Article

Subjective Information and Survival in a Simulated Biological System

Tyler S. Barker [1,†], Massimiliano Pierobon [1,*,†] and Peter J. Thomas [2,†]

1 School of Computing, College of Engineering, University of Nebraska-Lincoln, Lincoln, NE 68588, USA;
   tyler.barker@huskers.unl.edu
2 Department of Mathematics, Applied Mathematics and Statistics, Case Western Reserve University,
   Cleveland, OH 44106, USA; pjthomas@case.edu
* Correspondence: maxp@unl.edu
† These authors contributed equally to this work.

Abstract: Information transmission and storage have gained traction as unifying concepts to characterize biological systems and their chances of survival and evolution at multiple scales. Despite the potential for an information-based mathematical framework to offer new insights into life processes and ways to interact with and control them, the main legacy is that of Shannon's, where a purely syntactic characterization of information scores systems on the basis of their maximum information efficiency. The latter metrics seem not entirely suitable for biological systems, where transmission and storage of different pieces of information (carrying different semantics) can result in different chances of survival. Based on an abstract mathematical model able to capture the parameters and behaviors of a population of single-celled organisms whose survival is correlated to information retrieval from the environment, this paper explores the aforementioned disconnect between classical information theory and biology. In this paper, we present a model, specified as a computational state machine, which is then utilized in a simulation framework constructed specifically to reveal emergence of a "subjective information", i.e., trade-off between a living system's capability to maximize the acquisition of information from the environment, and the maximization of its growth and survival over time. Simulations clearly show that a strategy that maximizes information efficiency results in a lower growth rate with respect to the strategy that gains less information but contains a higher meaning for survival.

1. Introduction

Information processing and its correlation with survival and evolution in living organisms have been proposed through the years as unifying concepts to study life across different scales [1]. The transmission and storage of information are indeed at the basis of very diverse life processes, ranging from the processing of information via molecular interactions within a cell to complex codes in inter-organismal communication, and the consequent emergence of long-term evolutionary patterns [2]. Interestingly, while the development of a suitable mathematical framework to study and quantify information in these contexts could revolutionize the way we characterize, analyze, and interact with biological systems, there is little agreement on the mathematical formalism to be applied [3], mostly based upon Shannon's proposed framework.

In its inception, the restriction of Shannon's mathematical theory of communication [4,5] to purely syntactic considerations, where the meaning of information is considered irrelevant to the engineering problem, was a conceptual *sine qua non* for its success and applicability within the emerging fields of electronics and telecommunications. In particular, the concept of information entropy as a universal metric to quantify information, or its absence, i.e., noise, and the definition of the capacity of a communication channel as the theoretical

maximum amount of information, or Mutual Information (MI) that can be unequivocally propagated between two points in space (transmission) or time (storage), have been the fundamental underpinnings in the study of mathematical algorithms where information can be represented and received with maximum efficiency.

Despite attempts to apply the aforementioned concepts in biology, from neuroscience to biochemistry [6], and data analytics for bioinformatics [7], even abstracting biological systems as communication channels [8–11], the syntactic nature of information theory provides an obstacle to its application to living systems. Intuitively, within biological systems, some messages are "more important" than others. There have been attempts to provide a quantitative basis for this idea [12–14] but a general framework remains lacking.

Here, we develop a simple but rigorous model that illustrates the notion of "**subjective information**", defined as the particular information that an individual organism focuses on, out of all the potentially *useful* information available to the organism. We show that, under certain conditions, maximizing the expected cell division rate (or fitness) requires maximizing information about a sequence of selected subsets of an input signal's components in a way that deliberately discards a larger quantity of less useful information in favor of a smaller quantity of more useful information. The work presented in the following builds on preliminary results included in a previous conference publication [15]. In this extended version, we formulate a mathematical model of the biological system alongside the computational model, and we include the modeling and simulation of death and division processes, and introduce a new metric to measure cell population growth to better capture the correlation between information and organism survival. Comprehensive results presented in Section 5 now enable us to study the effects of cell stress on survival and its correlation to subjective information, and to propose a simple metric to quantify the emergence of subjective information in a population.

Other models and frameworks to study information in the context of agents receiving information to maximize their fitness have been proposed, either in the context of biology or in more general terms. In particular, a trade-off between a reward or fitness purpose (value-to-go) experienced by an organism versus how much informational "bandwidth" can be afforded (information-to-go) for that purpose is analyzed in [14] with a Markov Decision Process (MDP) formalism. These authors find a trade-off, under the constraint of a cost for information acquisition, by applying a reinforcement learning optimization to an MDP model of the organism. In [16], a thermodynamic-based framework (instead of purely information-theoretic) is employed to study decision-making processes where a constraint on the information processing resources needed to achieve optimal decision for maximum reward is considered. Under this framework, a trade-off between reward and information processing cost is found as a function of this constraint through the theory of free-energy differences and thermodynamic energy potentials. When processing costs are ignored, the maximum utility is reached with maximum use of information. In [17], a clear distinction about relevant and irrelevant information for an organism is made by proposing a metric to account for the fitness value of information, or *Gould–Kelly* information [18], i.e., the increase in the organism's fitness resulting from a cue conveying (limited) information about the environment. In addition, an upper bound limit to the fitness value is found as the mutual information between this cue and the environment, which can be easily interpreted as the performance of the communication channel between environment and organism. In a more recent publication [19], the fitness value of information and its upper bound are generalized for situations where the environment can vary fitness related parameters in time and/or space, possibly characterized by gradients, where the value of an informative cue to the organism becomes a function of the probabilistic distribution of these parameters. A definition of "semantic information" is proposed in [20] as the minimum amount of (relevant) information for an organism to ensure maximum viability (fitness). This quantity is found by gradually degrading a single information channel between the environment and the organism through a "coarse-graining" function until a decrease in the organism

fitness is observed. It is here assumed that providing the organism with more information than the semantic information does not contribute to its fitness.

All these previous contributions, amongst other results, seem to conclude that, for a biological organism (or a model/generalization thereof), more "relevant" information is always better, with its cost of acquisition or processing being the constraint, i.e., there is always a trade-off between utility/reward and work to be done. In contrast, our contribution shifts the focus from constraints on information cost, to constraints in the capability of an individual organism to opportunistically manipulate the information channel between the environment and itself. As we demonstrate through our model, where we consider two essential metabolic sources for growth and survival corresponding to competing space-varying chemotactic information cues from the environment, the variability of this information channel across the organism population has the two-fold effects of increasing the overall fitness while also decreasing the average information necessary to achieve this fitness.

The rest of the paper is organized as follows: in Section 2, we introduce a mathematical abstraction and consequent mathematical and computational models of a living system exhibiting the aforementioned characteristics, based on an organism that requires two essential molecular species, and information on their distribution, to survive and grow. In Section 3, we define two metrics to quantify information and survival from the output of simulations based on the model organism. In Section 4, we introduce and discuss how the simulation framework is implemented from the computational model. In Section 5, we present numerical results obtained from our simulations and the subsequent estimation of the metrics, while, in Section 6, we discuss our findings, finally proposing a simple metric to quantify the emergence of subjective information in the simulated living system. Finally, we conclude the paper in Section 7.

2. Two-Resource Foraging Model

In this paper, we consider an abstract model of an organism that requires two essential metabolic substrates to survive and grow, and whose behavior is simple but essential for appreciating the concept of "subjective information" and its correlation with survival. In this model, the two substrates have different spatial distributions, and the organism detects their local concentrations and gradients through a noisy receptor-binding process [21,22], which in turn informs its chemotaxis [23]. This model is then used to compare the organism survival (in terms of population growth rates [18,24]) upon adoption of strategies based either on maximizing information on the two essential substrates [25] or, alternatively, reducing this information by focusing on what is more important for survival [12], which corresponds to the emergence of a "subjective information".

2.1. Mathematical Model

Our organism is a motile species of cells of length ℓ, inhabiting a one-dimensional circular (periodic) environment of length $L \gg \ell$, for a time $t \in [0, T]$ as illustrated in Figure 1. In particular, our mathematical model is defined according to the following assumptions:

- The cell is rod-like (an abstraction of many motile bacteria), for which we distinguish "right" and "left" ends relative to its position axis.
- The metabolic substrates, denoted A and B, are present in the environment at determinate concentrations for each environment location $x \in [0, L]$. We impose a nonuniform distribution of metabolites A and B, given by local concentration functions $A(x)$ and $B(x)$. For simplicity, we consider these concentrations to be static within the timescale of microbial population growth T.
- Each cell at environment location $x(t)$ at time t maintains an internal storage of both substrate A and B molecules, i.e., $A_{\text{in}}(t)$ and $B_{\text{in}}(t)$, by **absorbing** A and B molecules from the environment proportionally to the concentrations $A(x(t))$ and $B(x(t))$, respectively, according to a determinate constant absorption coefficient k.
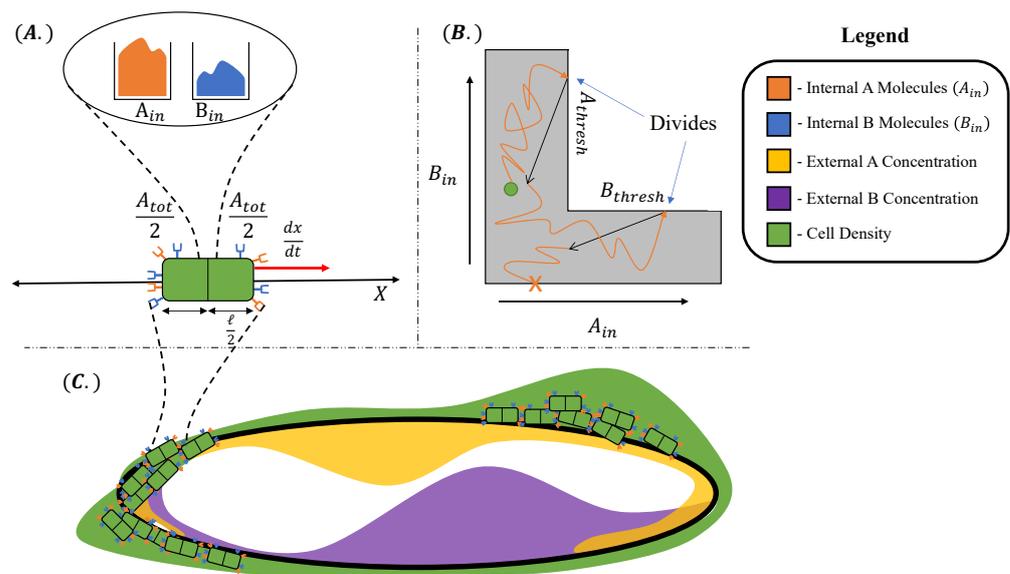
Figure 1. Model components and topology. (**A**) Each individual cell has an internal storage of metabolic resources A_{in} and B_{in}, and receptors $A_{\text{tot}}/2$ and $B_{\text{tot}}/2$ on the right and $A_{\text{tot}}/2$ and $B_{\text{tot}}/2$ on the left end of the cell. The cell has length ℓ and moves through space with velocity dx/dt driven by differences in bound receptors. (**B**) Each cell's internal resources exist in an L-shaped domain given by $\{0 \leq A_{\text{in}} \leq A_{\text{thresh}}, 0 \leq B_{\text{in}}\} \cup \{0 \leq A_{\text{in}}, 0 \leq B_{\text{in}} \leq B_{\text{thresh}}\}$. Starting from an initial resource profile (green dot), the cell accumulates metabolic resources (orange trajectory) until it reaches a cell division boundary $\mathcal{S}_{\text{div}}$ (see text for details). Upon reaching $\mathcal{S}_{\text{div}}$, the cell divides into two, partitioning its metabolic resources (black arrow). On the contrary, a cell dies if its internal storage of either A or B decreases to zero (orange "X"). (**C**) The geometry of the environment can be thought of as a 1D ring, $0 \leq x < L$. The environmental concentrations of metabolites $A(x)$, $B(x)$ vary with position around the ring.

- The cell receives information about its environment by **sensing**, which is realized through the binding of distinct chemical receptor proteins to A and B molecules.
- We endow each cell with a fixed budget of R_{tot} receptor proteins in total, equally distributed between the right and the left sides. The cell has $A_{\text{tot}}(t)$ receptors for A molecules and $B_{\text{tot}}(t)$ receptors for B molecules at time t, respectively, where $A_{\text{tot}}(t) + B_{\text{tot}}(t) \equiv R_{\text{tot}}$ and R_{tot} is constant over time.
- The cell reacts to its surroundings by **moving** along the direction of and proportionally to an estimation of the gradients of the concentrations $A(x(t))$ and $B(x(t))$ from the numbers of bound receptors for A and B molecules on the right and left sides, denoted as A_{R}^*, B_{R}^*, A_{L}^*, and B_{L}^*, respectively.
- The cell can control the amount of received information about the concentrations $A(x(t))$ and $B(x(t))$ and their gradients by way of **(re-)allocating** the R_{tot} receptors between $A_{\text{tot}}(t)$ receptors for A molecules and $B_{\text{tot}}(t)$ receptors for B molecules. In this paper, we contrast two different strategies, namely, a constant *equal receptor allocation* and an *adaptive receptor allocation*, the latter with the goal of acquiring more information about the more scarce substrate in its internal storage. We make the (strong) assumption that these cells can rapidly convert receptors between A-specific and B-specific types, without incurring a substantive metabolic cost for the transition.
- The cell consumes the substrates from its internal storage with a rate corresponding to a basal maintenance rate of metabolism S [26,27]. To grow and divide, a cell must maintain positive internal storage $A_{\text{in}}(t)$ and $B_{\text{in}}(t)$. The occurrence of cell division and cell death events are consequently decided by **assessing** the values of the current internal storage $A_{\text{in}}(t)$ and $B_{\text{in}}(t)$: when both the amounts of $A_{\text{in}}(t)$ and $B_{\text{in}}(t)$ exceed a specified threshold, the cell divides into two daughter cells, each receiving half of

the mother cell's internal storage. If either of the amounts of $A_{\text{in}}(t)$ and $B_{\text{in}}(t)$ in a particular cell becomes zero, the cell dies.

According to these assumptions, we express the behavior of a cell according to the following mathematical model.

Absorbing. The internal concentrations of A and B, $A_{\text{in}}(t)$ (respectively $B_{\text{in}}(t)$), of a cell located at position $x(t)$ evolves according to

$$\frac{dA_{\text{in}}(t)}{dt} = kA(x(t)) - S, \tag{1}$$

where the absorption coefficient k [28] translates from concentration to rate of molecules absorbed, and quantifies how easily a cell can absorb molecules. $B_{\text{in}}(t)$ is computed by substituting B in place of A. The parameter S represents a constant metabolic maintenance cost [27], assumed to be the same for both A and B. Table 1 lists the simulation parameters used.

Sensing. We model the state of the receptors at time t as a sample taken at equilibrium from a binomial binding/release process [22,29]. Thus, the receptor noise is binomial, meaning the numbers of bound receptors satisfy

$$
\begin{aligned}
A_{\text{R}}^* &\sim \text{Binom}\left(p_{\text{bind},A_{\text{R}}}, \frac{A_{\text{tot}}}{2}\right), & p_{\text{bind},A_{\text{R}}} &= \frac{\left(A(x) + \frac{\ell}{2} \cdot \nabla A(x)\right)}{K_d + \left(A(x) + \frac{\ell}{2} \cdot \nabla A(x)\right)}, \\
A_{\text{L}}^* &\sim \text{Binom}\left(p_{\text{bind},A_{\text{L}}}, \frac{A_{\text{tot}}}{2}\right), & p_{\text{bind},A_{\text{L}}} &= \frac{\left(A(x) - \frac{\ell}{2} \cdot \nabla A(x)\right)}{K_d + \left(A(x) - \frac{\ell}{2} \cdot \nabla A(x)\right)}, \\
B_{\text{R}}^* &\sim \text{Binom}\left(p_{\text{bind},B_{\text{R}}}, \frac{B_{\text{tot}}}{2}\right), & p_{\text{bind},B_{\text{R}}} &= \frac{\left(B(x) + \frac{\ell}{2} \cdot \nabla B(x)\right)}{K_d + \left(B(x) + \frac{\ell}{2} \cdot \nabla B(x)\right)}, \\
B_{\text{L}}^* &\sim \text{Binom}\left(p_{\text{bind},B_{\text{L}}}, \frac{B_{\text{tot}}}{2}\right), & p_{\text{bind},B_{\text{L}}} &= \frac{\left(B(x) - \frac{\ell}{2} \cdot \nabla B(x)\right)}{K_d + \left(B(x) - \frac{\ell}{2} \cdot \nabla B(x)\right)},
\end{aligned}
\tag{2}
$$

where $p_{\text{bind},A_{\text{R}}}$ is the probability of binding A molecules for a single receptor at the right side of the cell, $A(x) \pm \frac{\ell}{2} \cdot \nabla A(x)$ is the the local concentration of A (resp. B) at the right and left end of the cell, ∇ is the gradient operator (here, $\frac{d}{dx}$), and K_d is the chemical dissociation constant of the receptor [30]. For simplicity, we omitted the time variable t, we presume that samples taken at successive time steps and at opposite sides of the cell are independent, and we assume the same dissociation constant for each receptor.

Moving. We set the cell velocity to be proportional to the difference between the number of receptors bound on the right versus left sides of the cell:

$$\frac{dx}{dt} = v\left(\frac{2v_{\text{max}}}{1 + e^{-\left(A_{\text{R}}^* + B_{\text{R}}^* - A_{\text{L}}^* - B_{\text{L}}^*\right)}} - v_{\text{max}}\right), \tag{3}$$

where v is a constant parameter that converts a number of bound receptors to velocity, and v_{max} corresponds to the maximum velocity.

(Re-)allocating. The *equal receptor allocation* strategy seeks the greatest possible information about the environmental concentrations $A(x)$ and $B(x)$, regardless of their importance for cell survival, by setting a static distribution $A_{\text{tot}} = B_{\text{tot}}$ for all t, regardless of the cell's internal storage of A and B molecules.

The *adaptive receptor allocation* strategy, in contrast, attempts to increase the probability of survival for the cell by distributing receptors among the two types in a way that takes into account $B_{in}(t)$ and $A_{in}(t)$. The receptor apportionment rule is given as follows:

$$B_{tot}(t) = R_{tot}\left(\frac{A_{in}(t)}{A_{in}(t) + B_{in}(t)}\right),$$

and $A_{tot}(t) = R_{tot} - B_{tot}(t)$. In this way, the cell redistributes its receptors in proportion to its relative deficit of one metabolite versus the other. For example, if the cell has fewer B than A molecules, it will move a greater portion of its receptors to receptor type B. It is this apportionment rule that implements the "subjectivity" of information reception in this model, which is reflected in the simulation results in Section 5.

Assessing. A cell divides if its internal metabolites $A_{in}(t)$ and $B_{in}(t)$ both surpass a division threshold $A_{thresh} = B_{thresh} = D$. We define the cell division time τ_{div} to be the first passage time of the cell's $(A_{in}(t), B_{in}(t))$ trajectory to the cell division "surface" $\mathcal{S}_{div} \equiv \mathcal{S}_A \cup \mathcal{S}_B$, where

$$\begin{aligned} \mathcal{S}_A &= \{(a,b) \mid a = A_{thresh}, \, b \geq B_{thresh}\} \\ \mathcal{S}_B &= \{(a,b) \mid a \geq A_{thresh}, \, b = B_{thresh}\}. \end{aligned} \tag{4}$$

Figure 1B illustrates the division surface. Thus, as $t \to \tau_{div}$, the cell's internal state approaches $(A_{in}^*, B_{in}^*) \in \mathcal{S}_{div}$. At time τ_{div}, the old cell is replaced with two daughter cells at the same current location, each with internal state

$$(A_{in}(\tau_{div}^+), B_{in}(\tau_{div}^+)) = \left(\frac{A_{in}^*}{2}, \frac{B_{in}^*}{2}\right), \tag{5}$$

and increment the total population count by one. The two daughter cells then move independently of one another.

A cell dies if its internal storage $(A_{in}(t), B_{in}(t))$ reaches the death threshold $\mathcal{S}_{death} \equiv \{(a,b) \mid a = 0 \text{ or } b = 0\}$. In this case, the old cell ceases to exist, and the total population count is decremented by one.

2.2. Computational Model

To appreciate the emergence of "subjective information" and its correlation with organism survival in the mathematical abstraction detailed in Section 2.1, we derive here a computational model that is later utilized to obtain numerical simulation results.

In the rest of this section, we consider a sampled environment $[0, L]$ so that each discrete location in the environment $\bar{x} \in \{iL/N \mid i = 0, \ldots, N-1\}$. Here, we impose periodic boundary conditions, where $\bar{x} = 0$ ($i = 0$) and $\bar{x} = L$ ($i = N$) represent the same location. Likewise, we sample the time variable into $\bar{t} \in \{j\Delta t \mid j = 0, \ldots, T/(\Delta t)\}$.

In Figure 2, we show a state machine diagram of our computational model, which is based on the definition and the cell behavior abstractions presented in Section 2.1. In particular, the discrete states of the cell and the transitions that may occur within a time sample Δt at time $\bar{t}$ include the following: **absorption** of A and B molecules, resulting into the changes $\Delta A_{in}(\bar{t})$ and $\Delta B_{in}(\bar{t})$ in the internal storage of A and B molecules, respectively; **sensing** through the chemical receptors, which results into the numbers of bound receptors for A and B molecules, i.e., A_R^*, B_R^*, A_L^*, and B_L^*, respectively, according to (2); **movement** by a discrete length Δi based on the estimated concentration gradients, and received information modulation via **receptor (re-)allocation**, which follows one of the two aforementioned strategies to allocate $A_{tot}(\bar{t} + \Delta t)$ and $B_{tot}(\bar{t} + \Delta t)$. Finally, the cell divides if it **assesses** a surplus of molecule A and B that both exceed the threshold D, or dies if its internal storage $(A_{in}(\bar{t}), B_{in}(\bar{t}))$ reaches the death threshold $\mathcal{S}_{death}$. Each of the aforementioned states reproduces the cell behaviors detailed in Section 2.1 as follows.
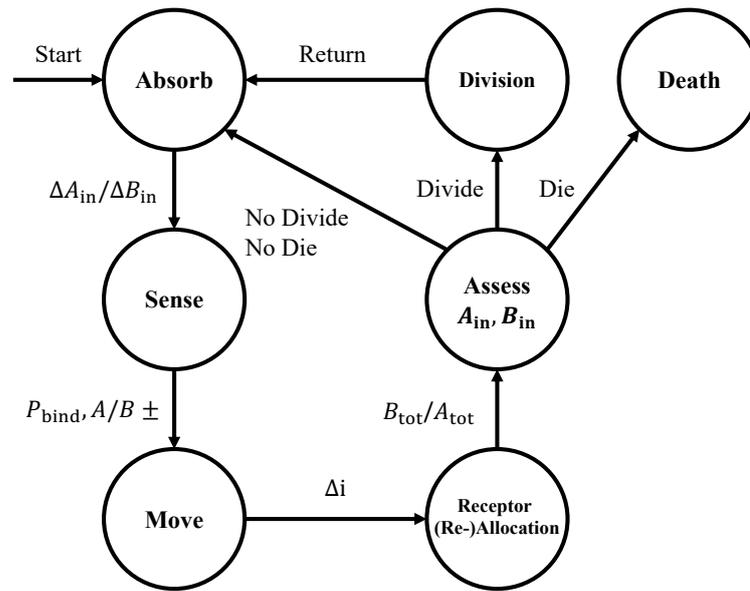
Figure 2. State machine diagram of the cell computational model formulated in this paper. See text for details.

Absorb. The cell absorbs A and B molecules according to the average of the external concentrations $A(x)$ and $B(x)$ at the left and right of the cell, $\bar{x}(\bar{t}) - \frac{\ell}{2}$ and $\bar{x}(\bar{t}) + \frac{\ell}{2}$, respectively. To compute the absorbed A (or B) molecules $\Delta A_{in}(\bar{t})$ (or $\Delta B_{in}(\bar{t})$), Equation (1) is changed into

$$\Delta A_{in}(\bar{t}) = \left\lfloor \left(k \frac{A(\bar{x}(\bar{t}) - \frac{\ell}{2}) + A(\bar{x}(\bar{t}) + \frac{\ell}{2})}{2} - S \right) \Delta t \right\rfloor, \tag{6}$$

where ΔB_{in} is computed by substituting B in place of A. ΔA_{in} and ΔB_{in} are then added to the internal storage of A and B molecules to obtain $A_{in}(\bar{t})$ and $B_{in}(\bar{t})$, respectively. If the resulting $A_{in}(\bar{t})$ and $B_{in}(\bar{t})$ are negative, then they are set to zero independently and cell death is assessed later on.

Sense. The state of the receptors at time $\bar{t}$ is a result of the same Binomial random processes as in (2), where this time we define the probability of binding A molecules at the right side of the cell as follows:

$$p_{bind,A_R} = \frac{A(\bar{x} + \frac{\ell}{2})}{K_d + A(\bar{x} + \frac{\ell}{2})}, \tag{7}$$

Similarly, we define the probability p_{bind,A_L} of binding A molecules for a single receptor at the left end of the cell (at location $\bar{x} - \frac{\ell}{2}$), with analogous definitions for the probabilities p_{bind,B_R} and p_{bind,B_L} for the B receptors. We assume in (7) that the time step Δt of the simulation is long enough that the receptors and surrounding concentration reach a steady state.

Move. We express the cell movement at time $\bar{t}$ according to an estimate of the local gradient of the concentration of molecules of type A and B, similarly as in (3):

$$\psi = A_R^* + B_R^* - A_L^* - B_L^*$$

$$\Delta i = \begin{cases} -v_{max} & \psi \leq -v_{max} \\ \psi & -v_{max} \leq \psi \leq v_{max} \\ v_{max} & \psi \geq v_{max} \end{cases} \tag{8}$$

where Δi is the change in cell location, corresponding to $\Delta \bar{x} = \Delta i \, L/N$, and $v_{\max}$ is the maximum allowed movement in a single step.

Receptor (Re-)Allocation. In the *equal receptor allocation* strategy, the cell behaves as detailed in Section 2.1. In the *adaptive receptor allocation* strategy, at time $\bar{t}$, the cell redistributes the receptors that will be considered in the absorption and sensing at time $\bar{t} + \Delta t$ among the two types based on the ratio of A to B molecules internal to the cell as follows:

$$B_{\text{tot}}(\bar{t} + \Delta t) = 2 \left\lfloor \frac{R_{\text{tot}}}{2} \frac{A_{\text{in}}(\bar{t})}{A_{\text{in}}(\bar{t}) + B_{\text{in}}(\bar{t})} \right\rfloor , \qquad (9)$$

where $B_{\text{tot}}(\bar{t} + \Delta t)$ is the new B-receptor count, $A_{\text{in}}(\bar{t})$, $B_{\text{in}}(\bar{t})$ are the internal A, B molecule storage after the cell absorption, and $A_{\text{tot}}(\bar{t} + \Delta t) = R_{\text{tot}} - B_{\text{tot}}(\bar{t} + \Delta t)$

Assess. The cell moves to the Divide state if the internal molecule numbers $A_{\text{in}}(\bar{t})$ and $B_{\text{in}}(\bar{t})$ both exceed the division threshold D or, equivalently, if the trajectory of $(A_{\text{in}}(\bar{t}), B_{\text{in}}(\bar{t}))$ has crossed the division surface $\mathcal{S}_{\text{div}}$, as defined in Section 2.1 and (4). The threshold D is set to the minimum energetic requirement needed in both A and B molecule storage for the cell to survive five time steps. If either of the internal molecule numbers $A_{\text{in}}(\bar{t})$ and $B_{\text{in}}(\bar{t})$ is equal to zero, the cell moves to the Death state. If a cell does not divide or die, then its state moves back to Absorb, and the process is repeated.

Divide. In this state, the cell loses half of A_{in} and B_{in} and an identical daughter cell is created at the same position $\bar{x}$ with an equal amount of $A_{\text{in}}(\bar{t} + \Delta t)$ and $B_{\text{in}}(\bar{t} + \Delta t)$ in the subsequent time step. The cell subsequently transitions to the Absorb state.

Death. In this state, the cell is considered dead and will be removed from the environment.

Figure 3 shows an example of the trajectory of $(A_{\text{in}}(\bar{t}), B_{\text{in}}(\bar{t}))$ in a simulation of the computational model. In this example, each time step is represented by an orange arrow and cell divisions by black arrows. The cell starts to gather B_{in}, moves through the divide threshold, and proceeds to divide twice in two time steps and ends with a higher A_{in} count below the division threshold. (This example shows how in a discrete computational model the cell could exist beyond the division threshold for multiple time steps.)
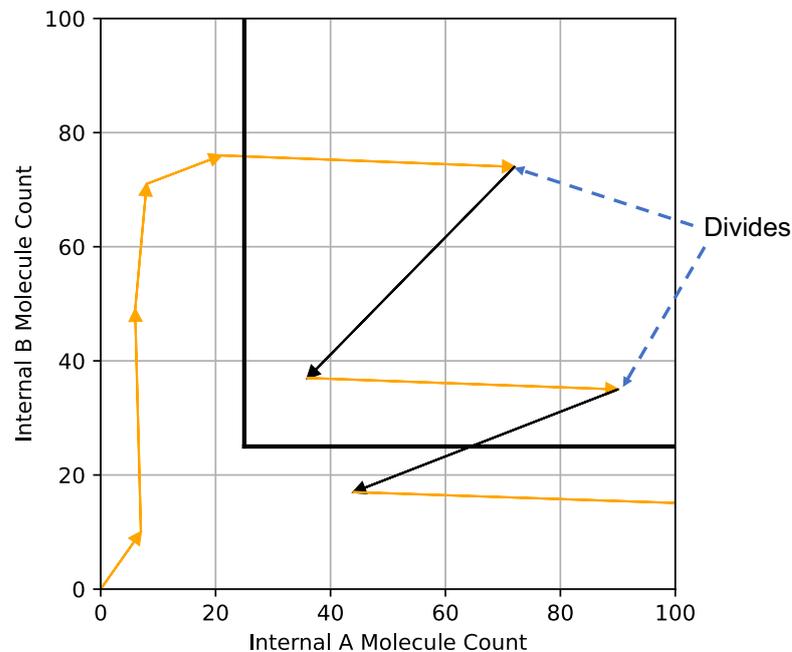

Figure 3. Trajectory of a simulated cell projected onto the A_{in} and B_{in} plane. The black line defines the threshold at which the cell can divide, and upon which it loses half its A_{in} and B_{in} to its daughter cell. The black arrows define cell division and the cell's internal state when division takes place. In this example, after one cell division (upper black arrow), the cell has enough A_{in} and B_{in} to divide again on the next timestep (lower black arrow).

The numerical results presented in Section 5 are obtained by simulating this computational model according to the parameter values listed in Table 1.

3. Performance Metrics

We compare the two receptor allocation strategies using two metrics: one based on the amount of information cells acquire from the environment, and the other based on the cell growth. The former is formulated as the average mutual information [5] between the input environmental concentration of A and B molecules and the output number of bound receptors A_R^*, B_R^*, A_L^*, and B_R^*, which are then used by the cell to estimate their gradients. The latter is expressed as the cell growth rate, which quantifies the exponential growth rate of the cells by utilizing the resources in the environment.

Information Efficiency. The Mutual Information (MI) formula is expressed as follows [5]:

$$MI_{X,Y} = \sum_{y \in Y} \sum_{x \in X} P_{(X,Y)}(x,y) \log_2 \frac{P_{(X,Y)}(x,y)}{P_X(x)P_Y(y)}, \tag{10}$$

where X is the set of input values, Y is the set of output values, $P_X(x)$ and $P_Y(y)$ are the marginal probabilities of X and Y, and $P_{(X,Y)}(x,y)$ is the joint probability of X and Y, respectively. The exact MI calculation in the simulation given a set of input and output data is defined in Algorithms 1–3.

Algorithm 1: Entropy (X)

Result: Entropy of a data set X
input :Set of data, X;
output: *entropyX*
initialize $(\_, bins) \leftarrow$ Binning(X);
$entropyX \leftarrow 0$;
for *each index i in bins* **do**
  | $entropyX \leftarrow entropyX - bins[i] \log_2(bins[i])$
end

Algorithm 2: $MI(X,Y)$

Result: Estimated Mutual Information Calculation
input : set of data X set of corresponding data Y
output: MI
$(boundsY, binnedY) \leftarrow$ Binning(Y);
$YallData \leftarrow$ two-dimensional array of size k initialized to empty arrays;
for *each index i in X/Y* **do**
  | **for** *each index j in boundsY* **do**
  | | **if** *boundsY[j + 1] > Y[i]* **then**
  | | | $YallData[j] = YallData[j]$ append $X[i]$
  | | **end**
  | **end**
end
$entropyXgivenY \leftarrow 0$;
for *each index i in YallData* **do**
  | $entropyXgivenY \leftarrow entropyXgivenY +$ Entropy$(YallData[i]) * binnedY[i]$
end
$MI \leftarrow$ Entropy$(X) - entropyXgivenY$;

Algorithm 3: Binning (*Y*)

Result: Array of bins containing the probability of selecting that bin
input : Y set of data points *Y*
output: bin bounds *bounds*, binned data points *bins*
min ← *Y*(0);
max ← *Y*(0);
ε ← 0.001;
k ← 20;
for *each receptor count i in Y* **do**
    if *Y*(*i*) < *min* **then**
        | *min* ← *Y*(*i*);
    else if *Y*(*i*) > *max* **then**
        | *max* ← *Y*(*i*);
end
min ← *min* - *ε*;
max ← *max* + *ε*;
bins ← zero indexed array the size of *k* with each index initialized to 0;
bounds ← zero indexed array the size of *k* + 1;
len ← length(*Y*);
for *each index i in bounds* **do**
    | $bounds(i) \leftarrow \frac{(min-max)}{k}i + min$;
end
for *each index i in Y* **do**
    for *each index j in bounds* **do**
        if *bounds*(*j* + 1) > *Y*(*i*) **then**
            | $bins(j) = bins(j) + \frac{1}{len}$;
        end
    end
end

The average MI at time *t* (*MI*(*t*)) of the cells in the environment is defined as

$$MI(t) = \mathrm{E}_{\text{Total \# of cells},t}[MI_{cell}], \tag{11}$$

where $\mathrm{E}_{\text{Total \# of cells},t}[\cdot]$ denotes the average computed over the entire population of cells present in the environment and over time up to time *t*, and MI_{cell} is the MI of a cell, computed as follows:

$$MI_{cell} = MI_{A_{\mathrm{R}},A_{\mathrm{R}}^*} + MI_{B_{\mathrm{R}},B_{\mathrm{R}}^*} + MI_{A_{\mathrm{L}},A_{\mathrm{L}}^*} + MI_{B_{\mathrm{L}},B_{\mathrm{L}}^*}, \tag{12}$$

where $A_{\mathrm{R}} = A(\bar{x} + \frac{\ell}{2})$, $B_{\mathrm{R}} = B(\bar{x} + \frac{\ell}{2})$, $A_{\mathrm{L}} = A(\bar{x} - \frac{\ell}{2})$, and $B_{\mathrm{L}} = B(\bar{x} - \frac{\ell}{2})$ are the environmental concentrations of molecules A or B on the right and left of a cell, respectively. The MI of the individual channels between the external concentration and bound receptors are assumed to be uncorrelated and are therefore added together. To express (12), we assume that each of the environmental concentrations and consequent number of bound receptors are independent from each other [30].

Growth Rate. We introduce death into the model and simulation to make the growth of the cells more realistic representations of in vivo environments. The growth rate over a time *T* is defined here in the same way as growth is defined in other systems, namely as the exponential rate of growth, that is, the doubling rate (or multiplication rate) over time,

$$G_T = \frac{\Delta t}{T} \sum_{j=0}^{(T/\Delta t)-1} \log_2\left(\frac{P_{j+1}}{P_j}\right), \tag{13}$$

where P_j is the population size at the end of the discrete time step with index j, as defined in Section 2.2. To obtain the growth rate results presented in Section 5, we discarded an initial transient and evaluated P_{j+1}/P_j by sampling from an ensemble at equilibrium.

4. Simulation Implementation

We simulate the model from Section 2.2 with many cells and their interaction with the environment using a computer program. The simulation results presented are obtained for different initial conditions and simulation parameters. The simulation is divided into time steps, where each step j is a cycle of the state machine. The main simulation parameters with their description and initialization values are shown in Table 1. Additional choices made for this simulation are as follows.

Table 1. Simulation parameters.

Variable	Description	Initialization
$A_{\text{tot}}, B_{\text{tot}}$	Total A/B receptor count	200 receptors
$A(\bar{x}), B(\bar{x})$	A/B external concentrations at location $\bar{x}$	0
$A_{\text{R}}^*, A_{\text{L}}^*, B_{\text{R}}^*, B_{\text{L}}^*$	A/B left/right bound receptor count	0
R_{tot}	Total number of receptors for a cell to allocate	400 receptors
$A_{\text{in}}, B_{\text{in}}$	Internal A and B molecule count	0
i	Cell Location	[51, 1–100]
k	Absorption Coefficient	[1.0–5.0]
K_d	Dissociation Constant	2.0
S	Basal Energetic Requirement	$[1, 5]$
D	Division Threshold	$5S$
$T/\Delta t$	Total Simulation Time Steps	[30–100]
v_{max}	Max Cell Velocity	10
$Cell_{\text{max}}$	Max Cell Count	[2000, 10,000]
$Cell_{\text{adjusted}}$	Adjusted Cell Count	$[1000, 9000]$
$Cell_{\text{multi}}$	Cell Multiplier	1

The cells' environment consists of discrete locations that define the external concentration and the position of the cells. The cells have length $\ell = 2$, with the left and right defined as being the current position minus one and the current position plus one, respectively. The simulation is optimized for limited computational resources through approximation of the real cell count in the environment by keeping the max number of environment cells lower than $Cell_{\text{max}}$ during any single time step. If $Cell_{\text{max}}$ is surpassed in a given time step, the number of cells at each location is reduced, and individual cells are removed at random so that the total number of environment cells is equal to a lower cell count $Cell_{\text{adjusted}}$, as defined by the system parameters in Table 1. These decrements of the total population are carried out in such a way that the overall spatial distribution of cells remains the same. A running multiplier $Cell_{\text{multi}}$ is then iterated upon, by multiplying it with the ratio between the initial high cell count and $Cell_{\text{adjusted}}$. A real cell count is then calculated by multiplying the current number of environmental cells by this multiplier. The cell stress $c_* = S/k$ is defined here as the of the metabolic cost S to the substrate absorption coefficient k, as defined in Section 2.1. Heuristically speaking, the cell stress parameter c_* represents the average concentration of either substrate that a cell would have to experience along its spatial trajectory, in order to maintain a net gain in metabolites (cf. (1)).

We employ a non-uniform distribution of A and B molecule concentrations in order for the cells to sense and respond to spatial differences in substrate concentrations. The concentrations of molecular substrates A and B, which remain static throughout the simulation time despite cell intake, are given by the von Mises distribution as

$$[C]_i = [C]_{\max} \exp\left[\kappa_C \frac{\cos(2\pi(\bar{x}_i - \mu_C)/L)}{(L\, I_0(\kappa_C))}\right], \tag{14}$$

where $C \in \{A, B\}$, $\mu_A = 25$ and $\mu_B = 75$, respectively, $\kappa_A = \kappa_B = 0.1$, and I_0 is the modified Bessel function of the first kind [31]. We choose the von Mises distribution because it is the maximal entropy distribution on a periodic support for a given mean and circular variance [32]. $[A]_{\max} = [B]_{\max} = 500$ for all simulations and they remain static over time and do not change in response to the cell's absorption process. The nonuniform distribution of molecules of type A and B across the environment results in a non-zero expected gradient in the number of bound receptors along a cell. Therefore, cells can use chemotaxis to improve their food intake, setting the stage for a non-trivial analysis of the two receptor allocation strategies.

5. Numerical Results

In the following, we present the numerical results obtained by running the simulations described in Section 4, based on the computational model detailed in Section 2.2, and computing the metrics defined in Section 3 from the obtained data.

In Figures 4 and 5, we show the cell's environment for the equal and the adaptive receptor allocation strategies, respectively. The lower plot of either figure shows the A and B external concentrations following the von Mises distribution for the entire simulation space L. The center of the figures show two cells, each represented by a different color, moving left or right in the environment space. 'x' is used here to represent if a cell is divided in a given time step. The simulation for this figure is based on the parameters from Table 1, where i is equal to [1–100], k is equal to 3.0, S is equal to 5, $Cell_{\max}$ is equal to 2000, $Cell_{\text{adjusted}}$ is equal to 1000, and $T/\Delta t$ is equal to 100 time steps. The upper plot of the figures represents the division density, where the cells divide, and the cell density, where the cells are. The cell's distribution statistics were compiled from 100 time steps and an initial cell at each location. Here, the cells have an infinite $v_{\max}$.

In Figures 6 and 7, we show a heat map of the cell density in the A_{in} and B_{in} plane for equal and adaptive receptor allocation strategies, respectively. A black line that runs from 25 to 200 in each direction represents the cell division surface $\mathcal{S}_{\text{div}}$, as defined in Section 2.1. The simulation parameters used in these figures are the same as in Table 1, specifically with i equal to [1–100], k equal to 5, S equal to 5, $Cell_{\max}$ equal to 10,000, $Cell_{\text{adjusted}}$ equal to 9000, and $T/\Delta t$ equal to 30. Here, the figure is split into a two-dimensional bin matrix where the corresponding color of each bin represents the portion of cells corresponding to the bin. The color representations are consistent between Figures 6 and 7 to the same portion of cells in each figure. Internal resource levels A_{in} and B_{in} are recorded in each time step once before the cell has a chance to divide, but after the absorption state, and again after the cell has a chance to divide. This method allows for a better visualization of the internal state of the cell.

In Figures 8 and 9, we show a heat map of the $A(\bar{x})$ and $B(\bar{x})$, respectively, given A_{in} and B_{in}, for the equal receptor allocation strategy. The simulation parameters used here are the same as Figures 6 and 7. Here, each bin in the figures represents the portion of $A(\bar{x})$ or $B(\bar{x})$ weighted on how many visits a cell makes to location $\bar{x}$. This figure gives some insight into where the average cell is when it has some internal state. As in Figures 6 and 7, $A(\bar{x})$ and $B(\bar{x})$ are recorded once before the cell can divide, and again after the cell has a chance to divide.
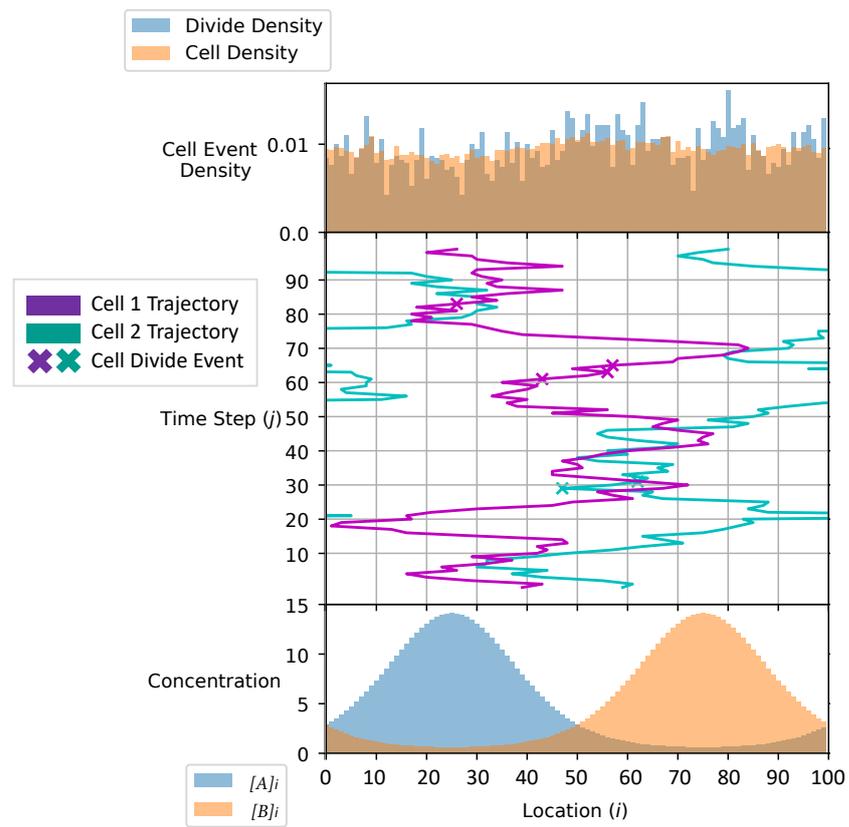
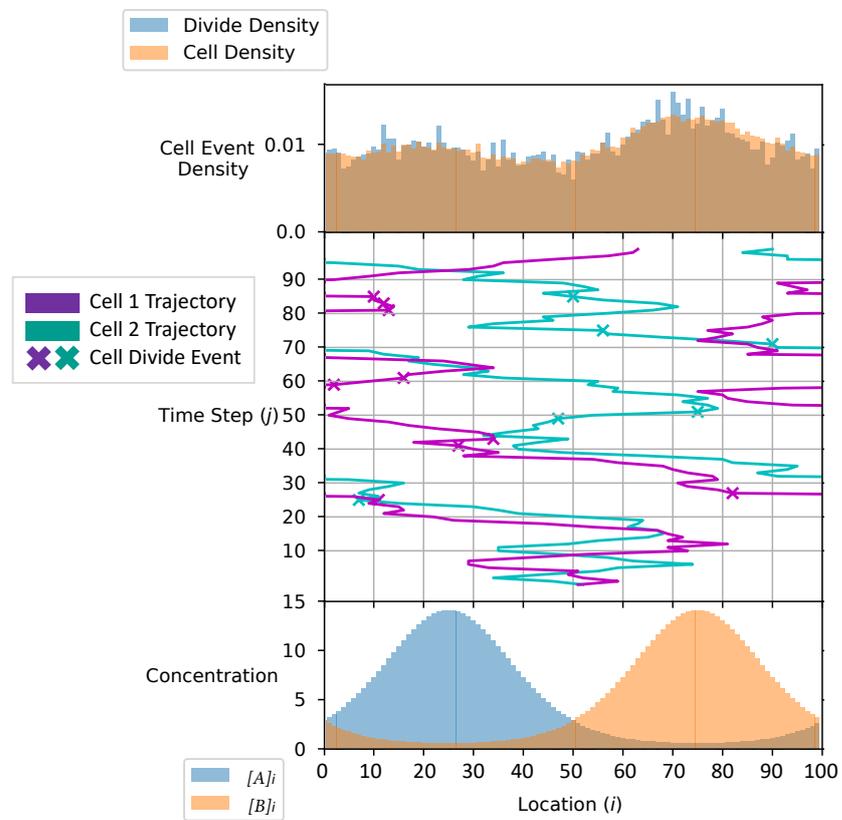**Figure 4.** Simulation environment of the equal receptor allocation strategy.

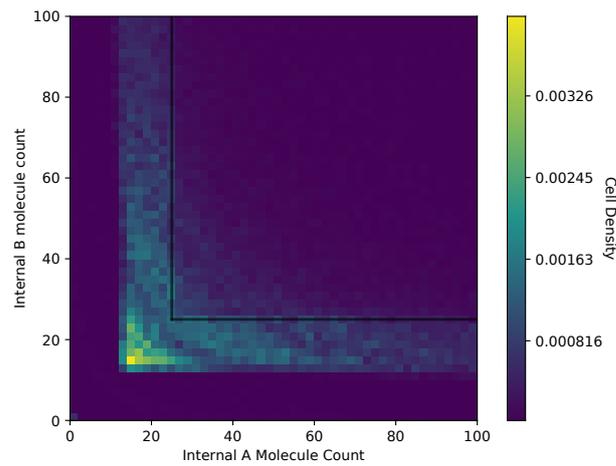**Figure 5.** Simulation environment of the adaptive receptor allocation strategy.

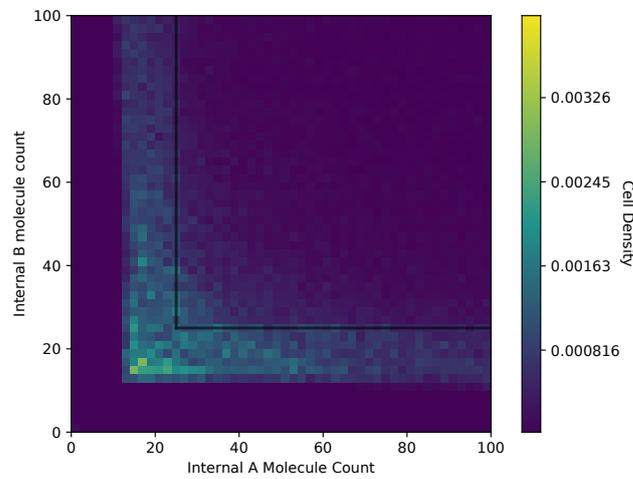**Figure 6.** Cell density in A_{in}, B_{in} space of the equal receptor allocation strategy.

Figure 7. Cell density in A_{in}, B_{in} space of the adaptive receptor allocation strategy.
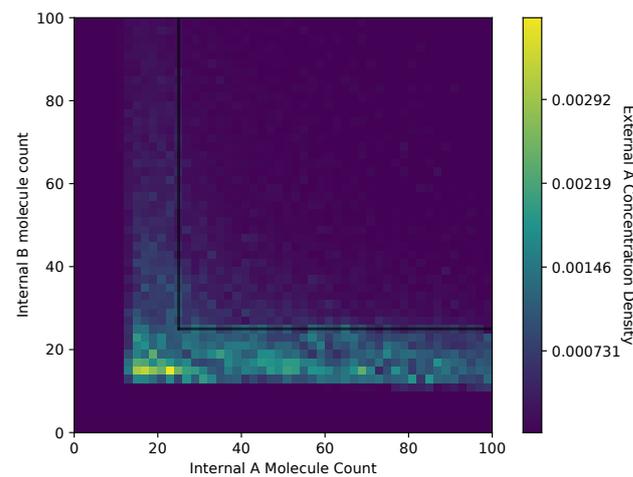

Figure 8. Heat map of the external A concentration with respect to A_{in} B_{in} in the equal receptor allocation strategy.

In Figures 10 and 11, we show a heat map of the $A(\bar{x})$ and $B(\bar{x})$, respectively, given A_{in} and B_{in}, for the adaptive receptor allocation strategy. These figures use the same simulation parameters as Figures 6 and 7. The color coding of the densities of $A(\bar{x})$ and $B(\bar{x})$ is consistent across Figures 8–11.
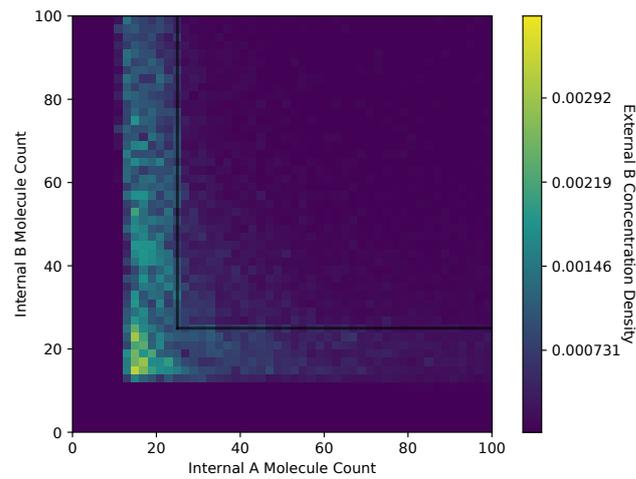
Figure 9. Heat map of the external B concentration with respect to A_{in} B_{in} in the equal receptor allocation strategy.
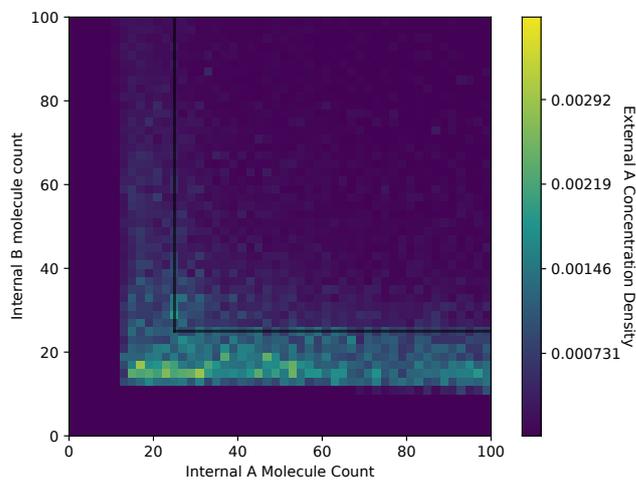
Figure 10. Heat map of the external A concentration with respect to A_{in} B_{in} for the adaptive receptor allocation strategy.
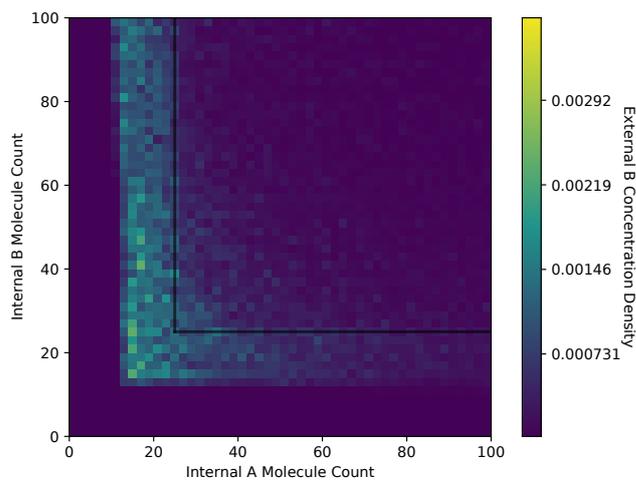
Figure 11. Heat map of the external B concentration with respect to A_{in} B_{in} for the adaptive receptor allocation strategy.

In Figures 12 and 13, we show the growth rate G_T of the cells and the average MI $MI(T)$, respectively, as defined in Section 3, for the equal receptor allocation strategy and

the adaptive receptor allocation strategy as a function of the cell stress c_*. The simulation parameters used to obtain these results are the same as in Table 1, where i is equal to [1–100], S is equal to 1, for $Cell_{max}$ equal to 10,000, $Cell_{adjusted}$ equal to 9000, and $T/\Delta t$ equal to 31. The growth rate G_T was found in another simulation with the same parameters, where the cell's movement was sampled from a uniform random distribution between $-v_{max}$ and v_{max} at cell stress $c_* = 1$ This growth rate was found to be 0.17 slightly below the equal receptor strategy.
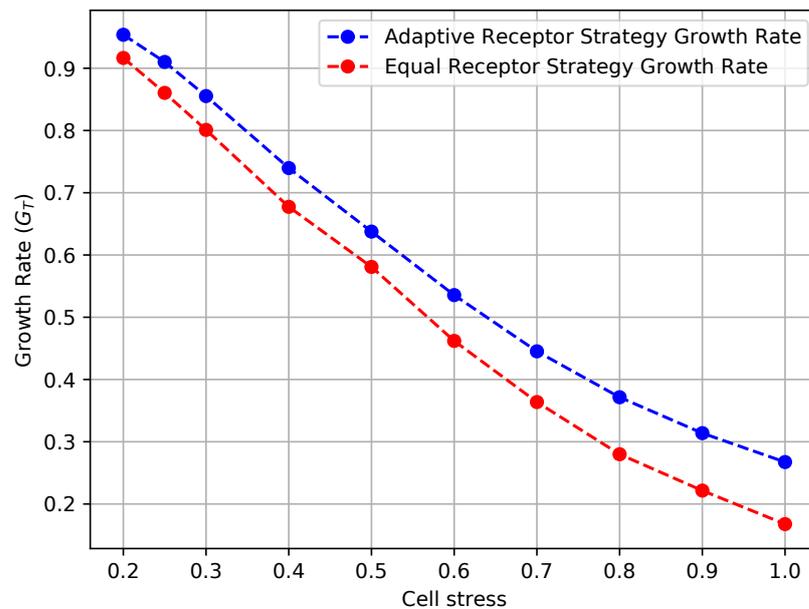
Figure 12. Growth rate of the receptor strategies with respect to cell stress.
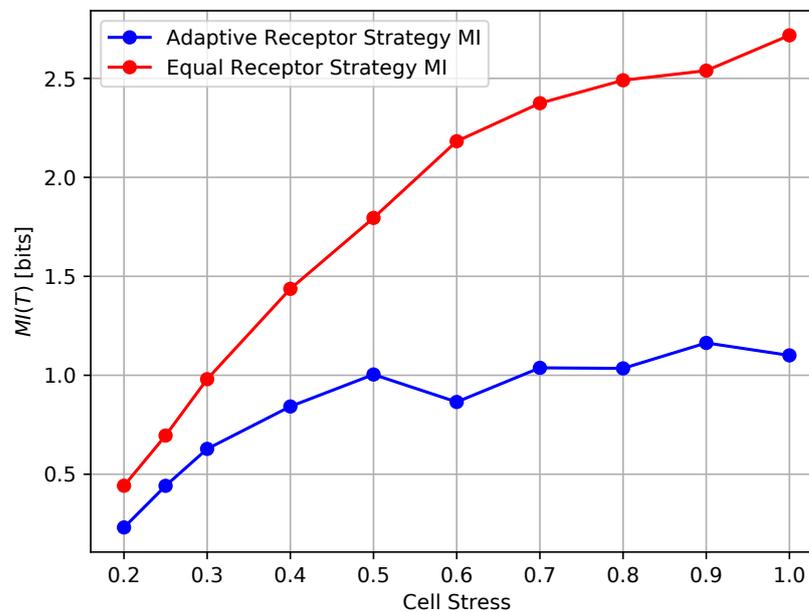
Figure 13. Mutual information of the receptor strategies with respect to cell stress.

In Figures 14–16, we show the average input entropy $E_{\text{# of cells},T}[H_{cell}(X)]$ and average conditional entropy $E_{\text{# of cells},T}[H_{cell}(X|Y)]$, the growth rate G_T of the cells, and the average MI $MI(T)$, respectively, of the Equal Receptor and Adaptive Receptor strategies as a function of the noise factor γ. The input entropy $H_{cell}(X)$ and conditional entropy $H_{cell}(X|Y)$ of a cell are defined as follows:

$$H(X) \quad = \quad -\sum_{x \in X} P_X(x) log_2 P_X(x), \tag{15}$$

$$H(X \mid Y) \quad = \quad -\sum_{x \in X} \sum_{y \in Y} P_{X,Y}(x,y) log_2 \frac{P_{X,Y}(x,y)}{P_Y(y)}, \tag{16}$$

$$H_{cell}(X) \quad = \quad H(A_R) + H(B_R) + H(A_L) + H(B_L), \tag{17}$$

$$H_{cell}(X|Y) \quad = \quad H(A_R|A_R^*) + H(B_R|B_R^*) + H(A_L|A_L^*) + H(B_L|B_L^*), \tag{18}$$

where A_R, B_R, A_L, and B_L are the environmental concentrations of molecules A or B on the right and left of a cell, respectively, and A_R^*, B_R^*, A_L^*, and B_R^* are the consequent output number of bound receptors on the cell. Finally, in Figure 17, we show a combined plot of the growth rate G_T for the Equal Receptor and Adaptive receptor strategies against the corresponding MI $MI(T)$, both as a function of an increasing γ factor. The simulation parameters used to obtain these results are the same as for the results shown in Figure 13.
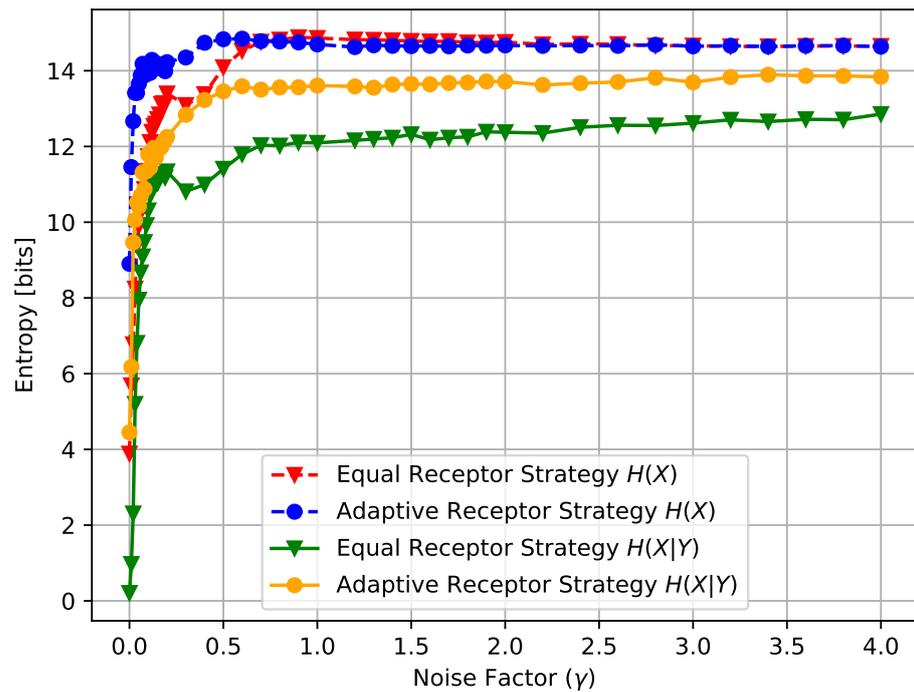

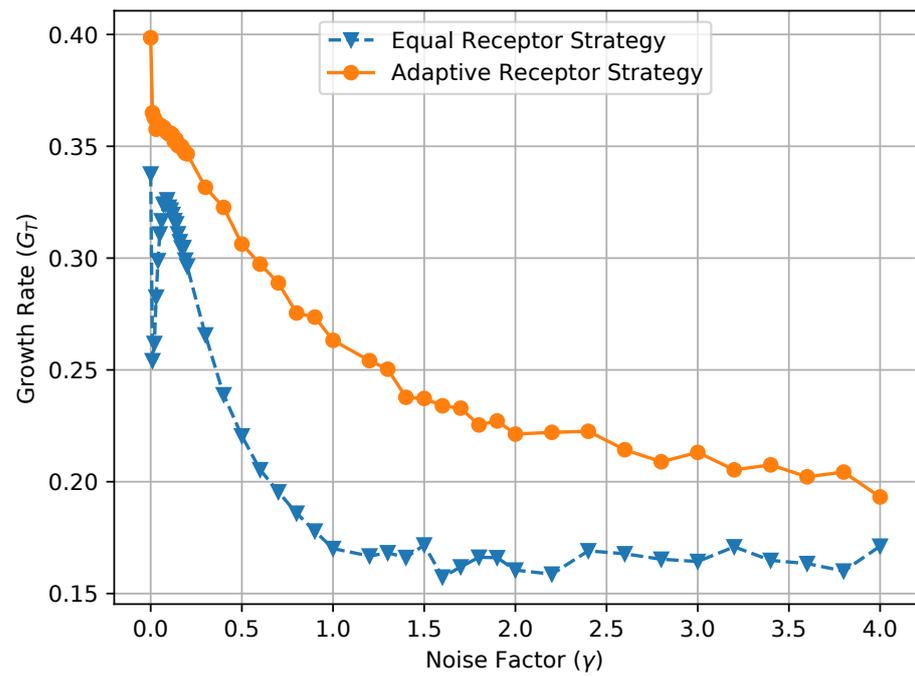**Figure 14.** $H(X)$ and $H(X|Y)$ for the equal receptor and adaptive receptor strategy.

Figure 15. Growth of the equal receptor strategy and adaptive receptor strategy with respect to receptor noise.
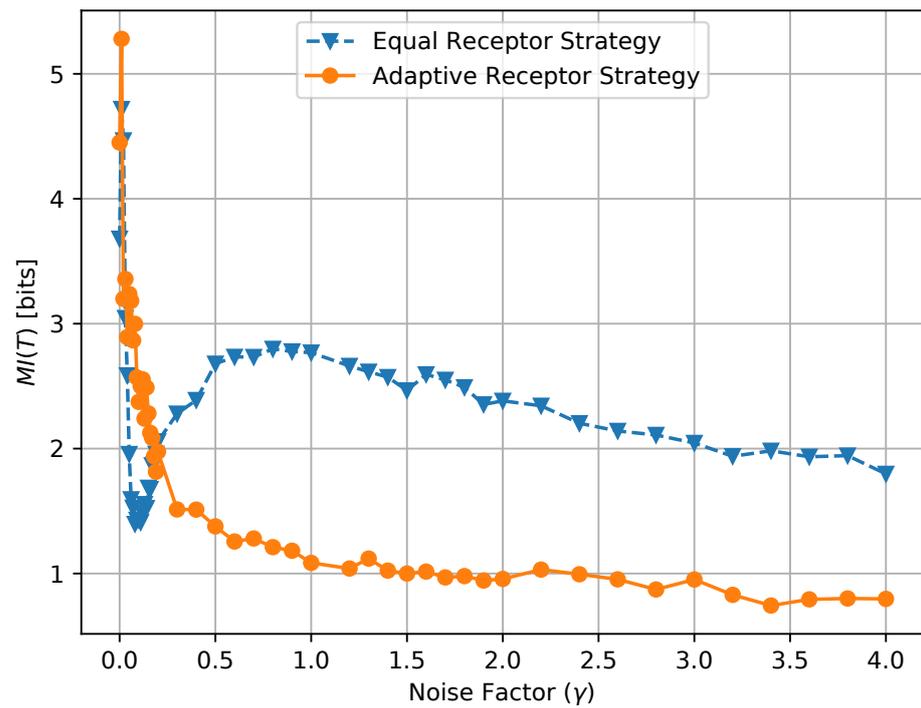

Figure 16. $MI(T)$ of the equal receptor strategy and the adaptive receptor strategy with respect to receptor noise.
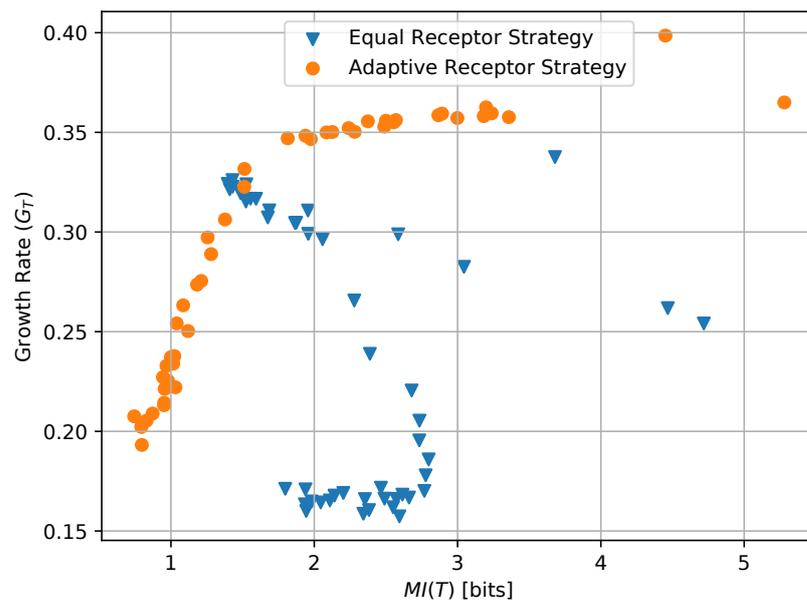
Figure 17. Growth of the equal receptor strategy and the adaptive receptor strategy with respect to the $MI(T)$ as the receptor noise is increased.

The noise factor γ is defined as the ratio between the variance of a Gaussian distribution and the variance of the Binomial distribution defined in (2), respectively. This Gaussian distribution is utilized in place of the Binomial in (2) to obtain the results in Figures 14–16 as a function of a varying receptor noise. This is expressed as a Gaussian with the same average as the corresponding Binomial, and a variance equal to the variance of the corresponding Binomial multiplied by the noise factor γ, as

$$A_R^* \sim \mathcal{N}\left(\frac{A_{\text{tot}}}{2} p_{\text{bind},A_R},\ \left(\frac{A_{\text{tot}}}{2} \right) p_{\text{bind},A_R} (1 - p_{\text{bind},A_R}) \gamma \right), \tag{19}$$

where A_{tot} and p_{bind,A_R} are defined in (2). We utilize similar Gaussian distributions with analogous definitions for A_L^*, B_R^*, and B_L^*, respectively.

6. Discussion

The cell division and cell density distributions plotted in Figures 4 and 5 show how the cell strategies are responding to stress and the static substrate distribution in the environment. Considering that the cells with the equal receptor allocation strategy have a lower growth rate, as shown in Figure 12, the higher density of cells in the interval of two concentration peaks is likely a necessity in that cells in this region would be close to either peak in order to stay alive. In contrast, the cells in the adaptive receptor allocation strategy are clustering around the two highest peaks of the $[A]_i$ and $[B]_i$ distributions. Considering that these cells have a comparatively higher growth rate, this can be interpreted as the cells' ability to stay in higher A or B concentration regions for longer.

When a large number of cells have a higher division rate in a particular region of the environment that region's average effect on the cells' state will result in a higher density. The cell density in Figure 6, for the equal receptor allocation strategy, is larger near the origin. Considering that the cell's divide density location in Figure 4 is centered around the middle of the two concentration masses, the larger number of cell divisions in this region could reveal that the majority of the distribution of the cells exist in a state having an equal and relatively low, compared to the adaptive allocation strategy, internal storage A_{in} and B_{in}. This is also consistent with Figure 12, where the equal receptor allocation strategy has a lower growth rate compared to the adaptive receptor allocation strategy. Conversely, in Figure 7, the adaptive receptor allocation strategy has a spread-out cell density with a larger portion contained in the area corresponding to less than 50 A_{in} and 50 B_{in}. This can

be interpreted as this strategy allowing for a larger internal storage of A_{in} and B_{in} during the simulation.

Both receptor allocation strategies have a sharp cutoff in the $A(\bar{x})$ and $B(\bar{x})$ when A_{in} or B_{in} are less than 10 (Figures 8–11), which is likely due to cell death in the simulation. In this region of the heat map plots, there is a non-zero cell density, albeit very low. This can be interpreted as the cells' chance of survival being greatly reduced for these low values of internal storage.

The equal receptor allocation strategy has a relatively uniform density of internal storage $A(\bar{x})$ and $B(\bar{x})$ along a strip between $B_{\text{in}} = [10, 25]$ and $A_{\text{in}} = [10, 25]$, as shown in Figures 8 and 9. The adaptive receptor allocation strategy in this same strip is not as uniform, while a high density in $A(\bar{x})$ and $B(\bar{x})$ occurs when $B_{\text{in}} = 60$ and $A_{\text{in}} = 60$, respectively.

The results shown in Figures 12 and 13 clearly demonstrate that, *for every value of cell stress imposed in the simulation, the adaptive receptor allocation strategy results in a **higher growth rate and correspondingly lower average** $MI(T)$ than the equal receptor allocation strategy*, where these differences increase as the cell stress increases. In addition, as the cell stress increases, both allocation strategies result in not only a lower growth rate, which is expected, but also a higher average $MI(T)$. This might be explained by how cells survive with respect to their strategy [33]. With higher stress, cells with a higher "capacity" to receive information (either because of their locations or their receptor allocations) from the environment have a higher probability to be naturally selected over others. This selection seems more pronounced for the equal receptor allocation and more subtle for the adaptive strategy, where cells tend to adapt their "capacity" to the more stringent constraint and increase their fitness.

In Figure 14, the initial rise in input entropy for low noise, as the noise increases, is due to a higher mobility of the cells to sample more diverse environmental locations, and a higher range of different A_{in} and B_{in} values. The corresponding growth rate in Figure 15 and mutual information in Figure 16 for the equal receptor allocation seem to reach optimal values (maximum growth with minimum mutual information) before plunging when the noise increases further. The same does not happen for the adaptive receptor strategy. The input entropy then stabilizes for both strategies after the noise factor γ reaches value 1. For higher noise, the subtle variations in the conditional entropy seem to dominate the variations in $MI(T)$, shown in Figure 16, which is therefore more stable, as the growth rate in Figure 15. As the noise increases, $MI(T)$ and the growth rate decrease for both strategies, as expected. Nevertheless, for all the noise values, the adaptive receptor allocation results in higher growth rates, achieved with a lower mutual information for all but in the aforementioned optimal noise region for the equal allocation.

The combined plot in Figure 17 offers the opportunity of a direct comparison with the work presented in [20]-(Figures 1c, 2b, 3b and 4b). In particular, while our model does not allow for an independent selection of the mutual information between organisms and environment during simulation, it is clear that the definition of "semantic information" from [20] is here not directly applicable since the degradation of the information channel operated by an increased noise is in our case countered by an optimization of the channel by each individual organism in the simulated population. This cannot be represented by a single average mutual information, and underscoring the need for a metric to measure the emergence of "subjective information", i.e., mutual information between environment and organism that varies across individuals in the same population.

Measuring the Emergence of Subjective Information

Given the results detailed above, we proposed to measure the emergence of "subjective information" (SI) over a time T by taking the average over time of the standard deviation of the MI of each individual cell MI_{cell} over the population, as

$$SI = \mathrm{E}_T[\sigma_{\text{Total \# of cells}}(MI_{cell})], \tag{20}$$

where $\sigma(.)$ denotes the standard deviation operator. According to this definition, the equal receptor allocation strategy has by definition SI equal to 0, as demonstrated also empirically in Figures 18 and 19, while the adaptive allocation strategy shows a positive SI in all considered situations. The proposed measure is relatively stable as a function of the noise, as shown in Figure 18, but, in the case of the adaptive receptor allocation strategy, it has an overall tendency to increase with cell stress. We reserve further study of "subjective information" measurements to future work.
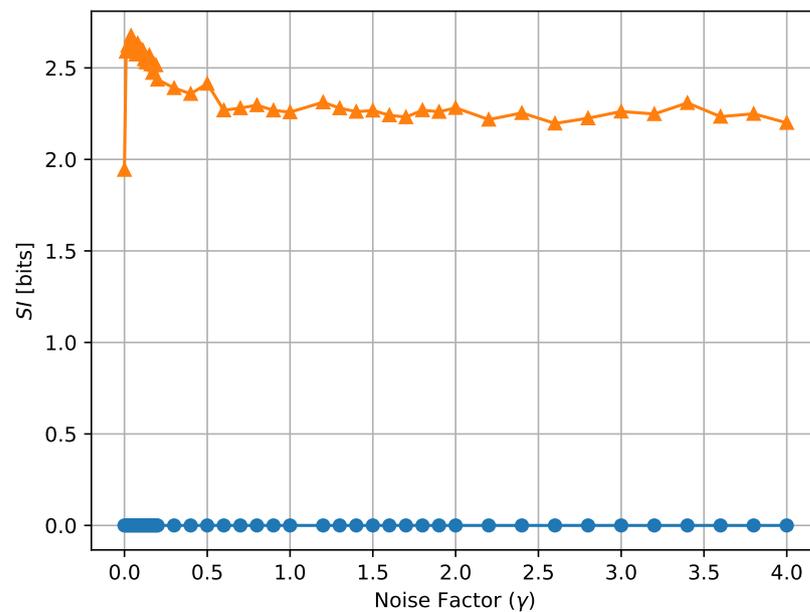
Figure 18. Subjective information of the equal receptor and adaptive receptor strategies as the noise factor (γ) is increased. The simulation parameters used to obtain these results are the same as for the results shown in Figure 13.
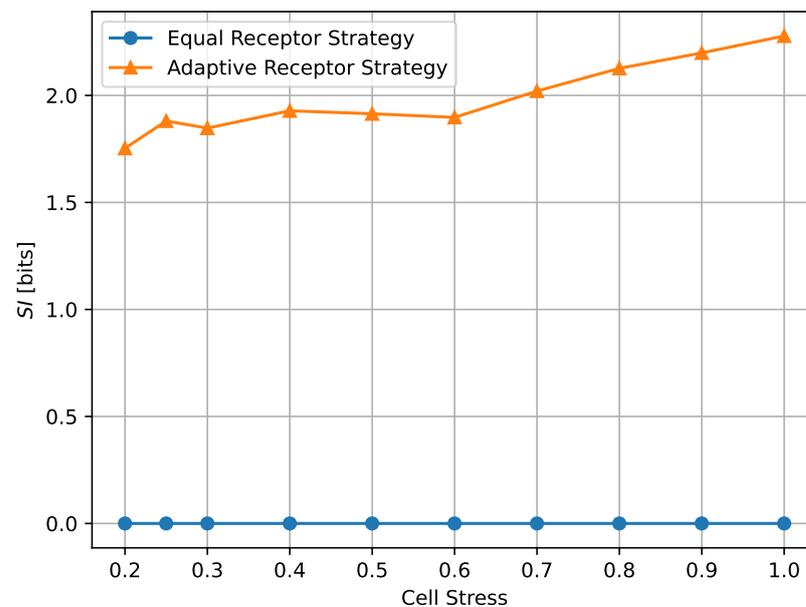

Figure 19. Subjective information of the equal receptor and adaptive receptor strategies as the cell stress factor is increased. The simulation parameters used to obtain these results are the same as for the results shown in Figure 13.

7. Conclusions

In this paper, we introduced a simulation-based scenario to analyze the emergence of a form of "subjective information" and its correlation with the survival of a biological organism. In particular, we defined an abstract mathematical model able to capture the parameters and behaviors of a population of single-celled organisms while they move through chemotaxis to absorb two essential nutrients from the environment. This model is then translated into a computational state machine, which is then utilized in a simulation framework constructed specifically to characterize and measure the emergence of the aforementioned information. For this, two different strategies adopted by these cells are considered, namely, based either on maximizing information on the two essential substrates or, alternatively, reducing this information by focusing on what is more important for survival. Simulation results based on these two strategies for different parameters related to cell's survival stress are compared in terms of information efficiency of the cells Vs growth rate. The obtained results clearly reveal that the strategy that maximizes information efficiency results in a lower growth rate with respect to the strategy that has lower information efficiency but optimizes the information channel of each individual to better focus on the pieces of information from the environment that are more important for survival, i.e., the "subjective information".

A more robust model and in vivo experiments will be required to better define and quantify "subjective information" of a living system. With this analysis, this new information concept may identify an important aspect of biological systems that can be used in tandem with other information theoretic principles studied in prior literature.

References

1. Farnsworth, K.D.; Nelson, J.; Gershenson, C. Living Is Information Processing: From Molecules to Global Systems. *Acta Biotheor.* **2013**, *61*, 203–222. [CrossRef] [PubMed]
2. Smith, J.M. The Concept of Information in Biology. In *In the Scope of Logic, Methodology and Philosophy of Science: Volume Two of the 11th International Congress of Logic, Methodology and Philosophy of Science, Cracow, August 1999*; Gärdenfors, P., Woleński, J., Kijania-Placek, K., Eds.; Springer: Dordrecht, The Netherlands, 2002; pp. 689–699.
3. Adami, C. Information theory in molecular biology. *Phys. Life Rev.* **2004**, *1*, 3–22. [CrossRef]
4. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]
5. Cover, T.M. *Elements of Information Theory*; John Wiley & Sons: Hoboken, NJ, USA, 1999.
6. Levchenko, A.; Nemenman, I. Cellular noise and information transmission. *Curr. Opin. Biotechnol.* **2014**, *28*, 156–164. [CrossRef]
7. Vinga, S. Information theory applications for biological sequence analysis. *Briefings Bioinform.* **2013**, *15*, 376–389. [CrossRef] [PubMed]
8. Suderman, R.; Bachman, J.A.; Smith, A.; Sorger, P.K.; Deeds, E.J. Fundamental trade-offs between information flow in single cells and cellular populations. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 5755–5760. [CrossRef] [PubMed]
9. Sakkaff, Z.; Immaneni, A.; Pierobon, M. Estimating the Molecular Information Through Cell Signal Transduction Pathways. In Proceedings of the 2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Kalamata, Greece, 25–28 June 2018; pp. 1–5. [CrossRef]
10. Harper, C.; Pierobon, M.; Magarini, M. Estimating Information Exchange Performance of Engineered Cell-to-cell Molecular Communications: A Computational Approach. In Proceedings of the IEEE INFOCOM 2018—IEEE Conference on Computer Communications, Honolulu, HI, USA, 16–19 April 2018; pp. 729–737. [CrossRef]
11. Akyildiz, I.F.; Pierobon, M.; Balasubramaniam, S. An Information Theoretic Framework to Analyze Molecular Communication Systems Based on Statistical Mechanics. *Proc. IEEE* **2019**, *107*, 1230–1255. [CrossRef]

12. Agarwala, E.K.; Chiel, H.J.; Thomas, P.J. Pursuit of food versus pursuit of information in a Markovian perception–action loop model of foraging. *J. Theor. Biol.* **2012**, *304*, 235–272. [CrossRef]
13. Bergstrom, C.T.; Rosvall, M. The transmission sense of information. *arXiv* **2008**, arXiv:0810.4168.
14. Tishby, N.; Polani, D. Information Theory of Decisions and Actions. In *Perception-Action Cycle*; Springer: New York, NY, USA, 2011. [CrossRef]
15. Barker, T.; Thomas, P.J.; Pierobon, M. Subjective Information in Life Processes: A Computational Case Study. In Proceedings of the Eight Annual ACM International Conference on Nanoscale Computing and Communication, Virtual, 7–9 September 2021; Association for Computing Machinery: New York, NY, USA, 2021; NANOCOM '21.
16. Ortega, P.A.; Braun, D.A. Thermodynamics as a theory of decision-making with information-processing costs. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **2013**, *469*, 20120683. [CrossRef]
17. Donaldson-Matasci, M.C.; Bergstrom, C.T.; Lachmann, M. The fitness value of information. *Oikos* **2010**, *119*, 219–230. [CrossRef] [PubMed]
18. Kelly, J., Jr. A new interpretation of the information rate. *Bell Syst. Tech. J.* **1956**, *35*, 917–926. [CrossRef]
19. Mafessoni, F.; Lachmann, M.; Gokhale, C.S. On the fitness of informative cues in complex environments. *J. Theor. Biol.* **2021**, *527*, 110819. [CrossRef] [PubMed]
20. Kolchinsky, A.; Wolpert, D.H. Semantic information, autonomous agency and non-equilibrium statistical physics. *Interface Focus* **2018**, *8*, 20180041. [CrossRef]
21. Thomas, P.J.; Spencer, D.J.; Hampton, S.K.; Park, P.; Zurkus, J.P. The diffusion-limited biochemical signal-relay channel. In *Advances in Neural Information Processing Systems*; Citeseer: Princeton, NJ, USA, 2004; pp. 1263–1270.
22. Thomas, P.J.; Eckford, A.W. Capacity of a simple intercellular signal transduction channel. *IEEE Trans. Inf. Theory* **2016**, *62*, 7358–7382. [CrossRef]
23. Kimmel, J.M.; Salter, R.M.; Thomas, P.J. An information theoretic framework for eukaryotic gradient sensing. In *Advances in Neural Information Processing Systems*; Curran Associates: Red Hook, NY, USA, 2007; pp. 705–712.
24. Rivoire, O.; Leibler, S. The value of information for populations in varying environments. *J. Stat. Phys.* **2011**, *142*, 1124–1166. [CrossRef]
25. Vergassola, M.; Villermaux, E.; Shraiman, B.I. 'Infotaxis' as a strategy for searching without gradients. *Nature* **2007**, *445*, 406–409. [CrossRef]
26. Lynch, M.; Marinov, G.K. The bioenergetic costs of a gene. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 15690–15695. [CrossRef]
27. Ilker, E.; Hinczewski, M. Modeling the growth of organisms validates a general relation between metabolic costs and natural selection. *Phys. Rev. Lett.* **2019**, *122*, 238101. [CrossRef]
28. Yang, N.; Hinner, M. Getting across the cell membrane: An overview for small molecules, peptides, and proteins. *Methods Mol. Biol.* **2015**, *1266*, 29–53. [CrossRef]
29. Wesel, R.D.; Wesel, E.E.; Vandenberghe, L.; Komninakis, C.; Medard, M. Efficient binomial channel capacity computation with an application to molecular communication. In Proceedings of the 2018 Information Theory and Applications Workshop (ITA), San Diego, CA, USA, 11–16 February 2018; pp. 1–5.
30. Thomas, P.J.; Eckford, A.W. Shannon capacity of signal transduction for multiple independent receptors. In Proceedings of the 2016 IEEE International Symposium on Information Theory (ISIT), Barcelona, Spain, 10–15 July 2016; pp. 1804–1808.
31. Mardia, K.V.; Jupp, P.E. *Directional Statistics*; John Wiley & Sons: Hoboken, NJ, USA, 2009; Volume 494.
32. Jammalamadaka, S.R.; SenGupta, A. *Topics in Circular Statistics*; World Scientific: Singapore, 2001.
33. Snyder, R.; Ellner, S. Pluck or Luck: Does Trait Variation or Chance Drive Variation in Lifetime Reproductive Success? *Am. Nat.* **2018**, *191*, E90–E107. [CrossRef] [PubMed]