MDPI

Article

Close to optimal cell sensing ensures the robustness of tissue differentiation process: the avian photoreceptor mosaic case

Arnab Barua[1], Alireza Beygi[1] and Haralampos Hatzikirou[1,2,*]

1 Centre for Information Services and High Performance Computing, Technische Universität Dresden, Nöthnitzer Straße 46, 01062 Dresden, Germany
2 Mathematics Department, Khalifa University, P.O. Box 127788, Abu Dhabi, UAE
* Correspondence: haralampos.hatzikirou@ku.ac.ae

Version February 14, 2021 submitted to Entropy

Abstract: The way that progenitor cell fate decisions and the associated environmental sensing are regulated to ensure the robustness of the spatial and temporal order in which cells are generated towards a fully differentiating tissue still remains elusive. Here, we investigate how cells regulate their sensing intensity and radius to guarantee the required thermodynamic robustness of a differentiated tissue. In particular, we are interested in finding the conditions where dedifferentiation at cell level is possible (microscopic reversibility) but tissue maintains its spatial order and differentiation integrity (macroscopic irreversibility). In order to tackle this, we exploit the recently postulated Least microEnvironmental Uncertainty Principle (LEUP) to develop a theory of stochastic thermodynamics for cell differentiation. To assess the predictive and explanatory power of our theory, we challenge it against the avian photoreceptor mosaic data. By calibrating a single parameter, the LEUP can predict the cone color spatial distribution in the avian retina and, at the same time, suggest that such a spatial pattern is associated with quasi-optimal cell sensing. By means of the stochastic thermodynamics formalism, we find out that thermodynamic robustness of differentiated tissues depends on cell metabolism and cell sensing properties. In turn, we calculate the limits of the cell sensing radius that ensure the robustness of differentiated tissue spatial order. Finally, we further constrain our model predictions to the avian photoreceptor mosaic.

Keywords: cell differentiation; cell decision-making; cell sensing; cell metabolism; stochastic thermodynamics; Least microEnvironmental Uncertainty Principle (LEUP)

1. Introduction

Decision-making is a process to identify important choices and responses which depends on some basic criteria [1]. The precise mechanisms and processes guiding cell fate decisions, as in the regulation of cell differentiation, still remain elusive. Thereby, cells irreversibly acquire new fates by following a hierarchical lineage, where pluripotent stem cells in a proper microenvironment (stem cell niche) differentiate into, for example, bone, muscle, epithelial, and further specialized cells. What we can observe during cell differentiation process includes the dramatic change of geometry, shape, gene expression inside the cell, etc. [2,3].

The theory behind cell differentiation has been formalized in a metaphorical way by C. H. Waddington [4–6], which it facilitates the study of cell differentiation, in particular, in the mathematical framework of dynamical systems, concerning single-cell fate decisions [7–9]. Waddington has represented the developmental process as a series of cell decisions that can be represented as valleys and forks in a developmental landscape. In his landscape diagram, changes

in the environment can be viewed as a small manipulation in slope that would lead to one channel in the landscape being favored over another, so that an adult cell could show a different phenotype starting from the same genotype. However, this theory does not take into account cell sensing and the corresponding interactions that take place in a tissue.

Cells can sense rapid changes of their milieus, where they mainly exploit two ways to decrease their sensing errors: (**1**) to escalate the number of receptors and (**2**) to increase the number of measurements per receptors over a particular time, which is controlled by the downstream signaling pathways [10]. There is an extensive literature on fundamental limits of cell sensing in both static and changing environments [11–13]. While cell differentiation has been well studied for single cells, it is still an open question how the interaction between individual cell fate decisions and microenvironmental sensing can ensure robust spatiotemporal patterns at the tissue scale.

In order to tackle the aforementioned questions and problems, in this paper, we apply the Least microEnvironmental Uncertainty Principle (LEUP) which is essentially a statistical mechanical theory for cell decision-making [14–16] to the problem of cell differentiation. Cells decide and change their fates (or, phenotypes) according to the LEUP, in order to minimize their local microenvironmental entropy. Formulated in the language of statistical physics and Bayesian statistics, cells try to optimize their prior to reduce the joint entropy of their own fate and microenvironmental variables. Cells can acquire knowledge about their microenvironment by spreading pseudopodia and reshaping their cytoskeleton (which is an energetically expensive process). Interestingly, the optimized prior can help to reduce the amount of the expenditure during the adaptation processes of the microenvironment [17]. A decrease of the microenvironmental entropy from a higher to a lower value can be observed in various biological settings, for example, mesenchymal-epithelial transition, wound healing, collective cell migration. Thus, from an empiricist point of view, one can take the assumption of the LEUP to guide the phenomenon of cell differentiation. Similar thoughts can also be found in an influential work by W. Bialek [18]. Since the LEUP remains a hypothesis, we use the avian photoreceptor mosaic to challenge our theoretical predictions, where we also establish a relation between quasi-optimal cell sensing and cone color spatial distribution in the avian retina [19].

In the present paper, we develop a thermodynamic-like theory using the tools of stochastic thermodynamics for a generic cell differentiation process. The main biological assumption is that differentiated cell can reverse to an undifferentiated state, for instance, by the process of carcinogenesis. Stochastic thermodynamics allows us to understand the conditions that, even though single-cell dedifferentiation is possible (microscopic reversibility), the system (tissue) is still able to be robust and it maintains its spatial order and differentiation integrity (macroscopic irreversibility). Stochastic thermodynamics is a suitable tool for systems where small scale dynamics matters (e.g., soft matters, active matters, and biological systems), in such cases the higher-order moments dominate [20–24]; for different experimental applications of stochastic thermodynamics, see Refs. [25,26]. In addition, to formulate the laws of thermodynamics in the mesoscopic scale (specifically, at the level of trajectory) has also been investigated recently [27]. To describe cell-level dedifferentiation process in the context of stochastic thermodynamics, we apply the fluctuation theorem, specifically the Crooks' relation [28,29]. The fluctuation theorem, which can be considered as the heart of stochastic thermodynamics, initially had been derived to explain how irreversibility at macroscopic level emerges from the underlying reversible dynamics and to estimate the probability of the violation of the second law of thermodynamics within a short amount of time for systems at small scales [30–32].

By combining stochastic thermodynamics considerations with the LEUP, we are aspire to answer questions such as: How single-cell fate decisions of progenitor cells are controlled by spatiotemporal order of the generated differentiated tissue and what should be the critical interaction radius for a particular cell type. In other words, we provide an understanding of how cells regulate their sensing radius to decide on their own fates and maintain the expected tissue integrity. Furthermore, we can determine the threshold values of the sensing radius which ensure the robustness of cell differentiation concerning the spatial order, in the particular case of the avian photoreceptor mosaic.

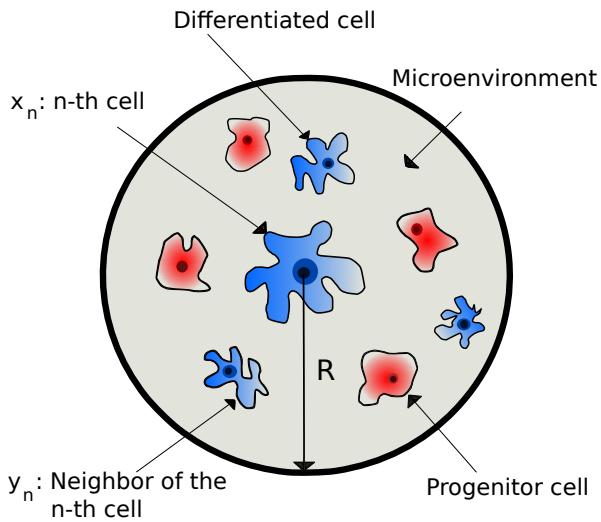

Figure 1. The schematic diagram of the microenvironment of a differentiated cell, where R is the sensing radius. Microenvironment is being composed of a distribution of pluripotent progenitors and differentiated cells.

This paper is organized as follows. In Sec. 2, we review the basic concepts of the LEUP and its connection to statistical mechanics. In Sec. 3, we demonstrate how optimal microenvironmental sensing is associated with differentiated tissue spatial configuration, therein, we examine our theory using the data obtained from the avian photoreceptor mosaic. We apply the fluctuation theorem to cell differentiation in Sec. 4 and demonstrate the thermodynamic robustness of this process in the case of the avian retina development. In Sec. 5, we show how cell sensing radius and total entropy production are related and determine the limits of the sensing radius in order the tissue development to be robust. Finally, we conclude and discuss our results in Sec. 6.

2. Least microEnvironmental Uncertainty Principle

Cell collects information from its microenvironment and based on that it takes actions (phenotypic decisions). In other words, cell reacts to the environmental information by changing its own state. By adopting the notations of Ref. [14], we denote internal variables (such as gene expression, RNA molecules, metabolites, etc.) of the nth cell as x_n and external variables (such as chemical signals, ligands, cellular density, translational proteins, etc.) of the nth cell as y_n, which it can also include internal states of other cells within the interaction radius of the nth cell, i.e., $y_n = \{x_m : |r_n - r_m| \leq R\}$, where internal variable(s) of the neighboring cell(s) is(are) denoted as x_m with position vector r_m, r_n is position vector of the nth cell, and R is the maximum microenvironmental sensing radius of the nth cell. We consider behavior of cell, as it reacts to its microenvironment, a Bayesian decision-maker [14,33], such that

$$P\left(x_n \mid y_n\right) = \frac{P\left(y_n \mid x_n\right) P\left(x_n\right)}{P\left(y_n\right)}, \tag{1}$$

where $P\left(y_n \mid x_n\right)$ is the probability distribution of microenvironmental information being collected by cell. In other words, it is the probability that cell perceives all other cells, chemicals, and nutrients in its surrounding. $P\left(x_n\right)$ is the prior probability of cell internal states. If these two distributions are multiplied and be divided by the prior probability of external states $P\left(y_n\right)$, then (1) implies that the resulting quantity is the posterior probability distribution of internal states, that is, the likely decision to be made by cell over the available information. In Fig. 1, we have shown the schematic diagram of the microenvironment of a differentiated cell as is being composed of both progenitors and differentiated cells.

Due to the fact that the collection of microenvironmental data and thus the evaluation of $P\left(y_n \mid x_n\right)$ is energetically expensive, in order cell to be energetically efficient, we propose the idea that cell optimizes its prior $P\left(x_n\right)$ and, at the same time, it minimizes total uncertainty of the composite system of cell and its microenvironment. By exploiting the ideas that are developed within the context of statistical mechanics, we formulate the problem as finding the appropriate prior probability distribution of cell internal states in such a way that it minimizes the joint entropy of cell-microenvironment system.

We define $S\left(x_n, y_n\right)$ as the joint entropy of the composite system of cell and its microenvironment, $S\left(x_n\right)$ as the internal entropy of cell, and $S\left(y_n \mid x_n\right)$ as the entropy of microenvironment which is sensed by cell. These quantities are connected through the relation: $S\left(x_n, y_n\right) = S\left(x_n\right) + S\left(y_n \mid x_n\right)$. We consider the definition of entropy in terms of probability distributions as $S = -\Sigma_i P_i \ln P_i$ (Shannon entropy), where in the case of continuous set of states the summation symbol is to be replaced by an integral. Now, our task is to find $P\left(x_n\right)$ such that it minimizes $S\left(x_n, y_n\right)$. This is equivalent to the optimization of $S\left(x_n\right)$ with respect to $P\left(x_n\right)$, which its corresponding variational problem reads

$$\begin{aligned}\frac{\delta}{\delta P(x_n)}\Bigg\{ & S(x_n) - \beta\left[\int P(x_n) S(y_n \mid x_n) \mathrm{d}x_n - \overline{S}(y_n \mid x_n)\right] \\ & - \mu\left[\int P(x_n) N(y_n \mid x_n) \mathrm{d}x_n - \overline{N}(y_n \mid x_n)\right] \\ & - \lambda\left[\int P(x_n) \mathrm{d}x_n - 1\right]\Bigg\} = 0,\end{aligned} \tag{2}$$

where $\delta / \delta P\left(x_n\right)$ is the functional derivative. We have introduced three constraints with the corresponding Lagrange multipliers as β, μ, and λ. The first two constraints impose bounds on the microenvironmental entropy which is sensed by cell and the number of neighboring cells in its surrounding, with the corresponding mean values of $\overline{S}\left(y_n \mid x_n\right)$ and $\overline{N}\left(y_n \mid x_n\right)$, respectively; and the last constraint is concerned with the normalization of the probability. We note that other knowledge about the system, such as an explicit model in terms of statistical observables, can be incorporated as additional constraints. The solution of Eq. (2) then reads

$$P\left(x_n\right) = \frac{e^{-\beta S(y_n|x_n) - \mu N(y_n|x_n)}}{Z}, \tag{3}$$

where $Z = \int e^{-\beta S(y_n|x_n) - \mu N(y_n|x_n)} \mathrm{d}x_n$ is a normalization constant. We note that μ resembles the chemical potential in statistical mechanics, i.e., it is the amount of energy pumped in system per new cell. Thus, μ can be thought of as a quantity which is proportional to cell division time: $\mu \propto \tau_{\text{div.}}$. In the next section, we elaborate on the relation of β and optimal microenvironmental sensing.

As a concrete example, if the microenvironmental probability distribution which is sensed by the nth cell follows a Gaussian distribution with the variance σ_n^2, assuming that y_n is scalar, and hence $S\left(y_n \mid x_n\right) = (1/2) \ln[2\pi e \sigma_n^2(y_n \mid x_n)]$, and by considering $\mu = 0$, then (3) reduces to

$$P\left(x_n\right) = \frac{\sigma_n^{-\beta}(y_n \mid x_n)}{\sum_j \sigma_j^{-\beta}(y_n \mid x_j)}. \tag{4}$$

Now, we are in a position to establish an explicit formula for cell internal entropy. To this end, by exploiting (3), we obtain

$$S\left(x_n\right) = -\int P\left(x_n\right) \ln P\left(x_n\right) \mathrm{d}x_n = \beta \langle S\left(y_n \mid x_n\right)\rangle_{x_n} + \mu \langle N\left(y_n \mid x_n\right)\rangle_{x_n} + \ln Z, \tag{5}$$

where $\langle ... \rangle$ denotes the expectation value. Relation (5) is the fundamental thermodynamic-like identity of the LEUP.

Making usage of the notions which are developed within the context of thermodynamics, we can define phenotypic internal energy which is sensed by cell as

$$U_{\text{int.}}(x_n, y_n) = \langle S(y_n \mid x_n)\rangle_{x_n} + \frac{\mu}{\beta}\langle N(y_n \mid x_n)\rangle_{x_n}, \tag{6}$$

and the Helmholtz-like free energy then reads

$$A_H = U_{\text{int.}} - \frac{1}{\beta}S = -\frac{1}{\beta}\ln Z. \tag{7}$$

3. Minimal cell sensing noise is associated with differentiated tissue spatial configuration

The steady-state distribution of (4) which is derived by the LEUP, under the assumption of Gaussian distributed microenvironment and $\mu = 0$, can be associated with different established statistical results. In particular, for $\beta = 1$, we reproduce distributions proportional to the well-known Jeffreys prior, see also Ref. [14]. Interestingly, if we assume that $\beta = 2$, then we recover the so-called *minimum variance estimator* when fusing multiple scalar estimates.

Following [34], let us assume extrinsic variables $y_1, y_2, ..., y_n$, such as ligand concentrations, cell densities, etc., that are pairwise statistically uncorrelated and are normally distributed, i.e., $y_i \sim \mathcal{N}(\mu_i, \sigma_i^2(x_i))$, where $1 \leq i \leq n$. In turn, we assume intrinsic variables $x_1, x_2, ..., x_n$, that correspond to cellular sensors and downstream process (e.g., they can represent receptor concentrations or downstream activated proteins after a ligand-receptor binding) of the aforementioned microenvironmental variables. We define the average sensed microenvironment as the average of the extrinsic signals y_i's weighted by the distribution of the cell sensors x_i's, i.e., $Z(y_1, y_2, ..., y_n) = \sum_{j=1}^{n} P_j y_j$, where $\sum_{j=1}^{n} P_j = 1$. The distribution of internal variables P_i that minimizes the variance, aka noise of the sensed microenvironment, is given by the following formula:

$$P_i = \frac{\sigma_i^{-2}(x_i)}{\sum_j \sigma_j^{-2}(x_j)}, \tag{8}$$

which has the same form as the steady state of the LEUP (4), for $\beta = 2$. The above result is pivotal since it connects the cell fate decision-making with optimal microenvironmental sensing in terms of minimization of sensing noise, which results into a specific spatial phenotypic distribution. In the following, we explore the validity of these results in the case of the avian photoreceptor mosaic.

3.1. Application: predicting the cone color distribution in the avian photoreceptor mosaic

In this subsection, we focus on a particular multicellular system, namely, the photoreceptor cone mosaic of the avian retina. Retina plays a vital role in the visual perception of vertebrates allowing for light sampling and signal transfer to the corresponding optic nerves. The photoreceptor layer is responsible for the efficient light collection and color encoding of the surrounding structures. The avian retina possesses one of the most sophisticated cone photoreceptor systems among vertebrates [19]. Birds have five types of cones including four single cones, namely, green, red, blue, and violet, and a double cone, whereas in humans there are only three: red, green, and blue. In the case of birds, these cones are ordered on the retina in a non-perfect hexagonal manner. The five avian cone types exist as five independent spatial mosaics that are all embedded within a single monolayered epithelium along with a population of rod photoreceptors. In Ref. [19], the spatial organization of these photoreceptor mosaics has been analyzed, where the authors have found that the five cone types are present in the characteristic ratios at the latest stage of development, as is illustrated in Fig. 2.

Our goal is to reproduce the in vivo observed cone distributions of [19] by using the aforementioned LEUP-driven results. In particular, we intend to fit the parameter β in formula (4) using the photoreceptors data. Therefore, let us consider we have $n = 5$ different cone phenotypes,

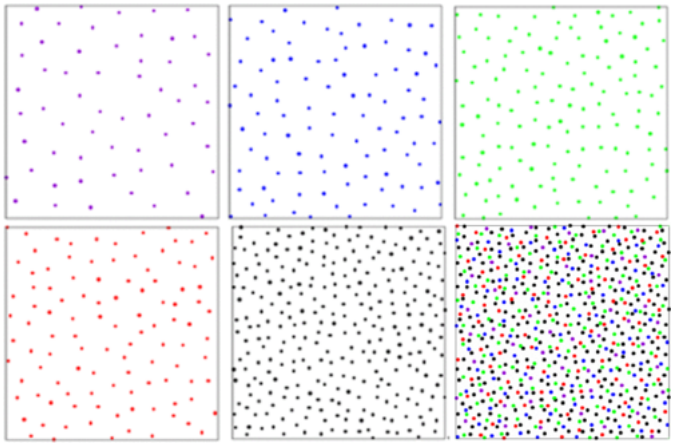

Figure 2. The experimentally obtained chicken cone photoreceptors arrangement in space. In the upper panels, the configurations from left to right correspond to violet, blue, and green species, respectively. In the lower panels, from left to right, the configurations correspond to red, double species, and the overall pattern, respectively. The figure is adapted from [35].

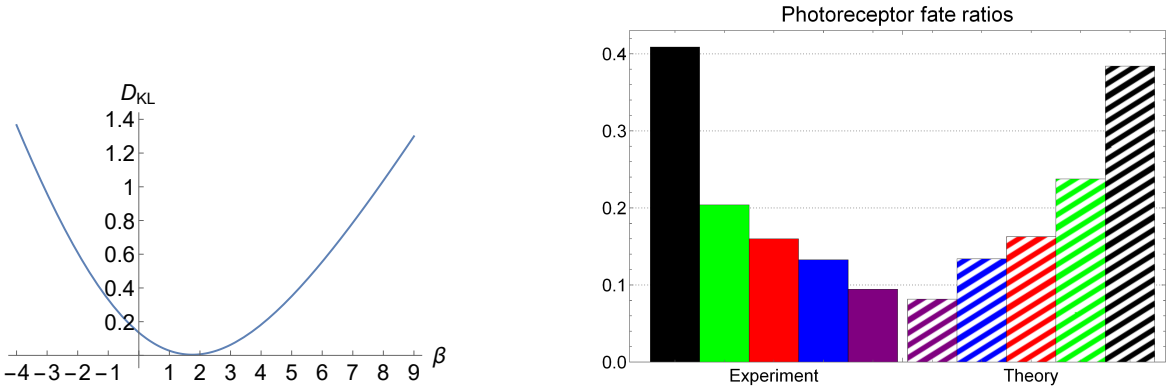

Figure 3. In the left panel, the Kullback-Leibler divergence D_{KL} is depicted as a function of β. D_{KL} reaches its minimum of ≈ 0.004 at $\beta \approx 1.754$. The experimentally obtained photoreceptor fate ratios is compared to the LEUP for this particular β, in the right panel.

that correspond to the different colors, with the corresponding cone occurrence probabilities P_i's, where $i = \delta, g, r, b, v$ (δ stands for double, g for green, r for red, b for blue, and v for violet). In turn, we denote σ_i as the standard deviation for a given spatially local neighborhood related to the Nearest Neighbor Distribution (NND). In Ref. [19], the authors have exploited the Delaunay triangulation in order to calculate the NND for each cone and they have reported the first and second moments of it. In order to compare the LEUP color distributions against the in vivo ones, we employ the Kullback-Leibler divergence. In turn, we numerically search for the β's which appropriately fit the physiologically observed photoreceptor fate ratios. In the left panel of Fig. 3, we have depicted the Kullback-Leibler divergence, i.e., $D_{KL} = \sum_i P_i \ln(P_i/Q_i)$, as a function of β, where i denotes different cone cells, and P_i and Q_i correspond to the probabilities obtained from the experimental data and the LEUP, respectively. The plot demonstrates that D_{KL} has a minimum of ≈ 0.004 at $\beta \approx 1.754$. We have compared the in vivo observed cone distributions to the LEUP, in the right panel of Fig. 3. Interestingly, the β values that provide the best fit are in the range $1.378 \leq \beta \leq 2.136$, where the values of the Kullback-Leibler divergence are in the order of magnitude 10^{-3}. Thus, it demonstrates that the values of $\beta \approx 2$ explain the observed data, implying that the avian cone mosaic distribution is associated with a cell sensing process that minimizes the corresponding noise.

4. Thermodynamic robustness of differentiated tissues: the fluctuation theorem

In this section, we determine the thermodynamic constraints of two coarse-grained cell states that correspond to pluripotent (s) and differentiated (d) states. Then, we apply the obtained results to the particular case of the avian cone cells differentiation and the formation of photoreceptor mosaics. To this end, first we show how microstates (internal variables) are related to the microenvironmental information and heat transfer (cell metabolism).

In this context, a cellular *microstate* corresponds to a cell phenotype that lives in a tissue, which could be gene expression, RNA molecules, receptor distribution, etc. In other words, microstate gives information about the internal states of cell. We label these internal variables as x_s and x_d corresponding to pluripotent and differentiated cells, respectively. We define a cellular *macrostate* as a statistical observable (e.g., average) of a cell microenvironment that involves multiple cells of different phenotypes. Macrostates contain information about external variables which are labeled as y_s and y_d. The macrostate y_s is assumed to describe a microenvironment of pluripotent progenitor or stem cells characterized by the microstate x_s; the macrostate y_d accordingly describes a microenvironment of differentiated cells characterized by the microstate x_d. We denote the number of pluripotent cells neighboring to a cell with microstate x_s as $N(y_s \mid x_s)$ and the number of differentiated cells neighboring to a cell with microstate x_d as $N(y_d \mid x_d)$. The total number of pluripotent and differentiated cells inside system are denoted as $N(s)$ and $N(d)$, respectively.

Now, based on (3), we can write the probability of cell to be in the microstate x_s with the corresponding macrostate s as

$$P(x_s) = \frac{e^{-\beta_s S(y_s|x_s) - \mu_s N(y_s|x_s)}}{Z_1}, \tag{9}$$

and the probability of being in the microstate x_d with the corresponding macrostate d as

$$P(x_d) = \frac{e^{-\beta_d S(y_d|x_d) - \mu_d N(y_d|x_d)}}{Z_2}. \tag{10}$$

It is known that for a system which is coupled to a set of heat baths and is in a time-symmetrically driven nonequilibrium state, the Crooks' theorem is applicable [28,29]. In this case, the dynamics follows Brownian motion, i.e., there are no extreme *jumps* in the system state. For any trajectory to be initially at $x_s(0)$ and is going through microstates $x_s(t)$ over time τ, the Crooks' theorem implies that [29],

$$\beta' \Delta Q = \ln \left[\frac{w[x_s(t)]}{w[x_s(\tau - t)]} \right], \tag{11}$$

where $\beta' \equiv 1/T$, which T is the temperature of heat bath, ΔQ is the total heat released into bath over the course of $x_s(t)$, and $w[x_s(t)]$ is the probability of trajectory $x_s(t)$. Equation (11) demonstrates that when there is a forward state change, system loses heat to reservoir and in the case of a time-reversed path, there is a heat gain from reservoir; this, in turn, implies that a forward trajectory is more probable than a time-reversed one, thus, (11) substantiates a relation between heat and irreversibility at a microscopic level (cell level). As is demonstrated in Ref. [29], Eq. (11) is also valid for other kinds of steady-state probability distributions besides the classical Boltzmann distribution. In our case, such Brownian jumps are interpreted as changes in phenotypes and the associated heat losses are assumed to be due to potential changes in cell metabolism. However, there are other heat loss sources that are disregarded since they act on shorter time scales, such as physical friction, changes in the cytoskeleton, etc.

By fixing the starting point of trajectory as $x_s(0) = x_s$ and the ending point as $x_s(\tau) = x_d$, then taking the average over all trajectories from x_s to x_d, results in

$$\frac{w(x_s \to x_d; \tau)}{w(x_d \to x_s; \tau)} = \langle \exp[\beta' \Delta Q^{\tau}_{x_s \to x_d}] \rangle_{x_s \to x_d}, \tag{12}$$

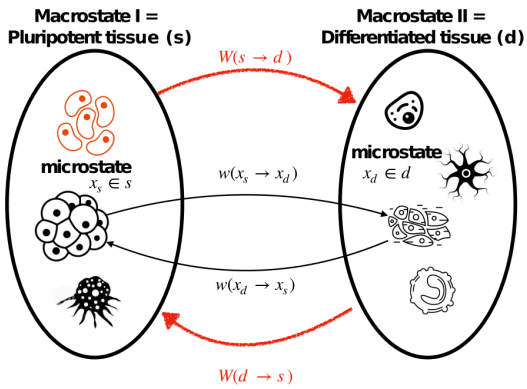

Figure 4. Microscopic and macroscopic transitions between two distinct cell types.

where $w(x_s \to x_d; \tau)$ is the transition probability that system is found to be in the microstate x_d at time τ, given that system was initially in the microstate x_s.

Now that we have microscopic relation of (12) at our disposal, we can study macroscopic (tissue) consequences of that. Indeed, the probabilistic description of macroscopic states of two distinct cell types, from which one can understand the phenomenon of irreversibility at a macroscopic level, can be constructed as [29],

$$W(s \to d) = \int_d \mathrm{d}x_d \int_s \mathrm{d}x_s \; P(x_s \mid s) w(x_s \to x_d), \tag{13}$$

and

$$W(d \to s) = \int_s \mathrm{d}x_s \int_d \mathrm{d}x_d \; P(x_d \mid d) w(x_d \to x_s), \tag{14}$$

where $P(x_s \mid s)$ is the probability that system to be in the microstate x_s, given that it is observed in the macrostate s, and $w(x_s \to x_d)$ is defined as before, where we have omitted τ for the notational convenience. The transition probability $W(s \to d)$ in (13) implies the likelihood that cell to be observed in the macrostate d while it was initially prepared in the macrostate s; accordingly, (14) is understood in the same fashion, i.e., the likelihood that a microenvironment being prepared in the state d to satisfy the microenvironment s after another time interval τ. These processes are illustrated in Fig. 4.

By taking the ratio of (13) and (14), we obtain

$$\frac{W(d \to s)}{W(s \to d)} = \frac{\int_s \mathrm{d}x_s \int_d \mathrm{d}x_d [P(x_d \mid d)/P(x_d)] P(x_d)\, w(x_d \to x_s)}{\int_d \mathrm{d}x_d \int_s \mathrm{d}x_s [P(x_s \mid s)/P(x_s)] P(x_s)\, w(x_s \to x_d)}, \tag{15}$$

where we have multiplied and divided the numerator and the denominator by $P(x_d)$ and $P(x_s)$, respectively. We define the pointwise mutual information for individual trajectories as $i_1 = \ln[P(x_s \mid s)/P(x_s)]$ and $i_2 = \ln[P(x_d \mid d)/P(x_d)]$, and then by taking these definitions into account and replacing $P(x_d)$ by its corresponding relation in (10), (15) reduces to

$$\begin{aligned} \frac{W(d \to s)}{W(s \to d)} &= \frac{\int_s \mathrm{d}x_s \int_d \mathrm{d}x_d e^{i_2} [e^{-\beta_d S(y_d|x_d) - \mu_d N(y_d|x_d)}/Z_2] w(x_d \to x_s)}{\int_d \mathrm{d}x_d \int_s \mathrm{d}x_s e^{i_1} P(x_s) w(x_s \to x_d)} \\ &= \frac{\int_s \mathrm{d}x_s \int_d \mathrm{d}x_d e^{G} e^{i_2 + i_1 - i_1} (Z_1/Z_2) [e^{-\beta_s S(y_s|x_s) - \mu_s N(y_s|x_s)}/Z_1] w(x_d \to x_s)}{\int_d \mathrm{d}x_d \int_s \mathrm{d}x_s e^{i_1} P(x_s) w(x_s \to x_d)}, \end{aligned} \tag{16}$$

where

$$G \equiv -[\beta_d S(y_d \mid x_d) - \beta_s S(y_s \mid x_s)] - [\mu_d N(y_d \mid x_d) - \mu_s N(y_s \mid x_s)].$$

By exploiting (9) and (12), we can rewrite (16) as

$$\frac{W(d \to s)}{W(s \to d)} = \frac{\int_s \mathrm{d}x_s \int_d \mathrm{d}x_d e^G e^{i_1} e^{\Delta i} (Z_1/Z_2) P(x_s) \langle e^{-\beta' \Delta Q^{\tau}_{x_s \to x_d}} \rangle_{x_s \to x_d} w(x_s \to x_d)}{\int_d \mathrm{d}x_d \int_s \mathrm{d}x_s e^{i_1} P(x_s) w(x_s \to x_d)}, \tag{17}$$

where $\Delta i \equiv i_2 - i_1$.

Now, (17) can be expressed in terms of the average over all trajectories from the ensemble of microstates x_s which correspond to the macrostate s to the ensemble of microstates x_d which correspond to the macrostate d while each path is weighted by its probability as

$$\frac{W(d \to s)}{W(s \to d)} = \langle \langle \exp[-\beta' \Delta Q^{\tau}_{x_s \to x_d}] \rangle_{x_s \to x_d} \exp[-\Delta S_{\mathrm{LEUP}}] \exp[-\Delta N] \exp[\Delta i] \exp[\ln(Z_1/Z_2)] \rangle_{s \to d}, \tag{18}$$

where

$$\begin{aligned} \Delta S_{\mathrm{LEUP}} &\equiv \beta_d S(y_d \mid x_d) - \beta_s S(y_s \mid x_s), \\ \Delta N &\equiv \mu_d N(y_d \mid x_d) - \mu_s N(y_s \mid x_s). \end{aligned} \tag{19}$$

By rearranging (18) as

$$\begin{aligned} &\langle \langle \exp[-\beta' \Delta Q^{\tau}_{x_s \to x_d}] \rangle_{x_s \to x_d} \exp[-\Delta S_{\mathrm{LEUP}}] \exp[-\Delta N] \\ &\qquad \times \exp[\Delta i] \exp[\ln(Z_1/Z_2)] \exp[\ln\{W(s \to d)/W(d \to s)\}] \rangle_{s \to d} = 1, \end{aligned} \tag{20}$$

and using the Jensen's inequality, i.e., $\exp[\langle X \rangle] \leq \langle \exp[X] \rangle$, and the fact that $e^x \geq 1 + x$, we arrive at

$$\langle \beta' \Delta Q^{\tau}_{x_s \to x_d} \rangle_{s \to d} + \langle \Delta S_{\mathrm{LEUP}} \rangle_{s \to d} + \langle \Delta N \rangle_{s \to d} - \langle \Delta i \rangle_{s \to d} - \ln \left[\frac{Z_1}{Z_2} \right] + \ln \left[\frac{W(d \to s)}{W(s \to d)} \right] \geq 0. \tag{21}$$

We note that if s and d correspond to the same identical classes, i.e., $W(d \to s) = W(s \to d)$, then the last term of (21) vanishes. Now, by defining the remaining terms as *total entropy production*, that is,

$$f \equiv \langle \beta' \Delta Q^{\tau}_{x_s \to x_d} \rangle_{s \to d} + \langle \Delta S_{\mathrm{LEUP}} \rangle_{s \to d} + \langle \Delta N \rangle_{s \to d} - \langle \Delta i \rangle_{s \to d} - \ln \left[\frac{Z_1}{Z_2} \right], \tag{22}$$

then (21) implies that total entropy production is always non-negative. In other words, (21) can be considered as a generalization of the second law of thermodynamics (see also Ref. [29]).

The above inequality is reformulation of the fluctuation theorem for tissue differentiation. It relates thermodynamic properties of system to the LEUP; and as is obtained under general assumptions, it is applicable to a general cell and tissue differentiation process. In the next subsection, we show how (21) implies the thermodynamic robustness of cell differentiation, in the particular case of the avian cone photoreceptor mosaics formation.

4.1. Application: differentiated photoreceptor mosaics are thermodynamically robust

In this subsection, we use the fluctuation theorem, culminated in (21), to demonstrate the robustness of the avian photoreceptor mosaics. By defining $s \to d$ as a pluripotent tissue differentiates to a more specialized one and denoting its corresponding forward transition probability as p_f, i.e., $W(s \to d) = p_f$, and $d \to s$ as differentiated tissue dedifferentiates to the pluripotent one with its backward transition probability of p_b, i.e., $W(d \to s) = p_b$, we can rewrite (21) in terms of p_f and p_b as

$$\frac{p_f}{p_b} \leq \exp \left[\langle \beta' \Delta Q^{\tau}_{x_s \to x_d} \rangle_{s \to d} + \langle \Delta S_{\mathrm{LEUP}} \rangle_{s \to d} + \langle \Delta N \rangle_{s \to d} - \langle \Delta i \rangle_{s \to d} - \ln \left[\frac{Z_1}{Z_2} \right] \right], \tag{23}$$

where we have already defined the exponent of the exponential in the right-hand side of (23) as total entropy production, see (22).

In order to obtain a thermodynamic constraint which ensures the robustness of differentiated tissue, first we assume that there exists a maximum forward transition probability from s to d in such a way that

$$\frac{p_{f\,\max}}{p_b} = \exp\left[\langle \beta' \Delta Q^{\tau}_{x_s \to x_d} \rangle_{s\to d} + \langle \Delta S_{\mathrm{LEUP}} \rangle_{s\to d} + \langle \Delta N \rangle_{s\to d} - \langle \Delta i \rangle_{s\to d} - \ln\left[\frac{Z_1}{Z_2}\right]\right]. \tag{24}$$

To simplify (24) more, we note that based on (5), we can write

$$\begin{aligned} S(x_s) &= \beta_s \langle S(y_s \mid x_s) \rangle_{x_s} + \mu_s \langle N(y_s \mid x_s) \rangle_{x_s} + \ln Z_1, \\ S(x_d) &= \beta_d \langle S(y_d \mid x_d) \rangle_{x_d} + \mu_d \langle N(y_d \mid x_d) \rangle_{x_d} + \ln Z_2, \end{aligned} \tag{25}$$

and then by subtracting these two and taking the average over all trajectories from s to d, we obtain

$$\langle \Delta S \rangle_{s\to d} \equiv \langle S(x_d) - S(x_s) \rangle_{s\to d} = \langle \Delta S_{\mathrm{LEUP}} \rangle_{s\to d} + \langle \Delta N \rangle_{s\to d} - \ln\left[\frac{Z_1}{Z_2}\right], \tag{26}$$

where we have used the definitions of ΔS_{LEUP} and ΔN as given in (19). Now, (24) can be written as

$$\frac{p_{f\,\max}}{p_b} = \exp\left[\langle \beta' \Delta Q^{\tau}_{x_s \to x_d} \rangle_{s\to d} + \langle \Delta S \rangle_{s\to d} - \langle \Delta i \rangle_{s\to d}\right]. \tag{27}$$

The immediate implication of (27) is that, differentiated tissue is thermodynamically robust if

$$\langle \beta' \Delta Q^{\tau}_{x_s \to x_d} \rangle_{s\to d} + \langle \Delta S \rangle_{s\to d} > \langle \Delta i \rangle_{s\to d}. \tag{28}$$

In the following, we present the avian cone cells differentiation and the corresponding photoreceptor mosaics formation as an example of which the above inequality is satisfied. In other words, we demonstrate the robustness of the avian retina development and the irreversibility of the time arrow in this particular process.

First, we note that as heat dissipation (the first term in the left-hand side of (28)) depends on metabolic pathways, this implies the crucial role of cell metabolism in the process of cell differentiation. In addition, consumption of glucose depends upon cell types. Progenitor or pluripotent stem cells use glucose as the primary metabolites for anaerobic glycolysis (fermentation) pathway: Glucose + 2 ADP + 2 Phosphate $\rightarrow$ 2 Lactate + 2 H^+ + 2 ATP, where as a result of the breakdown of glucose to lactic acid the amount of energy around 109.4 kJ/mol is released [36]. Differentiated cells use glucose to produce carbon dioxide and water by using aerobic glycolysis (respiration) reaction: Glucose + 6 O_2 + 36 ADP + 36 Phosphate $\rightarrow$ 6 CO_2 + 6 H_2O + 36 ATP, where the released energy is around 2820 kJ/mol [36]. Thus, $\langle \Delta Q^{\tau}_{x_s \to x_d} \rangle_{s\to d}$ which is the total heat released during the journey from s to d is 2929.4 kJ/mol. As these reactions have taken place in $T = 310$ K, we have: $\langle \beta' \Delta Q^{\tau}_{x_s \to x_d} \rangle_{s\to d} \approx 9.450$, where we have dropped its dimension as we have set the Boltzmann constant to be 1, in this paper.

Now, in order to investigate (28) for the avian cone cells differentiation, we need to calculate the value of $\langle \Delta S \rangle_{s\to d}$ for this particular process. Kram et al. [19] have reported different color percentages of the green, red, blue, violet, and double cone cells inside the avian retina (see Supporting Information of the mentioned reference). By assuming these numbers as the occurrence probabilities of the

corresponding cones, we can write: $P_g \approx 0.204$, $P_r \approx 0.160$, $P_b \approx 0.133$, $P_v \approx 0.094$, and $P_\delta \approx 0.409$. Thus, we are able to calculate entropies of the individual cone cells and as a result that of $S(x_d)$ as

$$\begin{aligned} S(x_d) &= -\sum_i P_i \ln P_i \\ &\approx 0.324 + 0.293 + 0.268 + 0.222 + 0.366 \\ &= 1.473. \end{aligned} \tag{29}$$

Now, the entropy difference between differentiated and stem cells reads

$$\langle \Delta S \rangle_{s\to d} = \langle S(x_d) - S(x_s) \rangle_{s\to d} = -0.136, \tag{30}$$

where $S(x_s)$ behaves as the entropy of a uniform distribution, i.e., $S(x_s) = -\sum_{i=1}^{5}(1/5)\ln(1/5) = \ln 5$; this is due to the fact that for pluripotent stem cells there are no yet color preferences.

In order differentiated tissue maintains its spatial order and integrity, (28) imposes a lower bound on heat dissipation (cell metabolism) as

$$\langle \beta' \Delta Q^{\tau}_{x_s\to x_d} \rangle_{s\to d} > -\langle \Delta S \rangle_{s\to d}, \tag{31}$$

where we have set $\langle \Delta i \rangle_{s\to d} \to 0$ for simplicity. Inequality (31) is strongly hold for the values obtained here, that is, $9.450 \gg 0.136$. This implies that the development of the avian retina is highly robust. In other words, the arrow of time is almost irreversible in this process.

5. Cell sensing radius limits for robust tissue development

In this section, we derive a relationship between total entropy production and cell sensing radius in the particular case of progenitor cells differentiation into the avian cone photoreceptors. We calculate the limits of the sensing radius and, in the parametric space of the LEUP parameters, we suggest the biologically and physically acceptable regions for robust tissue differentiation.

In (22), we have introduced total entropy production as

$$f = \langle \beta' \Delta Q^{\tau}_{x_s\to x_d} \rangle_{s\to d} + \langle \Delta S_{\text{LEUP}} \rangle_{s\to d} + \langle \Delta N \rangle_{s\to d} - \langle \Delta i \rangle_{s\to d} - \ln\left[\frac{Z_1}{Z_2}\right], \tag{32}$$

where thermodynamic properties of system are related to the LEUP quantities. In order to calculate $\langle \Delta S_{\text{LEUP}} \rangle_{s\to d}$, by assuming the microenvironmental probability distributions as Gaussians, we have

$$\langle \Delta S_{\text{LEUP}} \rangle_{s\to d} = \frac{\beta_d}{2}\ln[2\pi e \sigma_d^2] - \frac{\beta_s}{2}\ln[2\pi e \sigma_s^2], \tag{33}$$

where we have used (19). The above can be simplified to

$$\langle \Delta S_{\text{LEUP}} \rangle_{s\to d} = (\beta_d - \beta_s)\ln[2\pi e]^{1/2} + \frac{\beta_d}{2}\ln\sigma_d^2 - \frac{\beta_s}{2}\ln\sigma_s^2. \tag{34}$$

Now, we need to postulate the forms of σ_d^2 and σ_s^2 and their scaling with the corresponding cell sensing radius. Jiao et al. in [35] have found that the differentiated retina mosaic is hyperuniform, that is, $\sigma_d^2 = VR^A$, where $A < D$, and $D = 2, 3$; on the other hand, for progenitor cells, we assume a Poisson distribution, i.e., $\sigma_s^2 = UR^D$. Here, R is the sensing radius, and V and U are the densities of the cone and progenitor cells neighbors, respectively. Plugging these formulas in (34) leads to

$$\langle \Delta S_{\text{LEUP}} \rangle_{s\to d} = (\beta_d - \beta_s)\ln[2\pi e]^{1/2} + \frac{1}{2}\ln\left[\frac{V^{\beta_d}}{U^{\beta_s}}\right] + \frac{A\beta_d - D\beta_s}{2}\ln R. \tag{35}$$

Another term of (32) which needs to be dealt with is $\langle \Delta N \rangle_{s\to d}$. From (19), we have

$$\langle \Delta N \rangle_{s\to d} = \langle \Sigma_{i=1}^{5} (\mu_{i,d} N_{i,d} - \mu_{i,s} N_{i,s}) \rangle_{s\to d}, \tag{36}$$

where i counts different types of cones: green, red, blue, violet, and double. We note that the population at each tissue reads as $N_{i,j} = \rho_{i,j} R^D$, where $\rho_{i,j}$ is the density of microenvironment and $j \in \{s, d\}$. Thus, (36) reduces to

$$\langle \Delta N \rangle_{s\to d} = \langle \Sigma_{i=1}^{5} (\mu_{i,d} \rho_{i,d} - \mu_{i,s} \rho_{i,s}) \rangle_{s\to d} R^D \equiv \Delta\tilde{\mu} R^D. \tag{37}$$

Putting all terms together, total entropy production (32) becomes

$$f(R, \beta_d, \beta_s) = C_0 + C_1(\beta_d - \beta_s) + \frac{1}{2} \ln \left[\frac{V^{\beta_d}}{U^{\beta_s}} \right] + \frac{A\beta_d - D\beta_s}{2} \ln R + \Delta\tilde{\mu} R^D, \tag{38}$$

where we have assumed $\langle \beta' \Delta Q^{\tau}_{x_s \to x_d} \rangle_{s\to d}$, $-\langle \Delta i \rangle_{s\to d}$, and $-\ln(Z_1/Z_2)$ as constants and have grouped them together as C_0, and $C_1 \equiv (1/2)\ln(2\pi e)$.

Now that we have an explicit formula for total entropy production at our disposal, we can obtain the optimal sensing radius for which total entropy production reaches its extremum. To this end, we note that the first derivative of (38) vanishes at

$$R_c = \left(\frac{-\Gamma}{\Delta\tilde{\mu} D} \right)^{1/D}, \tag{39}$$

where $\Gamma \equiv (A\beta_d - D\beta_s)/2$. When R_c is a positive quantity, then the inequality $\Gamma/\Delta\tilde{\mu} < 0$ should always be satisfied. In order to determine the conditions for which R_c minimizes or maximizes total entropy production, we calculate the second derivative of (38) at (39), and obtain

$$\left. \frac{\partial^2 f}{\partial R^2} \right|_{R_c} = -\frac{\Gamma D}{R_c^2}. \tag{40}$$

Thus, R_c minimizes total entropy production if $\Gamma < 0$ and it maximizes if $\Gamma > 0$.

In the left panel of Fig. 5, we have illustrated total entropy production (38) as a function of R for a particular set of the LEUP parameters which minimizes total entropy production based on (40). The values of the curve above the R-axis (gray line) are biologically and physically relevant as total entropy production is positive, i.e., it ensures the robustness of differentiation process. Moreover, this plot shows that as we move away from R_c, (positive) total entropy production is rapidly increasing. The right panel of the figure illustrates f as a function of R and β_d, where the acceptable regions lie above the gray plane. Interestingly, minimization of total entropy production occurs only for $\Delta\tilde{\mu} > 0$, which assumes a decrease in the proliferative activity of differentiated phenotypes.

In Fig. 6, the case for which total entropy production reaches its maximum is shown. This figure illustrates that (positive) total entropy production is bounded in this case. The immediate implication of this is tissue dedifferentiation is possible (negative total entropy production) for a large sensing radius. If we want to have a positive total entropy production, in order to avoid reversibility of differentiated tissue, we must have some sort of fine-tuning in order to restrict the values of R in such a way that they lead to a positive f. According to Bialek's postulated biophysical principles [18], fine-tuning in Nature is not favorable. In the case of the avian photoreceptors, we can further constrain our parameters and identify further arguments to exclude this case.

5.1. Application: the average sensing radius of the avian cone cell

In this subsection, we calculate the average sensing radius for each photoreceptor and by exploiting (38) we look for the regions, in the parametric space of (β_s, β_d) and $(\beta_s, \beta_d, \Delta\tilde{\mu})$, which result in a positive total entropy production and ensure the robustness of the differentiated tissue.

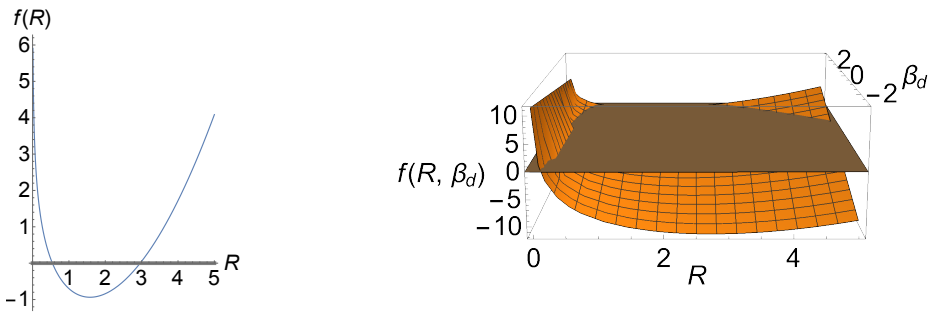

Figure 5. The left panel shows total entropy production as a function of sensing radius for $C_0 = -1$, $\beta_d = \beta_s = 3$, $V = U = 1$, $A = 1$, $D = 2$, and $\Delta\tilde{\mu} = 0.3$. Based on (40), this specific set of parameters leads to a total entropy production which has a minimum of ≈ -0.937 at $R_c \approx 1.581$. In the right panel, β_d is also treated as a variable. Due to the fact that total entropy production is always positive, the curve/surface above the gray line/plane is only biologically and physically acceptable.

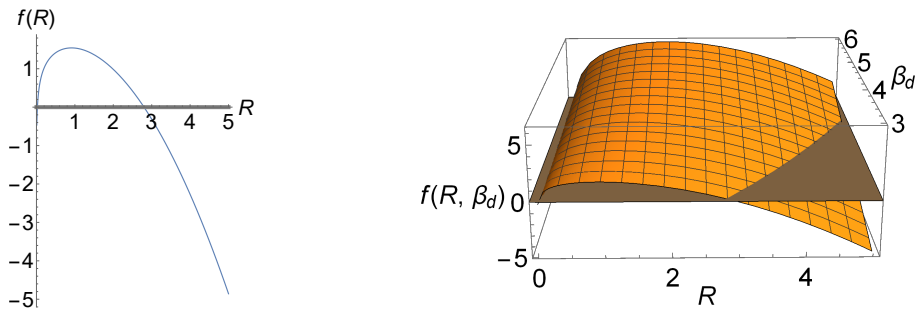

Figure 6. Total entropy production f as a function of R and (R, β_d) is shown in the left and right panels, respectively. In the left panel, we have fixed the parameters as: $C_0 = -1$, $\beta_d = 3$, $\beta_s = 1$, $V = U = 1$, $A = 1$, $D = 2$, and $\Delta\tilde{\mu} = -0.3$. $f(R)$ reaches its maximum of ≈ 1.542 at $R_c \approx 0.913$. In the right panel, β_d is also considered as a variable. In both panels, the only biologically and physically acceptable regions lie above the gray line and the gray plane, as total entropy production should be positive; this makes total entropy production a bounded function in contrast to the case depicted in Fig. 5.

In the retina of a bird, as explained in the previous sections, the five types of photoreceptor cells, namely, green, red, blue, violet, and double cones, form mosaic structures. From the experimental data (see Supporting Information of [19]), we have information about the average standard deviation of the Nearest Neighbor Distribution (NND) for each color as: $\sigma_g \approx 1.248$, $\sigma_r \approx 1.548$, $\sigma_b \approx 1.729$, $\sigma_v \approx 2.292$, and $\sigma_\delta \approx 0.948$. In addition, Jiao et al. [35] have found how the variance of an avian cone σ^2 is related to its sensing radius R as

$$\sigma^2(R) = M_1 R^2 + M_2 R \ln R + M_3 R, \tag{41}$$

where M_1, M_2, and M_3 have been calculated for each color (see Table I of Ref. [35]). By having this information at our disposal, we are able to calculate the average sensing radius for each photoreceptor as

$$\overline{R}_\delta \approx 0.786, \quad \overline{R}_g \approx 1.426, \quad \overline{R}_r \approx 2.006, \quad \overline{R}_b \approx 2.288, \quad \overline{R}_v \approx 3.553. \tag{42}$$

In the beginning of this section, we have introduced the variance of the differentiated cells as $\sigma_d^2 = VR^A$; now, by approximating (41) to have such a particular form and setting $V = 1$, and using the values of (42), we can find A for each cone as

$$A_\delta \approx 0.440, \quad A_g \approx 1.247, \quad A_r \approx 1.256, \quad A_b \approx 1.322, \quad A_v \approx 1.308, \tag{43}$$

where they are in agreement with the assumption of hyperuniformity made previously, that is, $A < 2$.

As an illustration, based on the values of (42) and (43), we have demonstrated the biologically (and physically) acceptable regions of total entropy production (38) in the parametric space of the LEUP parameters: (β_s, β_d) and $(\beta_s, \beta_d, \Delta\tilde{\mu})$ in Fig. 7, in the case of the double cone photoreceptor.

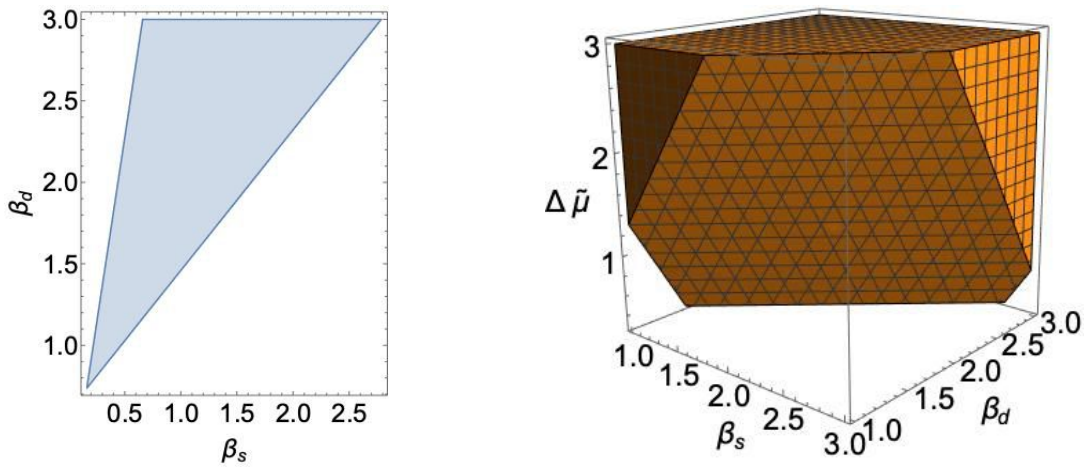

Figure 7. The biologically (and physically) acceptable total entropy production (shaded regions) for the double cone photoreceptor of the avian retina. In the left panel, we have fixed: $C_0 = -1$, $V = U = 1$, $D = 2$, and $\Delta\tilde{\mu} = 0.3$. This plot demonstrates the regions where (38) is positive, in the parametric space of the LEUP parameters (β_s, β_d). We note that we have also imposed the condition regarding the existence of the optimal sensing radius R_c, which in this case reads as $\Gamma < 0$, see also (39) and (40). In the right panel, we have relaxed the restriction on $\Delta\tilde{\mu}$.

At this point, we should remind that the parameter $\Delta\tilde{\mu}$ is proportional to the difference of differentiated and pluripotent cells division times. It is well known that the division time of any differentiated cell is much larger than the division time of the corresponding progenitor one [37]. Therefore, it is safe to assume that $\Delta\tilde{\mu} > 0$. This allows us to identify a relationship for the corresponding β's of the cones and their progenitors, where Fig. 7 implies that $\beta_d > \beta_s$. The latter is a sensible result since differentiated cells are expected to be more attentive to their microenvironment in order to ensure optimal cooperation and tissue integrity.

6. Discussion

The Least microEnvironmental Uncertainty Principle (LEUP) has been proposed recently to understand cell decision-making in multicellular systems and so far it has been applied to cell migration force distribution [14], collective cell migration [15], and phenotypic plasticity [16]. In this paper, we have used this formalism as a starting point and have developed a thermodynamic-like theory for a general cell and tissue differentiation process. We have applied the obtained results to the particular case of the avian photoreceptor mosaics formation.

Our LEUP-driven mathematical model is based on the two premises: (**i**) the multicellular system (tissue), where cell is differentiating, follows a Markovian dynamics with the assumption of microscopic reversibility and (**ii**) the spatial distribution of the early pluripotent cells microenvironment follows a Poisson distribution. Based on these assumptions, we have arrived at three important results: (**I**) predicting the color percentage of the cone cells in the avian retina without any knowledge about the underlying biophysical and biochemical mechanisms, (**II**) demonstrating the robustness of cell-tissue differentiation in thermodynamic terms, and (**III**) determining the limits of the cell sensing radius by establishing a relation between total entropy production and microenvironmental sensing. We elaborate on these results in the following.

Prediction of the cone color distribution: By calibrating a single parameter in the LEUP, we are able to predict the cone color percentage in the avian retina accurately. Our finding regarding the LEUP parameter which reads as $\beta \approx 1.754$ (close to 2), gives a strong indication that cells sense their environment quasi-optimally when choosing a particular cell fate during differentiation process. For future study and investigation, we want to examine the validity of this result for photoreceptor mosaics of other species [38–41].

Robustness and measurement of information gain in differentiation: We have constructed the fluctuation theorem for tissue differentiation and have derived a generalization of the second law of thermodynamics for this process based on a Markovian dynamics. It would be interesting and more realistic to relax this assumption and to analyze the problem using the LEUP on the basis of non-Markovian or memory processes [42], which can lead to different results. The only requirement is that system should have a unique stationary state. We also note that the LEUP-driven second law of thermodynamics can also be seen as a generalization of the Bayesian second law of thermodynamics [43] and the conditional second law of thermodynamics in a strongly coupled system [44].

Our theory, which in the current paper has been applied to the specific case of the avian photoreceptor mosaics, suggests that differentiated tissue is (highly) thermodynamically robust, that is, the arrow of time is almost irreversible, and this robustness depends on microenvironmental sensing and cell metabolism. It should be remarked that, as is demonstrated in (28), if we have the values of $\langle \beta' \Delta Q^{\tau}_{x_s \to x_d} \rangle_{s \to d}$ and $\langle \Delta S \rangle_{s \to d}$, then we can determine the upper bound of pointwise mutual information difference which is denoted as $\langle \Delta i \rangle_{s \to d}$ within the context of the LEUP. However, in the present work, we have only considered transitions between equilibrium end-states and have set $\langle \Delta i \rangle_{s \to d} = 0$. In a future study, we plan to investigate nonequilibrium dynamics of transitions between progenitor and differentiated cell states and establish the upper bound of information gain in cell differentiation. (We also note that the value of $\langle \Delta i \rangle_{s \to d}$ can also be obtained directly from experiments, see Refs. [45–48].)

Limits of sensing radius: By studying total entropy production as a function of cell sensing radius and the LEUP parameters, we have provided an understanding of how a cell regulates its sensing radius according to its microenvironment to ensure the thermodynamic robustness of differentiated tissue. We have shown two cases where (**a**) the entropy production goes to infinity beyond a certain threshold radius which is depicted in Fig. 5 and (**b**) the entropy production goes to a maximum value as in Fig. 6. Interestingly, we conclude that the former is the most biologically relevant case since it requires the division time of differentiated cells to be larger than that of the pluripotent ones and biologically systems operate away from fine-tuned parameter regimes to withstand noisy perturbations. On the technical side, we have assumed that the spatial distribution of pluripotent tissue resembles a

Poisson distribution. It would be interesting to relax this assumption and to derive this distribution from real tissue data.

In a nutshell, we have shown how the LEUP facilitates the inference of cellular intrinsic states (or, cell phenotypes) by means of local microenvironmental entropies or fluctuations. This allows the evaluation of cellular states without the detailed grasp of the underlying mechanisms. The sole knowledge about extrinsic variables distributions (or, collective cell decision-making) suffices. Therefore, we can apply the LEUP to cell differentiation problems where the biological or biophysical knowledge is unclear or unknown.

Funding: Ar. Ba., Al. Be. and H. H. would like to acknowledge the VolkswagenStiftung funding within the *Life?* program (96732). H. H. is supported by MiEDGE (01ZX1308D) of the ERACOSYSMED initiative. The authors also thank the Centre for Information Services and High Performance Computing (ZIH) at TU Dresden for providing an excellent infrastructure.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Simon, H.A. *The New Science of Management Decision*; Prentice Hall: USA, 1977.
2. Slack, J.M.W. Metaplasia and transdifferentiation: from pure biology to the clinic. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 369–378. doi:10.1038/nrm2146.
3. Alberts, B.; Johnson, A.D.; Lewis, J.; Morgan, D.; Raff, M.; Roberts, K.; Walter, P. *Molecular Biology of the Cell*, sixth ed.; W. W. Norton & Company: New York City, NY, 2014.
4. Goldberg, A.D.; Allis, C.D.; Bernstein, E. Epigenetics: A Landscape Takes Shape. *Cell* **2007**, *128*, 635–638. doi:10.1016/j.cell.2007.02.006.
5. Allen, M. Compelled by the Diagram: Thinking through C. H. Waddington's Epigenetic Landscape. *Contemporaneity: Historical Presence in Visual Culture* **2015**, *4*, 119–142. doi:10.5195/contemp.2015.143.
6. Casey, M.J.; Stumpf, P.S.; MacArthur, B.D. Theory of cell fate. *Wiley Interdiscip. Rev.: Syst. Biol. Med.* **2020**. doi:10.1002/wsbm.1471.
7. Ferrell, J.E. Bistability, Bifurcations, and Waddington's Epigenetic Landscape. *Curr. Biol.* **2012**, *22*, R458–R466. doi:10.1016/j.cub.2012.03.045.
8. Paździorek, P.R. Mathematical Model of Stem Cell Differentiation and Tissue Regeneration with Stochastic Noise. *Bull. Math. Biol.* **2014**, *76*, 1642–1669. doi:10.1007/s11538-014-9971-5.
9. Jilkine, A. Mathematical Models of Stem Cell Differentiation and Dedifferentiation. *Curr. Stem Cell Rep.* **2019**, *5*, 66–72. doi:10.1007/s40778-019-00156-z.
10. Berg, H.C.; Purcell, E.M. Physics of chemoreception. *Biophys. J.* **1977**, *20*, 193–219. doi:10.1016/S0006-3495(77)85544-6.
11. Bialek, W.; Setayeshgar, S. Physical limits to biochemical signaling. *PNAS* **2005**, *102*, 10040–10045. doi:10.1073/pnas.0504321102.
12. ten Wolde, P.R.; Becker, N.B.; Ouldridge, T.E.; Mugler, A. Fundamental Limits to Cellular Sensing. *J. Stat. Phys.* **2016**, *162*, 1395–1424. doi:10.1007/s10955-015-1440-5.
13. Mora, T.; Nemenman, I. Physical Limit to Concentration Sensing in a Changing Environment. *Phys. Rev. Lett.* **2019**, *123*, 198101. doi:10.1103/PhysRevLett.123.198101.
14. Hatzikirou, H. Statistical mechanics of cell decision-making: the cell migration force distribution. *J. Mech. Behav. Mater.* **2018**, *27*, 1–7. doi:10.1515/jmbm-2018-0001.
15. Barua, A.; Nava-Sedeño, J.M.; Meyer-Hermann, M.; Hatzikirou, H. A least microenvironmental uncertainty principle (LEUP) as a generative model of collective cell migration mechanisms. *Sci. Rep.* **2020**, *10*, 22371. doi:10.1038/s41598-020-79119-y.
16. Barua, A.; Syga, S.; Mascheroni, P.; Kavallaris, N.; Meyer-Hermann, M.; Deutsch, A.; Hatzikirou, H. Entropy-driven cell decision-making predicts 'fluid-to-solid' transition in multicellular systems. *New J. Phys.* **2020**, *22*, 123034. doi:10.1088/1367-2630/abcb2e.

17. Purvis, J.E.; Lahav, G. Encoding and Decoding Cellular Information through Signaling Dynamics. *Cell* **2013**, *152*, 945–956. doi:10.1016/j.cell.2013.02.005.
18. Bialek, W. *Biophysics: Searching for Principles*; Princeton University Press: Princeton, NJ, 2012.
19. Kram, Y.A.; Mantey, S.; Corbo, J.C. Avian Cone Photoreceptors Tile the Retina as Five Independent, Self-Organizing Mosaics. *PLoS ONE* **2010**, *5*, 1–13. doi:10.1371/journal.pone.0008992.
20. Seifert, U. Stochastic thermodynamics: An introduction. *AIP Conference Proceedings* **2011**, *1332*, 56. doi:10.1063/1.3569487.
21. Seifert, U. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.* **2012**, *75*, 126001. doi:10.1088/0034-4885/75/12/126001.
22. Van den Broeck, C. Stochastic thermodynamics: A brief introduction. *Proceedings of the International School of Physics "Enrico Fermi"* **2013**, *184*, 155–193. doi:10.3254/978-1-61499-278-3-155.
23. Klages, R.; Just, W.; Jarzynski, C., Eds. *Nonequilibrium Statistical Physics of Small Systems: Fluctuation Relations and Beyond*; Reviews of Nonlinear Dynamics and Complexity, Wiley-VCH: Weinheim, Germany, 2013.
24. Seifert, U. Stochastic thermodynamics: From principles to the cost of precision. *Physica A* **2018**, *504*, 176–191. doi:10.1016/j.physa.2017.10.024.
25. Ciliberto, S. Experiments in Stochastic Thermodynamics: Short History and Perspectives. *Phys. Rev. X* **2017**, *7*, 021051. doi:10.1103/PhysRevX.7.021051.
26. Ciliberto, S. From space-time chaos to stochastic thermodynamics. *CR Phys.* **2019**, *20*, 529–539. doi:10.1016/j.crhy.2019.09.001.
27. Van den Broeck, C.; Esposito, M. Ensemble and trajectory thermodynamics: A brief introduction. *Physica A* **2015**, *418*, 6–16. doi:10.1016/j.physa.2014.04.035.
28. Crooks, G.E. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E* **1999**, *60*, 2721–2726. doi:10.1103/PhysRevE.60.2721.
29. England, J.L. Statistical physics of self-replication. *J. Chem. Phys.* **2013**, *139*, 121923. doi:10.1063/1.4818538.
30. Evans, D.J.; Cohen, E.G.D.; Morriss, G.P. Probability of second law violations in shearing steady states. *Phys. Rev. Lett.* **1993**, *71*, 2401. doi:10.1103/PhysRevLett.71.2401.
31. Evans, D.J.; Searles, D.J. The Fluctuation Theorem. *Adv. Phys.* **2002**, *51*, 1529–1585. doi:10.1080/00018730210155133.
32. Sevick, E.M.; Prabhakar, R.; Williams, S.R.; Searles, D.J. Fluctuation Theorems. *Ann. Rev. Phys. Chem.* **2008**, *59*, 603–633. doi:10.1146/annurev.physchem.58.032806.104555.
33. Abe, S. Conditional maximum-entropy method for selecting prior distributions in Bayesian statistics. *EPL* **2014**, *108*, 40008. doi:10.1209/0295-5075/108/40008.
34. Pei, Y.; Biswas, S.; Fussell, D.S.; Pingali, K. An elementary introduction to Kalman filtering. *Commun. ACM* **2019**, *62*, 1–12. doi:10.1145/3363294.
35. Jiao, Y.; Lau, T.; Hatzikirou, H.; Meyer-Hermann, M.; Corbo, J.C.; Torquato, S. Avian photoreceptor patterns represent a disordered hyperuniform solution to a multiscale packing problem. *Phys. Rev. E* **2014**, *89*, 022721. doi:10.1103/PhysRevE.89.022721.
36. Marín, D.; Sabater, B. The cancer Warburg effect may be a testable example of the minimum entropy production rate principle. *Phys. Biol.* **2017**, *14*, 024001. doi:10.1088/1478-3975/aa64a7.
37. Soufi, A.; Dalton, S. Cycling through developmental decisions: how cell cycle dynamics control pluripotency, differentiation and reprogramming. *Development* **2016**, *143*, 4301–4311. doi:10.1242/dev.142075.
38. Curcio, C.A.; Sloan, K.R. Packing geometry of human cone photoreceptors: Variation with eccentricity and evidence for local anisotropy. *Visual Neurosci.* **1992**, *9*, 169–180. doi:10.1017/S0952523800009639.
39. Galli-Resta, L.; Novelli, E.; Kryger, Z.; Jacobs, G.H.; Reese, B.E. Modelling the mosaic organization of rod and cone photoreceptors with a minimal-spacing rule. *Eur. J. Neurosci.* **1999**, *11*, 1461–1469. doi:10.1046/j.1460-9568.1999.00555.x.
40. Mitkus, M.; Chaib, S.; Lind, O.; Kelber, A. Retinal ganglion cell topography and spatial resolution of two parrot species: budgerigar (Melopsittacus undulatus) and Bourke's parrot (Neopsephotus bourkii). *J. Comp. Physiol. A* **2014**, *200*, 371–384. doi:10.1007/s00359-014-0894-2.

41. Bueno, J.M.; Cruz-Castillo, R.; Avilés-Trigueros, M.; Bautista-Elivar, N. Arrangement of the photoreceptor mosaic in a diabetic rat model imaged with multiphoton microscopy. *Biomed. Opt. Express* **2020**, *11*, 4901–4914. doi:10.1364/BOE.399835.
42. Speck, T.; Seifert, U. The Jarzynski relation, fluctuation theorems, and stochastic thermodynamics for non-Markovian processes. *J. Stat. Mech.* **2007**, *2007*, L09002. doi:10.1088/1742-5468/2007/09/l09002.
43. Bartolotta, A.; Carroll, S.M.; Leichenauer, S.; Pollack, J. Bayesian second law of thermodynamics. *Phys. Rev. E* **2016**, *94*, 022102. doi:10.1103/PhysRevE.94.022102.
44. Crooks, G.E.; Still, S. Marginal and conditional second laws of thermodynamics. *EPL* **2019**, *125*, 40005. doi:10.1209/0295-5075/125/40005.
45. Fuller, D.; Chen, W.; Adler, M.; Groisman, A.; Levine, H.; Rappel, W.J.; Loomis, W.F. External and internal constraints on eukaryotic chemotaxis. *PNAS* **2010**, *107*, 9656–9659. doi:10.1073/pnas.0911178107.
46. Hu, B.; Chen, W.; Rappel, W.J.; Levine, H. Physical Limits on Cellular Sensing of Spatial Gradients. *Phys. Rev. Lett.* **2010**, *105*, 048104. doi:10.1103/PhysRevLett.105.048104.
47. Hu, B.; Fuller, D.; Loomis, W.F.; Levine, H.; Rappel, W.J. Phenomenological approach to eukaryotic chemotactic efficiency. *Phys. Rev. E* **2010**, *81*, 031906. doi:10.1103/PhysRevE.81.031906.
48. Uda, S. Application of information theory in systems biology. *Biophys. Rev.* **2020**, *12*, 377–384. doi:10.1007/s12551-020-00665-w.

© 2021 by the authors. Submitted to *Entropy* for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).