R. A. The basics of epithelial-mesenchymal transition. *The J. of Cli. Inves.* **120,** 1786–1786. doi:10.1172/JCI39104C1 (2010).

7. Radisky, D. C. Epithelial-mesenchymal transition. *J. Cell Sci.* **118,** 4325–4326. doi:10.1242/jcs.02552. eprint: https://jcs.biologists.org/content/118/19/4325.full.pdf (2005).

8. Thiery, J. P. Epithelial–mesenchymal transitions in development and pathologies. *C. Opi. Cell Bio.* **15,** 740–746. doi:https://doi.org/10.1016/j.ceb.2003.10.006 (2003).

9. Hatzikirou, H., Basanta, D., Simon, M., Schaller, K. & Deutsch, A. 'Go or Grow': the key to the emergence of invasion in tumour progression? *Math. Med. Biol.* **29,** 49–65. doi:10.1093/imammb/dqq011 (2010).

10. Jerby, L. *et al.* Metabolic associations of reduced proliferation and oxidative stress in advanced breast cancer. *Cancer Res.* **72,** 5712–5720. doi:10.1158/0008-5472.CAN-12-2215 (2012).

11. Hoek, K. S. *et al.* In vivo switching of human melanoma cells between proliferative and invasive states. *Cancer Res.* **68,** 650–656. doi:10.1158/0008-5472.CAN-07-2491 (2008).

12. Kohrman, A. Q. & Matus, D. Q. Divide or Conquer: Cell Cycle Regulation of Invasive Behavior. *T. Cell Bio.* **27,** 12–25. doi:10.1016/j.tcb.2016.08.003 (2017).

13. Mehta, P. & Schwab, D. J. Energetic costs of cellular computation. *Proceedings of the National Academy of Sciences* **109,** 17978–17982. doi:10.1073/pnas.1207814109. eprint: https://www.pnas.org/content/109/44/17978.full.pdf (2012).

14. Sourjik, V. & Berg, H. C. Receptor sensitivity in bacterial chemotaxis. *Proceedings of the National Academy of Sciences* **99,** 123–127. doi:10.1073/pnas.011589998. eprint: https://www.pnas.org/content/99/1/123.full.pdf (2002).

15. Ueda, M. & Shibata, T. Stochastic Signal Processing and Transduction in Chemotactic Response of Eukaryotic Cells. *Biophysical Journal* **93,** 11–20. doi:10.1529/biophysj.106.100263 (2007).

16. Purvis, J. E. & Lahav, G. Encoding and decoding cellular information through signaling dynamics. *Cell* **152,** 945–956. doi:10.1016/j.cell.2013.02.005. arXiv: NIHMS150003 (2013).

17. Hatzikirou, H. Statistical mechanics of cell decision-making: the cell migration force distribution. *J. Mech. Behav. Mater.* **27,** 1–7. doi:10.1515/jmbm-2018-0001 (2018).

18. Box, G. E. P., Hunter, W. G. & Hunter, J. S. *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building* (1979).

19. Hakim, V. & Silberzan, P. Collective cell migration: a physics perspective. *Rep. on Prog. in Phys.* **80,** 076601. doi:10.1088/1361-6633/aa65ef (2017).

20. Schienbein, M. & Gruler, H. Langevin equation, Fokker-Planck equation and cell migration. *Bull. Math. Bio.* **55,** 585–608. doi:10.1007/BF02460652 (1993).

21. Bhatnagar, P. L., Gross, E. P. & Krook, M. A Model for Collision Processes in Gases. I. Small Amplitude Processes in Charged and Neutral One-Component Systems. *Phys. Rev.* **94,** 511–525. doi:10.1103/PhysRev.94.511 (3 1954).

22. Jiao, Y., Berman, H., Kiehl, T.-R. & Torquato, S. Spatial Organization and Correlations of Cell Nuclei in Brain Tumors. *PLOS ONE* **6,** 1–9. doi:10.1371/journal.pone.0027323 (2011).

23. Milster, S., Nötel, J., Sokolov, I. M. & Schimansky-Geier, L. Eliminating inertia in a stochastic model of a micro-swimmer with constant speed. *The Euro. Phys. Jour. Sp. Topics* **226,** 2039–2055. doi:10.1140/epjst/e2017-70052-8 (2017).

24. Murray, J. D. *Mathematical Biology II. Spatial Models and Biomedical Applications* 3rd ed. (ed Murray, J. D.) 1067. doi:10.1007/b98869 (Springer New York, New York, NY, 2003).

25. Jolly, M. K. *et al.* Towards elucidating the connection between epithelial - mesenchymal transitions and stemness. *J. Roy. Soc. Int.* **11,** 20140962. doi:10.1098/rsif.2014.0962. eprint: https://royalsocietypublishing.org/doi/pdf/10.1098/rsif.2014.0962 (2014).

26. Boareto, M. *et al.* Jagged–Delta asymmetry in Notch signaling can give rise to a Sender/Receiver hybrid phenotype. *PNAS* **112,** E402–E409. doi:10.1073/pnas.1416287112. eprint: https://www.pnas.org/content/112/5/E402.full.pdf (2015).

27. Barua, A., Nava-Sedeño, J. M. & Hatzikirou, H. A least microenvironmental uncertainty principle (LEUP) as a generative model of collective cell migration mechanisms. *bioRxiv.* doi:10.1101/404889. eprint: https://www.biorxiv.org/content/early/2019/10/04/404889.full.pdf (2019).

28. Vicsek, T., Czirók, A., Ben-Jacob, E., Cohen, I. & Shochet, O. Novel type of phase transition in a system of self-driven particles. *Phys. rev. lett.* **75,** 1226 (1995).

29. Grossmann, R., Schimansky-Geier, L. & Romanczuk, P. Self-propelled particles with selective attraction–repulsion interaction: from microscopic dynamics to coarse-grained theories. *New J. Phys.* **15,** 085014 (2013).

30. Banyard, J. & Bielenberg, D. R. The role of EMT and MET in cancer dissemination. *Conn. Tis. Res.* **56.** PMID: 26291767, 403–413. doi:10.3109/03008207.2015.1060970. eprint: https://doi.org/10.3109/03008207.2015.1060970 (2015).

31. Hanahan, D. & Weinberg, R. A. The Hallmarks of Cancer. *Cell* **100,** 57–70. doi:10.1016/S0092-8674(00)81683-9 (2000).

32. Dahlman, E. & Watanabe, Y. SU-F-T-109: A Shortcoming of the Fisher-Kolmogorov Reaction-Diffusion Equation for Modeling Tumor Growth. *Med. Phys.* **43,** 3486–3487. doi:10.1118/1.4956245. eprint: https://aapm.onlinelibrary.wiley.com/doi/pdf/10.1118/1.4956245 (2016).

33. Böttger, K. *et al.* An Emerging Allee Effect Is Critical for Tumor Initiation and Persistence. *PLOS Comp. Bio.* **11** (ed Alber, M. S.) e1004366. doi:10.1371/journal.pcbi.1004366 (2015).

34. Barriere, G., Fici, P., Gallerani, G., Fabbri, F. & Rigaud, M. Epithelial Mesenchymal Transition: a double-edged sword. *Clinic.Trans. Med.* **4,** 14. doi:10.1186/s40169-015-0055-4 (2015).

35. Yan, L. & Bi, D. Multicellular Rosettes Drive Fluid-solid Transition in Epithelial Tissues. *Phys. Rev. X* **9,** 011029. doi:10.1103/PhysRevX.9.011029 (1 2019).

36. Malinverno, C. *et al.* Endocytic reawakening of motility in jammed epithelia. *Nature Mat.* **16,** 587–596. doi:10.1038/nmat4848 (2017).

37. Heller, E. & Fuchs, E. Tissue patterning and cellular mechanics. *J. Cell Bio.* **211,** 219–231. doi:10.1083/jcb.201506106 (2015).

38. Pham, K. *et al.* Density-dependent quiescence in glioma invasion: instability in a simple reaction–diffusion model for the migration/proliferation dichotomy. *J. Bio. Dyn.,* 1–18. doi:10.1080/17513758.2011.590610 (2011).

39. Puliafito, A. *et al.* Collective and single cell behavior in epithelial contact inhibition. *Proceedings of the National Academy of Sciences of the United States of America* **109,** 739–44. doi:10.1073/pnas.1007809109 (2012).

40. Lanza, R. in *Essentials of Stem Cell Biology (Third Edition)* (eds Lanza, R. & Atala, A.) Third Edition, xxi (Academic Press, Boston, 2014). doi:https://doi.org/10.1016/B978-0-12-409503-8.00046-9.

41. Jones, D. L. & Wagers, A. J. No place like home: anatomy and function of the stem cell niche. *Nature Reviews Molecular Cell Biology* **9,** 11–21. doi:10.1038/nrm2319 (2008).

42. Artavanis-Tsakonas, S., Rand, M. D. & Lake, R. J. Notch Signaling: Cell Fate Control and Signal Integration in Development. *Science* **284,** 770–776. doi:10.1126/science.284.5415.770. eprint: https://science.sciencemag.org/content/284/5415/770.full.pdf (1999).

43. Bocci, F., Onuchic, J. N. & Jolly, M. K. Understanding the Principles of Pattern Formation Driven by Notch Signaling by Integrating Experiments and Theoretical Models. *Frontiers in Physiology* **11,** 929. doi:10.3389/fphys.2020.00929 (2020).

Supplementary Information

Symbol	Explanation
X_i	Phenotype of i-th cell
Y_i	Phenotype of i-th neighbourhood cells
N_i^0	Number of cells in i-th cell's microenvironment having phenotype $(X_i = 0)$
N_i^1	Number of cells in i-th cell's microenvironment having phenotype $(X_i = 1)$
N_i^ϕ	Number of free slots in i-th cell's microenvironment
V	Total capacity
N_T	Total number of phenotypes
ℓ	Radius of the microenvironment
ν	Exchange rate
r	Growth rate
β	LEUP sensitivity
$\rho_0(x)$	Mean density of the resting cells at position x
$\rho_1(x)$	Mean density of the migratory cells at position x
D_0	Diffusion coefficient of resting cells
D_1	Diffusion coefficient of migratory cells

S.1. Calculation of microenvironmental entropy

We assume that a maximum number of N cells is present inside the cell's microenvironment, where ℓ is the radius of the microenvironment (Fig.1a). Since there is the possibility of free slots, if there are less than N cells in the cell neighbourhood, we have started by assuming a trinomial distribution for the i-th cell is $P\left(N_i^1, N_i^\phi\right)$ where the number N_i^0 of cells having phenotype $X_i = 0$ and by the number N_i^1 of cells having phenotype $X_i = 1$. In addition, a number N_i^ϕ of empty slots are included inside the microenvironment. The joint probability is defined by

$$P\left(N_i^1, N_i^\phi\right) = \frac{N!}{N_i^1! N_i^\phi! \left(N - N_i^\phi - N_i^1\right)!} p^{N_i^1} \theta^{N_i^\phi} \left(1 - p - \theta\right)^{\left(N - N_i^\phi - N_i^1\right)}, \tag{50}$$

where p and θ are the probabilities of having a number N_i^1 of cells with phenotype $(X_i = 1)$ out of N cells and having a number N_i^ϕ of free slots. The conditional probability of having a

number N_i^1 of cells present in the microenvironment given a number N_i^ϕ of free slots is

$$\begin{aligned} P\left(N_i^1 \mid N_i^\phi\right) &= \frac{P\left(N_i^1, N_i^\phi\right)}{P\left(N_i^\phi\right)} \\ &= \frac{\frac{N!}{N_i^1! N_i^\phi! \left(N-N_i^\phi-N_i^1\right)!} p^{N_i^1} \theta^{N_i^\phi} \left(1-p-\theta\right)^{\left(N-N_i^\phi-N_i^1\right)}}{\frac{N!}{N_i^\phi!\left(N-N_i^\phi\right)!} \theta^{N_i^\phi} \left(1-\theta\right)^{\left(N-N_i^\phi\right)}} \\ &= \binom{N-N_i^\phi}{N_i^1} \left(\frac{p}{1-\theta}\right)^{N_i^1} \left(1-\frac{p}{1-\theta}\right)^{N-N_i^1-N_i^\phi} \\ &= \boldsymbol{B}\left(N-N_i^\phi, \frac{p}{1-\theta}\right) \\ &= \boldsymbol{B}\left(N_T, \frac{p}{1-\theta}\right) \end{aligned} \tag{51}$$

Now, we use the form of binomial distribution to calculate the entropy of the Binomial distribution. Please note that $P(N_i^1 \mid N_i^\phi)$ can be written as $P(N_i^1 \mid N_T)$ due to Binomial distribution.

$$S\left(N_i^1 \mid N_T\right) = S\left(Y_i{=}\left[N_i^1 \mid N_T\right]\right) = \frac{1}{2}\log_2\left(2\pi e \frac{N_T p}{1-\theta}\left(1-\frac{p}{1-\theta}\right)\right), \tag{52}$$

where

$$\begin{aligned} \frac{p}{1-\theta} &= \frac{N_i^1}{N}\frac{1}{1-\frac{N_i^\phi}{N}} = \frac{N_i^1}{N-N_i^\phi} = \frac{N_i^1}{N_T}, \\ 1-\frac{p}{1-\theta} &= \frac{N_i^0}{N-N_i^\phi} = \frac{N_i^0}{N_T}. \end{aligned} \tag{53}$$

According to LEUP we have to evaluate the microenvironmental entropy of $S\left(Y_i \mid X_i = 0\right)$ and $S\left(Y_i \mid X_i = 1\right)$ to calculate the probability of the internal states $(X_i = 0)$ and $(X_i = 1)$

$$\begin{aligned} S\left(Y_i{=}\left[N_i^1 \mid N_T\right] \mid X_i = 1\right) &= \frac{1}{2}\log_2\left(2\pi e\left(\frac{\left(N_T-1\right)\left(N_i^1-1\right)}{\left(N_T-1\right)}\right)\left(\frac{N_i^0}{\left(N_T-1\right)}\right)\right), \\ S\left(Y_i{=}\left[N_i^1 \mid N_T\right] \mid X_i = 0\right) &= \frac{1}{2}\log_2\left(2\pi e\left(\frac{\left(N_T-1\right)\left(N_i^1\right)}{\left(N_T-1\right)}\right)\left(\frac{N_i^0-1}{\left(N_T-1\right)}\right)\right). \end{aligned} \tag{54}$$

From the Gaussian approximation we can write the entropy difference as

$$\begin{aligned} \Delta S &= S(Y_i{=}\left[N_i^1 \mid N_T\right] | X_i = 1) - S(Y_i{=}\left[N_i^1 \mid N_T\right] | X_i = 0), \\ &= \frac{1}{2}\ln\left[\frac{N_i^0\left(N_i^1-1\right)}{N_i^1\left(N_i^0-1\right)}\right]. \end{aligned} \tag{55}$$

S.2. Calculation of cell proliferation rate

In our go-or-grow model only resting cells are allowed to proliferate but not the moving cells. The growth rate depends in a specific way on the number of moving and resting cells in the

microenvironment. In particular, we assume that the per- capita growth rate for resting cells is a linearly decreasing function of N_i^0 and N_i^1, and is also decreasing with the number of migratory cells and a constant per-capita death rate d_1. Accordingly,

$$\begin{aligned}\frac{d\langle N_i^0\rangle}{dt} &= \left(\langle W^+\rangle_{N_i^0,N_i^1} - \langle W^-\rangle_{N_i^0,N_i^1}\right) \\ &= h_1\rho_0 - q\left(\rho_0+\rho_1\right)\rho_0 - d_1\rho_0 \\ &= r\rho_0\left(1-\rho_0-\rho_1\right)\end{aligned} \tag{56}$$

where $\frac{q}{r} = 1$ and $r = h_1 - d_1$.

S.4. Phase space diagram for null sensitivity case

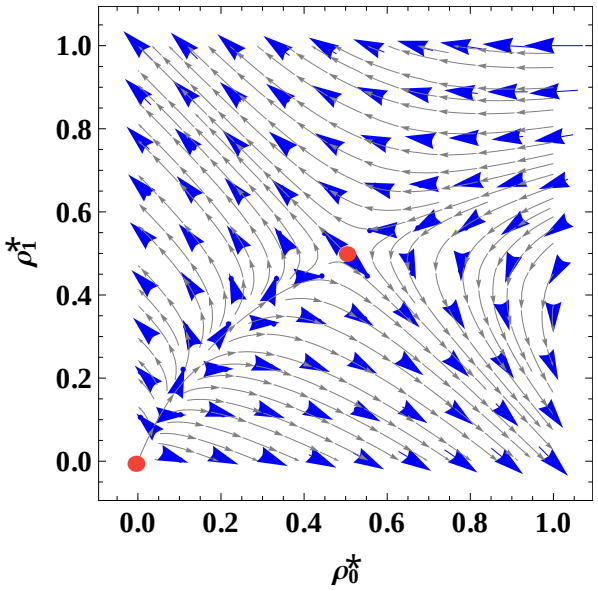
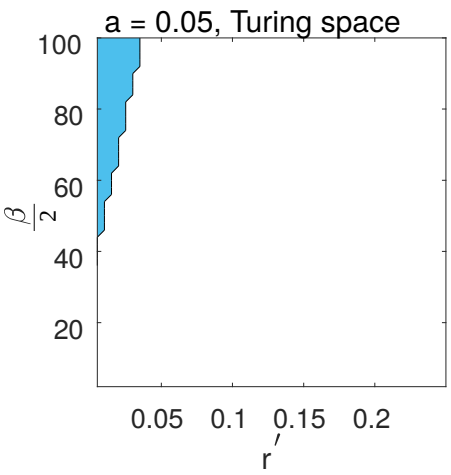
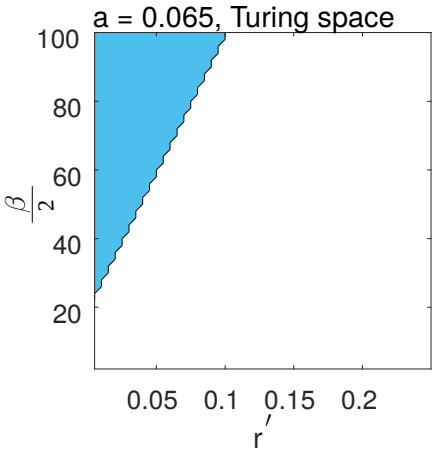


Figure 9. Phase space diagram of ρ_0^* and ρ_1^* for $r = 1$ where two fixed points are marked by red circles.

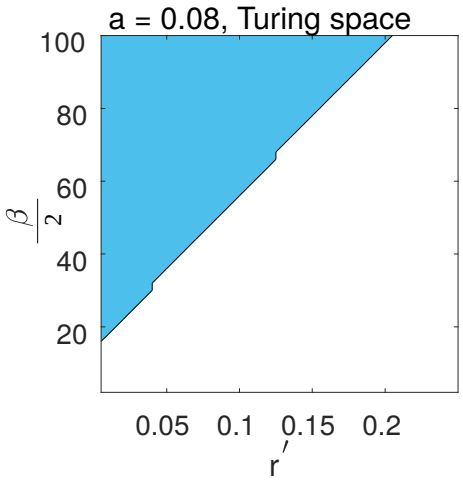
S.5. Turing space for pattern formation



(a) a is fixed at 0.05.



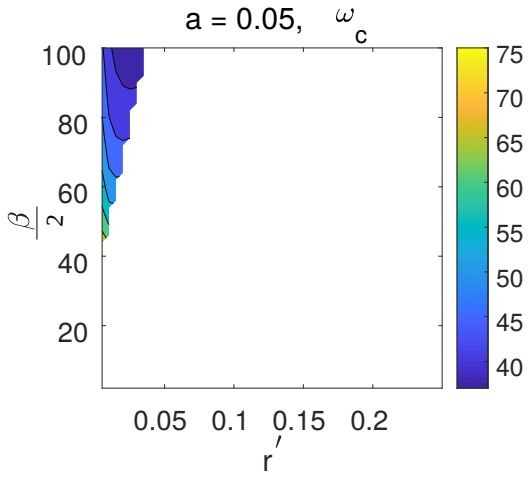
(b) a is fixed at 0.065.



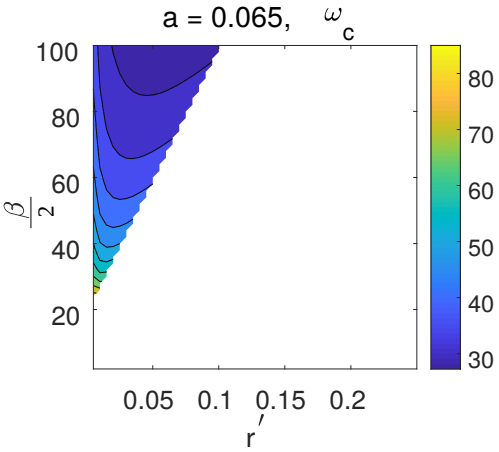
(c) a is fixed at 0.08.

Figure 10. Phase space diagram of β vs. $r^{'}$ for different values of a (**a-c**), where the Turing space is marked in blue. a is defined by $\frac{1}{V}$.

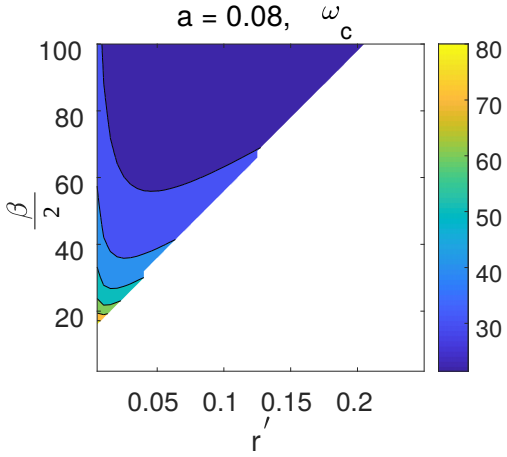
S.6. Critical wavelengths inside the Turing space for pattern formation



(a) a is fixed at 0.05.



(b) a is fixed at 0.065.



(c) a is fixed at 0.08.

Figure 11. Phase space diagram of β vs. $r^{'}$ for different values of a (**a-c**), where critical wavelengths (ω_c) have been plotted inside the Turing space. a is defined by $\frac{1}{V}$.

S.7. Plot of resting probability

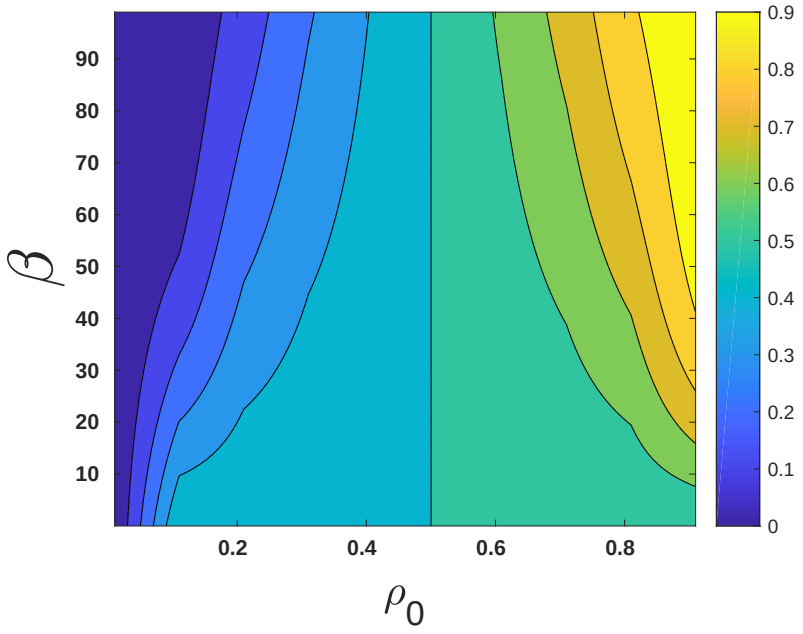


Figure 12. Plot of resting probability with respect to β and ρ_0. We have fixed the value of $\frac{1}{V}$ to 0.01.

S.8. Critical sensitivity vs. inverse capacity graph from IBM

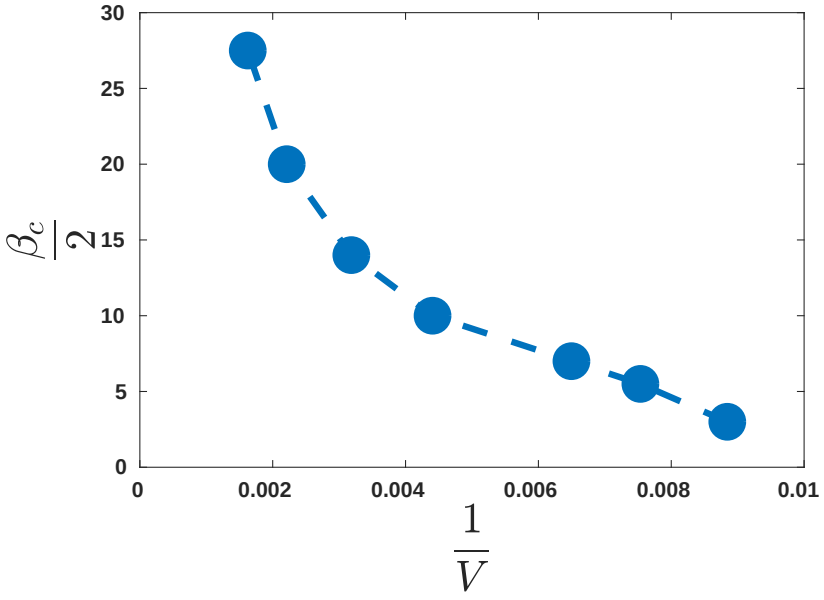


Figure 13. Critical sensitivity with respect to the inverse capacity of the IBM over 5 simulations. Throughout the simulations, we kept the total density at 0.2 and the total number of cells was 500.