

Special Issue on Biological Applications of Information Theory in Honor of Claude Shannon's Centennial—Part 1

The Use of Rate Distortion Theory to Evaluate Biological Signaling Pathways

Pablo A. Iglesias, *Senior Member, IEEE*

(*Invited Paper*)

Abstract—Cells must make decisions based on noisy measurements of their environment. One way for cells to manage this noise is to attempt to minimize its effect. However, this may not always be possible, or may prove to be costly. On the other hand, the noise may be managed so that the correct decision is made most of the time. Shannon introduced rate distortion theory to evaluate the efficiency of systems like this in which error free communication may not be possible or needed. Here, we show how rate distortion theory can be used to analyze cellular decision processes. We show how several stimulus-response curves that are frequently observed in biological signaling pathways arise naturally as the optimal decision strategy based on rate distortion theory.

Index Terms—Biological information theory, computational systems biology, rate distortion theory, biological interactions.

I. INTRODUCTION

ONE OF the principal features of living organisms is their ability to sense and respond to changes in the environment. We see this at the smallest possible scale, where single-cell organisms in the presence of chemical signals respond by moving towards food sources [1], stopping their cell cycle in preparation for mating [2], or inducing differentiation [3]. The complete process of responding consists of a number of separable, but interrelated steps. The first is the sensing mechanism, whereby the environmental signal is detected. The second step is the decision process. The third and final step is the process of generating the response.

Because each of these steps is achieved through molecular interactions, the signals that are transmitted are subject to stochastic fluctuations [4]. The effect of these noisy signals is particularly acute in genetic regulatory networks in which the copy numbers of the various molecules can be considerably small — in the range of single or double digits. However, the consequence of stochastic fluctuations is not limited to these systems. Signaling in larger cells is still subject to stochastic disturbances; in fact, these fluctuations can help to regulate desirable cell functions. For example, cell migration in the amoeba *Dictyostelium discoideum* is now believed

Manuscript received May 31, 2016; revised August 17, 2016; accepted October 25, 2016. Date of publication November 1, 2016; date of current version January 6, 2017. The associate editor coordinating the review of this paper and approving it for publication was P. J. Thomas.

The author is with the Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD 21218 USA (e-mail: pi@jhu.edu). Digital Object Identifier 10.1109/TMBMC.2016.2623600

to be largely driven by stochastic triggering of an excitable system [5], [6].

That stochastic fluctuations are an important fact of life in biological systems has been appreciated from the earliest days of systems biology [7]. Substantial effort has been expended in developing the theory dealing with these noisy signaling schemes [8], simulation techniques that can properly account for the variability [9], as well as experimental protocols that clearly demonstrate the effect of this noise [10].

To understand how biological systems function in the presence of noise, there has been considerable interest in trying to use tools from engineering. For example, a number of studies aim at understanding how the noise filtering machinery in cells functions [11], [12]. Recently, a molecular filtering mechanism was designed and implemented in cells [13].

These studies show how noise can be mitigated, but do not address a number of fundamental questions. How does a cell make decisions given the inevitable presence of stochastic fluctuations in all sensed signals? Do the regulatory systems in cells take noise into account in their decisions? For guidance in this matter, there has been an interest in looking at engineering principles. Of particular relevance is Shannon's theory of communication, which set the framework for understanding data transmission in noisy environments. In the last decade there has been much effort trying to understand how cell signaling can be interpreted in the context of Shannon's theory; see, for example, the reviews [14]–[20]. Most of this work has focused on Shannon's channel capacity theorem. Here we argue that his work on rate distortion may be a more reasonable avenue for discerning how biological signaling systems function as information-carrying networks.

The rest of this paper is organized as follows. In Section II we introduce some background requisite material on information, rate distortion theory and provide a short literature review of their uses in biology. We then consider two types of signaling systems and analyze them using rate distortion theory. In Section III, we first study a binary decision system. In Section IV, we adopt this approach to the study of a more general signaling system, in which a cell must determine the correct direction of navigation based on receptor-ligand occupancy. Finally, we conclude with a general discussion in Section V.

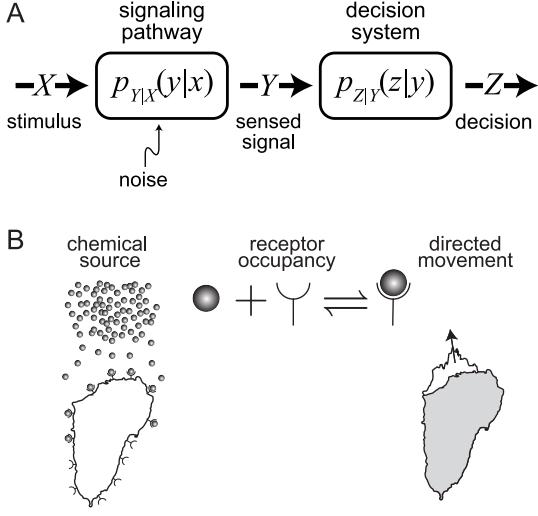


Fig. 1. A. Cell signaling as a communication channel. Cellular stimuli (X) are sensed through a noisy signaling processes giving rise to the sensed signal (Y) that is then used to make a particular cellular decision (Z). B. This cartoon illustrates how directed cell motion fits this framework. A chemical source provides molecules that bind to cell-surface receptors. The resultant receptor occupancy is used by cells to guide movement. In this context, the chemical source is the stimulus, receptor occupancy is the sensed signal, and the direction of motion is the decision.

II. PRELIMINARIES

A. Information Theory and Cell Signaling

The basic set-up is the familiar communication system illustrated in Fig. 1. The signal X , which can represent the concentration of the external environmental signal, is to be used by the cell to make a particular decision, Z . However, making this decision first requires a signaling pathway that senses X and generates Y . This process invariably involves the corruption of the signal so that the decision must be made based on signal Y , which is an imprecise measure of X .

We denote by Y and Z the received and decision signals, respectively, and assume that both are random variables with individual probability distributions $p_Y(y)$ and $p_Z(z)$, respectively, joint distribution $p_{Y,Z}(y,z)$ and conditional distribution $p_{Z|Y}(z|y)$. For the random variable Y , the entropy is the familiar

$$H(Y) = - \sum_y p_Y(y) \log p_Y(y)$$

which is a measure of uncertainty, and the mutual information

$$I(Y; Z) = H(Z) - H(Z|Y) \quad (1)$$

$$= \sum_{y,z} p_{Y,Z}(y,z) p_Y(y) \log \frac{p_{Y,Z}(y,z)}{p_Y(y)p_Z(z)}, \quad (2)$$

which quantifies the reduction of uncertainty in Z after measuring Y [21].

Shannon's capacity theorem deals with lossless information transfer across the channel. In particular, he showed that if communication across a noisy channel with negligible error is to be achieved, then the rate of transmission (in bits/s) must not exceed the channel capacity, given

$$C = \sup_{p_Y(y)} I(Y; Z).$$

In rate distortion theory, however, we assume that some loss of information is inevitable so that, based on z , the signal y cannot be reproduced. To characterize this error we introduce a *distortion function*, $d(y, z) \geq 0$. In engineering, popular choices are the squared error distortion

$$d(y, z) = (y - z)^2$$

and, particularly for y and z elements of a common finite alphabet, the Hamming distortion

$$d(y, z) = \begin{cases} 0, & \text{if } y = z, \\ 1, & \text{if } y \neq z. \end{cases}$$

Given a distortion function and a joint distribution $p_{Y,Z}(y,z)$, we can compute the expected distortion

$$\langle d(y, z) \rangle = \sum_{y,z} p_{Y,Z}(y,z) d(y, z).$$

Now, suppose that we set D as the limit to the allowable loss of fidelity; that is, we want to make sure that

$$\langle d(y, z) \rangle \leq D.$$

Our goal is then to find a $p_{Y,Z}(y,z)$ that achieves this. However, in general there will be many such distributions. So we can then ask whether there is one (or more) that is optimal in some sense. In rate distortion theory, this amounts to finding the $p_{Z|Y}(z|y)$ with the smallest mutual information; i.e.,

$$p_{Z|Y}^*(z|y) = \arg \min_{p_{Z|Y}(z|y): \langle d(y, z) \rangle \leq D} I(Y; Z).$$

The resultant mutual information is known as the *rate distortion function*

$$R(D) = \min_{p_{Z|Y}(z|y): \langle d(y, z) \rangle \leq D} I(Y; Z).$$

Note that $C \geq R(D)$.

The solution to this variational problem is, in general, intractable, though there are some known solutions to some specific cases; for example a Gaussian memoryless channel with quadratic distortion function, or a binary source with Hamming distortion [21]. In general, however, the Blahut-Arimoto algorithm, outlined below, can generate both the optimal $p_{Z|Y}^*(z|y)$ as well as $R(D)$.

We note that in this approach, the mutual information is the cost associated with minimization, as increasing the allowed mutual information leads to lower expected distortion. It is natural to ask whether such a currency is appropriate in the context of biological signaling. The general assumption is that increasing mutual information imposes real-life costs to the cell, which are likely based on energetic constraints. We return to this question in Section V.

B. Computation of the Rate Distortion Function

The Blahut-Arimoto algorithm [21, Sec. 13.8] facilitates computation of $R(D)$ using the method of Lagrange multipliers. The principal steps are:

- 1) Initialize the probability distribution of Z , $p_Z(z)$, as a uniform distribution.
- 2) Given $p_Z(z)$, compute the conditional probability distribution $p_{Z|Y}(z|y)$ that minimizes the mutual information

$I(Y; Z)$ while meeting the distortion constraint. The optimal is:

$$p_{Z|Y}(z|y) = \frac{p_Z(z)e^{-\lambda d(y,z)}}{\sum_z p_Z(z)e^{-\lambda d(y,z)}}$$

- 3) Given $p_{Z|Y}(z|y)$, compute the marginal probability distribution $p_Z(z)$ that minimizes the mutual information $I(Y; Z)$, which is given by:

$$p_Z(z) = \sum_y p_{Z|Y}(z|y)p_Y(y)$$

- 4) Repeat steps 2 and 3 until $p_{Z|Y}(z|y)$ and $p_Z(z)$ converge. As these steps are executed, the limiting mutual information $I(Y; Z)$ equals $R(D)$, where $\langle d(y, z) \rangle = D$ is determined by the Lagrange multiplier λ chosen through the minimization. This process can then be repeated with varying values of λ to generate a collection of optimal strategies $p_{Z|Y}(z|y)$ for which the distortion-information point $(D, I(Y; Z))$ lies on $R(D)$.

C. Experimental Computation of Mutual Information

A number of studies have tried to compute or measure the channel capacity of specific biological systems using characterization (1) for the mutual information. This requires estimates of entropies, which itself involves numerous measurements.

In an early study, Tkačik *et al.* [22] used previously measured concentrations of the proteins bicoid and hunchback to estimate the informational flow during transcriptional regulation. Bicoid is a morphogen, a protein whose concentration gradient is used by a developing fruit fly embryo to establish positional information. Depending on the local concentration of the morphogen, specific genes are turned on or off, thus generating a pattern within the embryo. In the context of each individual cell, the concentration of the sensed morphogen is used to determine that cell's fate. In particular, bicoid turns on the gene hunchback and this process is used to establish the anterior and posterior axis of the fruit fly. Using simultaneous measurements of bicoid (X) and the protein hunchback (Y), Tkačik *et al.* [22] estimated both $p(x)$ and $p(y|x)$ and hence were able to assess the mutual information between these two signals. Their results suggested that the channel capacity of this system was 1.7 bits and the measured mutual information 1.5 bits, which represents approximately 90% efficiency.

This relatively low level of mutual information has been seen in a number of other systems. For example, Cheong *et al.* [23] used TNF (tumor necrosis factor)-dependent activation of two transcription factors. TNF is a signaling protein that regulates the immune system. Binding of TNF to the cell induces a number of cellular responses, including activation of NF-κB (a protein complex that controls transcription) and the transcription factor ATF2. Using a microfluidic set-up, they applied a large range of concentrations of TNF and simultaneously measured the activities of NF-κB and ATF2. Computing the corresponding mutual information allowed them to estimate the channel capacity of the network. From their studies they concluded that both the NF-κB and ATF2 responses transmit only 0.92 and 0.85 bits, respectively, which

is equivalent to being able to distinguish reliably only between the presence or absence of TNF. More recently, Uda *et al.* [24] examined the response of rat PC12 cells to a number of growth factors and found that the channel capacity was approximately one bit, though they did find one growth factor for which the capacity was only ~ 0.5 bits.

One possible way to mitigate the effect of this relatively low capacity for information transmission is by increasing the signal-to-noise ratio of the signal [21]. A common engineering methodology for this is to filter extrinsic noise temporally by, for example, averaging the response over time.

Selimkhanov *et al.* [25] considered whether such dynamic signaling could be at work in the ERK (extracellular signal-regulated kinase), calcium (Ca^{2+}), and NF-κB signaling systems. Their results confirmed that, using static single time point measurements, the signaling pathways transmitted less than one bit. However, using multiple time points led to higher information transmission capacities. This data integration need not be temporal. For example, Hansen and O'Shea [26] showed that the information of the yeast transcription factor Msn2 could also be increased by integrating information from multiple genes. Finally, Voliotis *et al.* [27] demonstrated that, in the ERK signaling pathway, negative feedback increases the amount of information transferred to the nucleus with maximal transfer at intermediate feedback levels [28].

D. Information Flow Through Multiple Channels and Multiple Signals

Beyond computing experimental values of the mutual information capacity, information theory has been able to elucidate some of the features of signaling pathways. For example, Cheong *et al.*'s [23] findings prompted them to consider the effect of network architecture on data transmission. They considered two broad classes of networks. The first, which they referred to as the *bush* architecture, takes a signal and uses this to generate a number of independent parallel channels. In the second, termed the *tree* architecture, information is first transmitted through a common channel (the *trunk*) before separating. Whereas in the former the channel capacity can grow with the number of parallel channels, in the latter the trunk may act as a bottleneck limiting the available information data transmission.

The results above deal with transmission of information about one signal (e.g., TNF) through multiple channels (NF-κB and ATF2). In contrast, Mehta *et al.* [29] used information theory to study how a single channel could be used to infer information from different signals. Their studies involve the *quorum sensing* ability of the bacterium *Vibrio harveyi*. Bacteria are known to regulate gene expression in response to changes in the population density. This requires a means for cells to communicate, and this is achieved through the production, secretion and subsequent sensing of chemicals known as *autoinducers*. Bacteria of the species *V. harveyi* use three different autoinducers to regulate gene expression, some which are produced only by *V. harveyi* cells while others are produced by cells from other bacterial species. Thus, the signals from each of the three different auto-inducers can serve as

a measure of the population density of *V. harveyi* as well as other bacteria in the environment. Their studies showed that the ability to infer information about population density is inhibited by noise but also by interference from the other signals. Moreover, to counter this interference, the cells use feedback loops to control various receptor numbers.

E. The Use of Rate Distortion Theory to Analyze Cell Decisions

The studies referred to above have primarily been aimed at measuring the information-carrying capacity of the biological signaling pathway. However, they do not address the ultimate goal of the information, which is for the cell to make a decision. To determine how much information is actually required to make reliable decisions we suggest that the correct mechanism is rate distortion theory. The use of rate distortion theory to analyze biological systems is still in its infancy, though there have been some attempts [30]–[32]. Below we illustrate how it arises in two classes of signaling systems.

One difference between our studies and those above is that, by focusing on the decision process, we look at how the signal z is generated, based on the perceived signal y . In particular, we assume that given x , there is a corrupted y (which can be described by a probability distribution, $p_Y(y)$). However, there is a correct response to this y . In what follows, we refer to $p_{Z|Y}(z|y)$ as a *stimulus-response* map. In particular, it represents how the response (z) depends on the sensed stimulus (y). Of course, in our context, this is a probabilistic function, as it only describes the likelihood of getting a particular response.

III. EXAMPLE: BINARY DECISION PROCESSES

We first present an example of how rate distortion theory can be used to evaluate the efficacy of a binary decision, akin to the bicoid/hunchback system referred to above. Our standing assumption is that the decision is made depending on whether the sensed signal, y , is above or below a threshold, y_{th} . We assume that the level of the sensed signal is a stochastic process with probability $p_Y(y)$. In particular, we denote by

$$\gamma = \int_0^{y_{\text{th}}} p_Y(y) dy$$

the probability that the external signal is below the threshold. Note that this is a probability based on the sensed signal. The problem could be reformulated based on the external signal x , by computing the probability (based on x) that y is above the threshold.

Let us denote the *decision* by the symbol $Z \in \{\text{low}, \text{high}\}$. If $y \leq y_{\text{th}}$ then the correct decision is $z = \text{low}$. We will assume that the typical situation is one in which the stimulus is low, so that $\gamma > 1/2$.

Based on this assumption, we use as our distortion function the Hamming-like function

$$d(y, z) = \begin{cases} & \begin{array}{cc} y < y_{\text{th}} & y \geq y_{\text{th}} \end{array} \\ \text{z = low} & \begin{array}{cc} 0 & 1 \end{array} \\ \text{z = high} & \begin{array}{cc} 1 & 0, \end{array} \end{cases}$$

For illustrative purposes, let us assume that the decision process is completely binary — that is, it assumes that the

decision is based solely on what y is, relative to y_{th} :

$$\begin{array}{c|cc} & y < y_{\text{th}} & y \geq y_{\text{th}} \\ \hline z = \text{low} & \alpha & 1 - \beta \\ z = \text{high} & 1 - \alpha & \beta, \end{array} \quad (3)$$

where α and β are represent the probabilities that each of the two decisions are reached based on the input. In this system, the mutual information is

$$\begin{aligned} I(Y; Z) &= \sum_{y,z} p_{Y,Z}(y, z) \log \frac{p_{Y,Z}(y, z)}{p(y)p_Z(z)} \\ &= \alpha\gamma \log \alpha + (1 - \alpha)\gamma \log(1 - \alpha) \\ &\quad + \beta(1 - \gamma) \log \beta + (1 - \beta)(1 - \gamma) \log(1 - \beta) \\ &\quad - (\alpha\gamma + (1 - \beta)(1 - \gamma)) \log(1 - \xi) \\ &\quad - ((1 - \alpha)\gamma + \beta(1 - \gamma)) \log \xi \end{aligned} \quad (4)$$

where $\xi = \beta + \gamma(1 - \alpha - \beta)$ and the expected distortion

$$\langle d(y, z) \rangle = 1 - \beta + \gamma(\beta - \alpha)$$

Let us consider a number of naïve strategies and compute their mutual information and expected distortion.

Case 1: $\alpha = 1, \beta = 0$. In this case, the decision is always $z = \text{low}$. One would then expect that no information about the input is needed, and the mutual information $I(y, z) = 0$ bears this out. The expected distortion is $\langle d(y, z) \rangle = 1 - \gamma$ which is the probability that the external signal is above the threshold.

Case 2: $\alpha = 0, \beta = 1$. This is the complete opposite of the previous case, as the decision is always $z = \text{high}$. Again, this requires no knowledge of y , with $I(y, z) = 0$. Moreover, the expected distortion is $\langle d(y, z) \rangle = \gamma$.

Case 3: $\alpha = 1, \beta = 1$. This represents making the correct decision all the time, and we can check that $\langle d(y, z) \rangle = 0$. This clearly requires that the cell follow y exactly. The mutual information needed for this decision is $I(y, z) = H(Y)$, that is, the mutual information equals all the available information about y , and this quantity is given by its entropy.

Case 4: $\alpha = 0, \beta = 0$. This is the case in which we make the wrong decision every time. Note that $\langle d(y, z) \rangle = 1$, but that the requisite information is also $I(y, z) = H(Y)$.

Case 5: $\alpha = \beta = 1 - \delta, \delta \in (0, 1)$. The previous four cases represent extremes in terms of either the amount of information required or the resultant distortion. This case represents a certain amount of hedging whereby the cell does not commit to either all low or all high responses, irrespective of the input. This strategy leads to $\langle d(y, z) \rangle = \delta$. The mutual information does not simplify much beyond the formula (4) but, when plotted as a function of the expected distortion we see a concave function of D , as shown in Fig. 2A. For $\delta \in (0, 1/2)$, the amount of mutual information is decreasing, reaching zero at $\delta = 1/2$. It is worth emphasizing that this is not an optimal strategy. In fact, we know from Case 1 that we can achieve $\langle d(y, z) \rangle = 1 - \gamma < 1/2$ even when there is no mutual information. We can carry out similar strategies whereby we search for combinations α and β ; see Fig. 2.

Having looked at these simple scenarios, we now ask what the optimal strategy is *without* assuming that the decision is piecewise constant. In particular, we set a maximum allowable distortion $D \in \{0, 1\}$ and require that $\langle d(y, z) \rangle \leq D$.

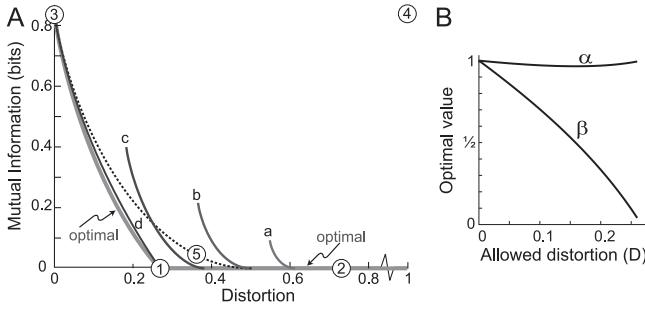


Fig. 2. Mutual information as a function of distortion. A. The four points (labeled, 1–4 inside circles) refer to Cases 1–4. In particular, the pairs 1&2 and 3&4 have the same mutual information, but different achieved distortions. Dashed line 5 maps out Case 5 as δ varies from zero to 1/2. Curves labeled a–d represent different pairs where α is held constant at 0.25, 0.5, 0.75 and 1 (respectively) and β is swept from 0.5 to 1. Also shown is the optimal rate distortion curve as computed by the Blahut-Arimoto algorithm. For these curves, we assumed that $p_Y(y)$ is an exponential function with finite support: $p_Y(y) = e^{-\lambda y} / (1 - e^{-\lambda y_T})$, where $y \in [0, y_T]$. In particular, we set $\lambda = 0.05$, $y_T = 40$ and $y_{th} = 20$; this gives $\gamma \approx 0.731$. B. Optimal values of α and β for the decision (3) as D varies.

Because the penalty is binary, the optimal strategy is also binary, as suggested in (3), where the optimal α and β depend on the allowed distortion D ; see Fig. 2B. Note that, perhaps unexpectedly, the optimal α 's dependence on D is non-monotonic. In this case, the rate distortion function can be computed to be [21]:

$$R(D) = \begin{cases} H(\gamma) - H(D), & D \in [0, \min\{\gamma, 1 - \gamma\}] \\ 0, & D > \min\{\gamma, 1 - \gamma\}. \end{cases}$$

A. The Choice of Distortion Function

So far, in this example we have assumed that there is a hard threshold, but that the response is otherwise binary. This leads to optimal decisions that are themselves binary. In practice, it is rare, if not unprecedented, for cell signaling systems to exhibit such sharp transitions. It is worth asking whether other distortion curves give rise to experimentally-observed stimulus-response maps. To this end, one possibility is to alter the shape of the distortion curve. For example, while it is reasonable to expect that erroneous decisions made when the observed y is far from the origin should be penalized heavily, perhaps those arising for concentrations near the origin should invoke a smaller penalty. With this in mind, we can consider the following graded or fuzzy distortion measure (Fig. 3A):

$$d_1(y, z) = \begin{cases} \frac{y - y_{th}}{y_T - y_{th}} & y < y_{th} \\ 0 & z = \text{low} \\ 1 - \frac{y}{y_{th}} & z = \text{high} \end{cases}$$

We have normalized the two functions that define d_1 so that they are one at the two concentration extremes for y . Moreover, if $y_T = 2y_{th}$ then these two curves are symmetric (Fig. 3A).

The rate distortion curve for this example is shown in Fig. 3B and the optimal stimulus-response strategy for three different values of D are plotted in Fig. 3C. These curves show a gradual transition from low to high decision. Moreover, they

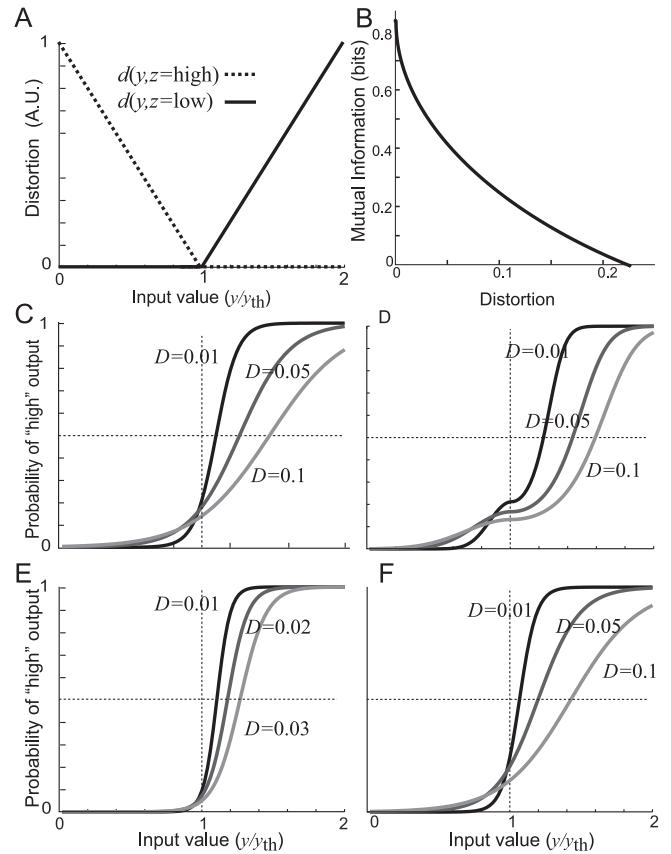


Fig. 3. Optimal stimulus-response curves. A. Graded distortion function. B. Rate-distortion curve for the linear distortion function assuming that the probability distribution is exponential. C. Corresponding optimal stimulus-response curves for three levels of D . D. Optimal response curve assuming a quadratic distortion function. E. Optimal response curve assuming a Gaussian probability distribution with finite support: $p(y) \propto \exp(-(y - y_0)^2/2\sigma)$, $y \in [0, y_T]$. The computation assumes that $y_0 = 3y_{th}/4$, $\sigma = 3y_{th}/2$ and $y_T = 2y_{th}$. F. Optimal response curve for a beta distribution.

are well approximated by the sigmoidal "Hill" function

$$h(y) = \frac{y^n}{y_0^n + y^n}$$

where n is the Hill coefficient and y_0 is the value that makes $h(y) = 1/2$. As D decreases, the Hill exponent n increases, and the 50% point (y_0) approaches y_{th} .

We can repeat this exercise with a quadratic distortion function ($d_2(y, z) = d_1^2(y, z)$) in which case the distortion function is quite similar but the optimal stimulus-response curves are no longer sigmoidal; see Fig. 3D.

B. Dependence on the Assumptions About the Source

In a similar manner, we can ask whether the choice regarding the input distribution affects the optimal response or rate distortion function. To test this, we considered a Gaussian distribution in which the mean and standard deviations are given by $\mu = 3y_{th}/4$ and $\sigma = 3y_{th}/2$, respectively, and y is restricted to $y \in [0, 2y_{th}]$. As seen in Fig. 3E, the resultant optimal stimulus-response curves are also sigmoidal. We repeated this computation assuming that y follows a beta distribution with finite support:

$$p(y) \propto (y/y_T)^{\alpha-1}(1 - y/y_T)^{\beta-1}, \quad y \in [0, 2y_{th}]$$

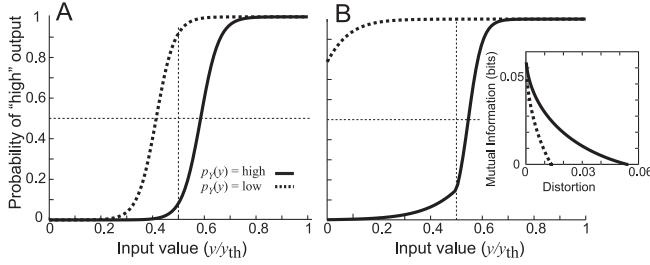


Fig. 4. A. Optimal stimulus-response curves for the linear distortion function assuming that the probability distribution is Gaussian with either low ($\mu = 3y_{\text{th}}/4$; solid line) or high ($\mu = 5y_{\text{th}}/4$; dotted line) mean, assuming that $D = 0.03$. The other parameters are as in Fig. 3D. B. Optimal stimulus-response curves assuming non symmetric distortion function with $D = 0.0134$ and $p(y)$ as in panel A. The inset shows the rate distortion function $R(D)$ for this case.

The computation used $\alpha = 2$, $\beta = 3$. Once again, the shape is qualitatively quite similar.

All the calculations above have assumed that $y < 1/2$; that is, there is a higher probability that the sensed signal would be below the threshold, so that a low response would be more typical. Reversing this assumption leads to a different optimal response curve; see Fig. 4A. There we plot the rate distortion curves and optimal stimulus-response curves using the linear distortion function $d_1(y, z)$ assuming that y is Gaussian with finite support, in which the mean is either below or above y_{th} . Note how the two optimal stimulus-response curves share the same shape, but differ as to where they cross from low to high.

This shift represents a change in the optimal response of the cell based on varying assumptions. Suppose, for example, that a cell has been conditioned to an environment with low signaling, in which $y < y_{\text{th}}$. In this case, the cell is operating at a low response. By setting an effective threshold above y_{th} it ensures that the sensed signal is significantly above the threshold before commencing a possibly metabolically expensive response. On the other hand, suppose that the environment has changed so that the default environment is one in which $y > y_{\text{th}}$. When the cell is operating in this high regime, it may want to recondition the optimal response so that it is only a sensed signal significantly below the threshold that triggers a return to the low state. This set of optimal strategies invoking varying effective thresholds is equivalent to a hysteretic signaling system [33] which makes the system insensitive to small perturbations.

C. The Use of an Asymmetric Distortion Function

So far we have assumed that the distortion function is symmetric, suggesting that there is an equal penalty for an error in either direction. This may not always be a reasonable assumption. For example, apoptosis marks a literally life-or-death cell decision in which cells, in response to environmental cues, commence a series of irreversible steps that lead to cell death.

To see how asymmetric distortion functions can alter the change in the shape of the response, we defined the function

$$d(x, y) = \begin{cases} \frac{1}{2}(1 - \cos(\theta_y - \theta_z)) & \text{if } y < y_{\text{th}} \\ 0 & \text{if } y \geq y_{\text{th}} \end{cases}$$

The resultant rate distortion curves and optimal stimulus-response maps are shown in Fig. 4B under the assumptions that the typical concentration is either below (solid lines) or above (dotted lines) the threshold. Note that the relatively high penalty of getting the $z = \text{low}$ decision wrong when $y \geq y_{\text{th}}$ makes this decision curve steep when the expected input is low. On the other hand, the relatively low penalty in the reverse situation virtually abolishes the transition from high to low. This makes the transition from low to high a nearly irreversible decision.

IV. EXAMPLE: GRADED SIGNALING SYSTEMS

We now consider a situation where the decision can take on multiple values, the directed motion of cells, such as the amoebae *Dictyostelium discoideum*. *Dictyostelium* cells typically live in the soil as independent, unicellular organisms feeding on bacteria. When faced with adverse environmental conditions, such as starvation due to a depletion in their available food source, they begin a developmental process whereby they become a multicellular organism. Approximately 4–8 hours after commencing development, up to 100,000 cells signal each other by synthesizing and releasing a chemical into the environment. Cells use spatial differences in the concentration of this to move towards each other and aggregate, eventually forming a spore. *Chemotaxis*, the ability to sense the direction of external chemical sources and respond by polarizing and migrating toward chemoattractants is crucial for the survival of *Dictyostelium* and is evolutionarily conserved, as many of the same biochemical pathways are used by cells of other organisms, including humans.

These cells have receptors evenly spaced throughout the plasma membrane. The receptors bind chemoattractant molecules and it is this level of receptor occupancy that is used to guide the motion of cells directionally [34]. Moreover, there is evidence that this occupancy signal undergoes stochastic fluctuations [35], [36]. A number of papers have shown how this noise limits the ability of the cell to infer the direction of the chemical source, both by calculating the expected level of the fluctuations [37], [38], the types of receptors [39] or the level of information loss at various points along the signaling cascade [40]. Here, we consider the response in terms of an optimal rate distortion problem — how much information about the external signal must the cell have in order to move directionally with a predetermined level of accuracy [30].

To consider this response in the framework of a rate distortion problem, we define $\theta_y \in [-\pi, \pi]$ as the directional angle sensed by the cells, and $\theta_z \in [-\pi, \pi]$ as their directional response. A natural notion of distortion in this case is the function

$$d(\theta_y, \theta_z) = \frac{1}{2}(1 - \cos(\theta_y - \theta_z))$$

This distortion is related to the common experimentally measured chemotactic index (CI) that is used to quantify chemotactic accuracy:

$$\text{CI} = \cos(\theta_y - \theta_z).$$

In our case, we have normalized CI in such a way so that there is no distortion when two angles coincide ($\text{CI} = 1$), and equals one when they differ by 180° ($\text{CI} = -1$).

Note that because the input and output signals are now defined in a continuum, we should use continuous versions of the formulae defined in Section II. Rather than doing this, we discretize the circular domain into N equal sectors where, in the plots that follow, we use $N = 72$. This is done for computational convenience, but also accounts for the fact that during movement, the protrusions of cells, known as *pseudopods*, are usually sufficiently large so as to cover 5–10% of the cell perimeter.

A. Cells With No a Priori Directional Expectation

We first assume that θ_y is uniformly distributed, representative of a cell with no *a priori* directional bias. In this case, the rate distortion function is shown by the solid line in Fig. 5A. It shows that if we allow $\langle d(\theta_y, \theta_z) \rangle \geq \frac{1}{2}$, then the cell requires no information regarding the external environmental cue. In this case, the optimal stimulus-response map is also uniformly distributed around the perimeter signifying random motion without any directional bias; see light gray line in Fig. 5B.

Increasing the fidelity requirement from $D = 0.5$ leads to an increase in the mutual information needed, so that, at $D = 0.2$ and $D = 0.1$, approximately 0.57 and 1.17 bits are required, respectively. A fidelity of $D = 0.1$ amounts to an expected chemotactic index of 0.8, which is similar to those measured experimentally in 10% chemoattractant gradients [41].

We can also compare the predicted stimulus-response curve to the experimentally-measured response of cells in chemoattractant gradients of immobilized cells [42], [43]. These show that when cells are stimulated by a constant chemoattractant gradient they redistribute internal markers with a spatial distribution that is steeper (3.1 ± 0.89 for *Dictyostelium* cells and 3.25 ± 2.0 for human neutrophils) than that of the external signal. That is, if the stimulus concentration varies by $\pm 10\%$ between front and back, the intracellular concentration will vary by $\pm 30\%$. The optimal response for a distortion of $D \approx 0.15$ amplifies the signal by a factor of ≈ 3.17 steepness, close to that of the experimental measurements [30].

B. Cells With a Priori Directional Expectation

We now assume that θ_y has a preferred direction. This can represent biased cells that expect gradients to come predominantly from a predetermined direction. In the context of *Dictyostelium* chemotaxis, these could be cells that are several hours into their aggregation. These cells have been moving in a constant direction for some time and hence might expect to continue moving in this direction. To account for this *a priori* information about the direction of the gradient, we assume that the angle of the source is defined by a von Mises distribution:

$$p_{\Theta_y}(\theta_y) = \frac{\exp(\kappa \cos(\theta_y - \mu))}{2\pi I_0(\kappa)}$$

where $I_0(\kappa)$ is the modified Bessel function of order zero. For this distribution the variance is monotonically decreasing

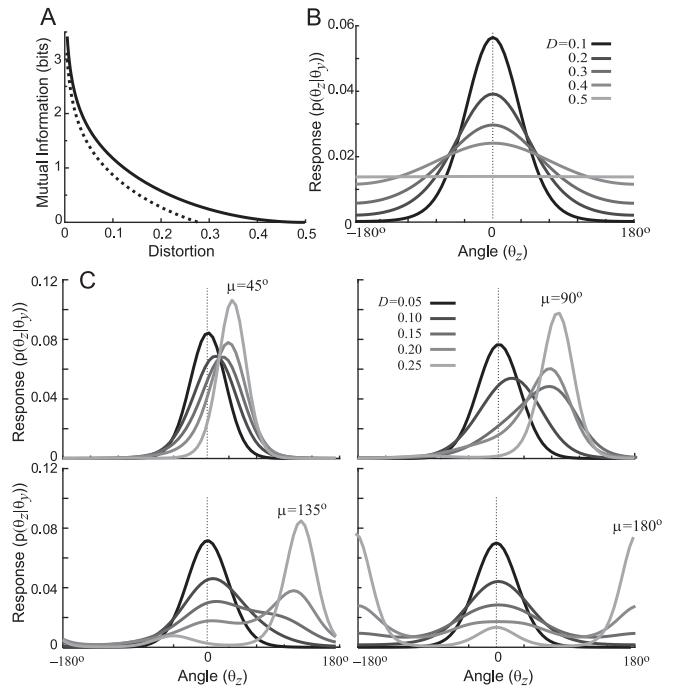


Fig. 5. Graded signaling systems. A. Rate distortion curve assuming that the external signal is uniformly distributed (solid line) or follows a von Mises distribution with $\kappa = 1$ and $\mu = 45^\circ$ (dotted line). B. Optimal stimulus-response curves assuming uniformly distributed external signal for various fidelity requirements. We assume that the sensed signal is $\theta_y = 0$ so that the “correct” response is to move with $\theta_z = 0$. C. Optimal stimulus-response curves assuming biased external signal for various fidelity levels. The expected direction is marked in each subplot, and note how the peak in the distribution for low fidelity levels ($D = 0.25$) points in these directions. As D is lowered, the peak shifts so that eventually it points in the direction of the external signal, assumed to be $\theta_y = 0$.

function of κ (equal to $1 - I_1(\kappa)/I_0(\kappa)$). We set $\kappa = 1$ which corresponds to a standard deviation approximately $\sigma_{\theta_y} \approx 43^\circ$.

The rate distortion function for this case is shown by the dotted line in Fig. 5A. Note that the *a priori* information allows the cell to require less information about the source to achieve the same level of fidelity. In all cases, a sufficiently lax fidelity requirement, for example, $D = 0.3$ can be met with no information about the direction of the external chemoattractant. This can be achieved by a cell that moves solely following the direction of the internal bias. For higher fidelity requirements (lower D), the optimal response depends on both the angles θ_y and θ_z ; we plot these in Fig. 5C under various choices for $\theta_y = 0$ and D . Note that for relatively low fidelity ($D = 0.25$) the best response is for the cell to follow mostly the expected direction (μ). As we increase the required fidelity, the optimal response shifts from the *a priori* direction to the sensed direction so that, at $D = 0.05$, the peak in $p_{\theta_z|\theta_y}$ occurs towards the external angle. When the external and internal angles agree or are close, the function $p_{\theta_z|\theta_y}$ is steeper than if there were no external information. This increase likely represents the fact that two sources of information (internal and external) are guiding the cell in one direction. On the other hand, when the internal and external angles differ significantly, there is either a bimodal optimal $p_{\theta_z|\theta_y}$ or one that peaks at an angle somewhere between the two.

V. CONCLUSION

The two examples that we have presented here consider two commonly-seen scenarios in biology. Moreover, our results show that many of the observed responses seen in cells arise as an optimal response assuming reasonable distortion functions. For example, sigmoidal Hill-type functions are ubiquitous throughout biology. Our results show that they are the natural consequence of a binary decision that does not use a strict binary distortion function, but that instead allows for fuzzy or graded penalties. Moreover, we see that by changing the assumption regarding the typical or expected input, the shape of the optimal-response curve remains unaltered, but the 50% point on the Hill function changes. Such hysteretic behavior is seen throughout biology [33] and represents a means by which cells adapt to their environment. Finally, we also see how irreversible curves also arise as the optimal response to asymmetric distortion functions.

In the case of chemoattractant gradient sensing, we also found strong agreement between the optimal responses predicted by the theory and experimentally-observed cell function. In particular, we see how different assumptions regarding the *a priori* information about the external gradient has an effect on the optimal response. *Dictyostelium* cells that are early into their developmental response to starvation are finding the direction of the external signal. As such, it is beneficial to assume no prior information. Experimentally, these cells are said to be *unpolarized*, possessing no preferred internal direction. In contrast, after finding the direction of the source, *Dictyostelium* cells continue migrating in this direction for several hours during which the direction of the gradient will not change. During this movement, the cells polarize, developing well defined anterior and posterior regions. Moreover, there is evidence that cells can remember the direction of movement in this polarized state [44]–[46]. Experiments on these cells show that they integrate the remembered, internal state of the cell and the external chemoattractant gradient in their direction [47], similar to our results from Fig. 5C. When the two angles agree, the response of these polarized cells show an increase in the amplification of the external signal, greater than that seen in unpolarized cells [42], [48], once again in agreement with the optimal response predicted by our model.

It is worth noting that nearly all the experimentally-measured values of channel capacity amount to a mutual information of about one bit. This is clearly enough for the binary decision processes considered here. In the context of chemotaxis, a rate of one bit amounts to approximately $D = 0.1$, which leads to a CI equal to 0.8 or, approximately being off by 35° . However, this represents the average distortion. Over time, we see cells compensate by including further *a priori* information in their decision. In this case, the same one bit of channel capacity can result in lowering the expected distortion.

Rate distortion theory was developed by Shannon to consider real-world situations in which distortion-free information transmission is required or a desirable. The rate distortion function provides a quantifiable trade-off between precision or fidelity and cost. In the traditional communication engineering

setting, this cost represents the information-carrying capacity of the channel. Lowering the fidelity requirement thereby saves resources. Earlier, we questioned whether mutual information was an appropriate cost for biological signaling pathways, or whether other “currencies,” such as energy [49], are more appropriate. We have seen that easing fidelity requirements in the situations depicted here result in a lower (mutual information) cost to the cell, but also result in less sharp stimulus-response curves (lower Hill coefficients). Biochemical circuits with high Hill coefficient, which lead to lower expected distortions, achieve these higher exponents through cooperative interactions. This involves having multiple molecules of the same species form multi-protein complexes, for which the metabolic cost is higher [50], [51].

A little known fact about Shannon [52] is that his 1940 MIT doctoral thesis addressed problems in population genetics. Now, as we celebrate his 100th anniversary, there is little doubt that he would have been pleased to see his most celebrated contribution, information theory, applied to the study of biology. We hope that this article will help spur interest in the use of rate distortion theory to analyze various signaling systems.

ACKNOWLEDGMENT

The author would like to thank Debojyoti Biswas and Sayak Bhattacharya for their helpful comments.

REFERENCES

- [1] M. Pan, X. Xu, Y. Chen, and T. Jin, “Identification of a chemoattractant G-protein-coupled receptor for folic acid that controls both chemotaxis and phagocytosis,” *Dev. Cell*, vol. 36, no. 4, pp. 428–439, Feb. 2016.
- [2] S. Paliwal *et al.*, “MAPK-mediated bimodal gene expression and adaptive gradient sensing in yeast,” *Nature*, vol. 446, no. 7131, pp. 46–51, Mar. 2007.
- [3] D. Jukam and C. Desplan, “Binary fate decisions in differentiating neurons,” *Current Opin. Neurobiol.*, vol. 20, no. 1, pp. 6–13, Feb. 2010.
- [4] C. V. Rao, D. M. Wolf, and A. P. Arkin, “Control, exploitation and tolerance of intracellular noise,” *Nature*, vol. 420, no. 6912, pp. 231–237, Nov. 2002.
- [5] Y. Xiong, C.-H. Huang, P. A. Iglesias, and P. N. Devreotes, “Cells navigate with a local-excitation, global-inhibition-biased excitable network,” *Proc. Nat. Acad. Sci. USA*, vol. 107, no. 40, pp. 17079–17086, Oct. 2010.
- [6] I. Hecht, D. A. Kessler, and H. Levine, “Transient localized patterns in noise-driven reaction-diffusion systems,” *Phys. Rev. Lett.*, vol. 104, no. 15, Apr. 2010, Art. no. 158301.
- [7] H. H. McAdams and A. Arkin, “Stochastic mechanisms in gene expression,” *Proc. Nat. Acad. Sci. USA*, vol. 94, no. 3, pp. 814–819, Feb. 1997.
- [8] J. Paulsson, “Summing up the noise in gene networks,” *Nature*, vol. 427, no. 6973, pp. 415–418, Jan. 2004.
- [9] D. T. Gillespie, “Stochastic simulation of chemical kinetics,” *Annu. Rev. Phys. Chem.*, vol. 58, pp. 35–55, May 2007.
- [10] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, “Stochastic gene expression in a single cell,” *Science*, vol. 297, no. 5584, pp. 1183–1186, Aug. 2002.
- [11] B. W. Andrews, T.-M. Yi, and P. A. Iglesias, “Optimal noise filtering in the chemotactic response of *Escherichia coli*,” *PLoS Comput. Biol.*, vol. 2, no. 11, Nov. 2006, Art. no. e154.
- [12] C.-S. Chou, L. Bardwell, Q. Nie, and T.-M. Yi, “Noise filtering tradeoffs in spatial gradient sensing and cell polarization response,” *BMC Syst. Biol.*, vol. 5, p. 196, Dec. 2011.
- [13] C. Zechner, G. Seelig, M. Rullan, and M. Khammash, “Molecular circuits for dynamic noise filtering,” *Proc. Nat. Acad. Sci. USA*, vol. 113, no. 17, pp. 4729–4734, Apr. 2016.
- [14] T. J. Perkins and P. S. Swain, “Strategies for cellular decision-making,” *Mol. Syst. Biol.*, vol. 5, Nov. 2009, Art. no. 326.

- [15] I. S. Mian and C. Rose, "Communication theory and multicellular biology," *Integr. Biol. (Camb)*, vol. 3, no. 4, pp. 350–367, Apr. 2011.
- [16] G. Tkačik and A. M. Walczak, "Information transmission in genetic regulatory networks: A review," *J. Phys. Condensed Matter*, vol. 23, no. 15, Apr. 2011, Art. no. 153102.
- [17] A. Rhee, R. Cheong, and A. Levchenko, "The application of information theory to biochemical signaling systems," *Phys. Biol.*, vol. 9, no. 4, Aug. 2012, Art. no. 045011.
- [18] Z. Mousavian, K. Kavousi, and A. Masoudi-Nejad, "Information theory in systems biology. Part I: Gene regulatory and metabolic networks," *Seminars Cell Develop. Biol.*, vol. 51, pp. 3–13, Mar. 2016.
- [19] Z. Mousavian, J. Díaz, and A. Masoudi-Nejad, "Information theory in systems biology. Part II: Protein–protein interaction and signaling networks," *Seminars Cell Develop. Biol.*, vol. 51, pp. 14–23, Mar. 2016.
- [20] S. Uda and S. Kuroda, "Analysis of cellular signal transduction from an information theoretic approach," *Seminars Cell Develop. Biol.*, vol. 51, pp. 24–31, Mar. 2016.
- [21] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 2006.
- [22] G. Tkačik, C. G. Callan, Jr., and W. Bialek, "Information capacity of genetic regulatory elements," *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.*, vol. 78, no. 1, Jul. 2008, Art. no. 011910.
- [23] R. Cheong, A. Rhee, C. J. Wang, I. Nemenman, and A. Levchenko, "Information transduction capacity of noisy biochemical signaling networks," *Science*, vol. 334, no. 6054, pp. 354–358, Oct. 2011.
- [24] S. Uda *et al.*, "Robustness and compensation of information transmission of signaling pathways," *Science*, vol. 341, no. 6145, pp. 558–561, Aug. 2013.
- [25] J. Selimkhanov *et al.*, "Accurate information transmission through dynamic biochemical signaling networks," *Science*, vol. 346, no. 6215, pp. 1370–1373, Dec. 2014.
- [26] A. S. Hansen and E. K. O’Shea, "Limits on information transduction through amplitude and frequency regulation of transcription factor activity," *Elife*, vol. 4, May 2015, Art. no. e06559.
- [27] M. Voliotis, R. M. Perrett, C. McWilliams, C. A. McArdle, and C. G. Bowsher, "Information transfer by leaky, heterogeneous, protein kinase signaling systems," *Proc. Nat. Acad. Sci. USA*, vol. 111, no. 3, pp. E326–E333, Jan. 2014.
- [28] K. L. Garner *et al.*, "Information transfer in gonadotropin-releasing hormone (GnRH) signaling: Extracellular signal-regulated kinase (ERK)-mediated feedback loops control hormone sensing," *J. Biol. Chem.*, vol. 291, no. 5, pp. 2246–2259, Jan. 2016.
- [29] P. Mehta, S. Goyal, T. Long, B. L. Bassler, and N. S. Wingreen, "Information processing and signal integration in bacterial quorum sensing," *Mol. Syst. Biol.*, vol. 5, no. 1, p. 325, 2009.
- [30] B. W. Andrews and P. A. Iglesias, "An information-theoretic characterization of the optimal gradient sensing response of cells," *PLoS Comput. Biol.*, vol. 3, no. 8, Aug. 2007, Art. no. e153.
- [31] J. R. Porter, B. W. Andrews, and P. A. Iglesias, "A framework for designing and analyzing binary decision-making strategies in cellular systems," *Integr. Biol. (Camb)*, vol. 4, no. 3, pp. 310–317, Mar. 2012.
- [32] I. Gkigkitzis, "Theoretical aspects and modelling of cellular decision making, cell killing and information-processing in photodynamic therapy of cancer," *BMC Med. Genomics*, vol. 6, Dec. 2013, Art. no. S3.
- [33] J. E. Ferrell, Jr., "Self-perpetuating states in signal transduction: Positive feedback, double-negative feedback and bistability," *Current Opin. Cell Biol.*, vol. 14, no. 2, pp. 140–148, Apr. 2002.
- [34] M. Ueda, Y. Sako, T. Tanaka, P. Devreotes, and T. Yanagida, "Single-molecule analysis of chemotactic signaling in *Dictyostelium* cells," *Science*, vol. 294, no. 5543, pp. 864–867, Oct. 2001.
- [35] Y. Miyanaga, S. Matsuoka, T. Yanagida, and M. Ueda, "Stochastic signal inputs for chemotactic response in *Dictyostelium* cells revealed by single molecule imaging techniques," *Biosystems*, vol. 88, no. 3, pp. 251–260, Apr. 2007.
- [36] M. Ueda and T. Shibata, "Stochastic signal processing and transduction in chemotactic response of eukaryotic cells," *Biophys. J.*, vol. 93, no. 1, pp. 11–20, Jul. 2007.
- [37] J. M. Kimmel, R. M. Salter, and P. J. Thomas, "An information theoretic framework for eukaryotic gradient sensing," in *Advances in Neural Information Processing Systems 19*, B. Schölkopf, J. C. Platt, and T. Hoffman, Eds. Cambridge, MA, USA: MIT Press, 2007, pp. 705–712.
- [38] W.-J. Rappel and H. Levine, "Receptor noise limitations on chemotactic sensing," *Proc. Nat. Acad. Sci. USA*, vol. 105, no. 49, pp. 19270–19275, Dec. 2008.
- [39] R. G. Endres and N. S. Wingreen, "Accuracy of direct gradient sensing by single cells," *Proc. Nat. Acad. Sci. USA*, vol. 105, no. 41, pp. 15749–15754, Oct. 2008.
- [40] B. Hu, D. A. Kessler, W.-J. Rappel, and H. Levine, "Effects of input noise on a simple biochemical switch," *Phys. Rev. Lett.*, vol. 107, no. 14, Sep. 2011, Art. no. 148101.
- [41] L. Bosgraaf and P. J. M. Van Haastert, "Navigation of chemotactic cells by parallel signaling to pseudopod persistence and orientation," *PLoS One*, vol. 4, no. 8, 2009, Art. no. e6842.
- [42] C. Janetopoulos, L. Ma, P. N. Devreotes, and P. A. Iglesias, "Chemoattractant-induced phosphatidylinositol 3,4,5-trisphosphate accumulation is spatially amplified and adapts, independent of the actin cytoskeleton," *Proc. Nat. Acad. Sci. USA*, vol. 101, no. 24, pp. 8951–8956, Jun. 2004.
- [43] M. D. Onsum, K. Wong, P. Herzmark, H. R. Bourne, and A. P. Arkin, "Morphology matters in immune cell chemotaxis: Membrane asymmetry affects amplification," *Phys. Biol.*, vol. 3, no. 3, pp. 190–199, Sep. 2006.
- [44] A. Samadani, J. Metetal, and A. van Oudenaarden, "Cellular asymmetry and individuality in directional sensing," *Proc. Nat. Acad. Sci. USA*, vol. 103, no. 31, pp. 11549–11554, Aug. 2006.
- [45] A. Nakajima, S. Ishihara, D. Imoto, and S. Sawai, "Rectified directional sensing in long-range cell migration," *Nat. Commun.*, vol. 5, Nov. 2014, Art. no. 5367.
- [46] M. Skoge *et al.*, "Cellular memory in eukaryotic chemotaxis," *Proc. Nat. Acad. Sci. USA*, vol. 111, no. 40, pp. 14448–14453, Oct. 2014.
- [47] M. J. Wang, Y. Artemenko, W.-J. Cai, P. A. Iglesias, and P. N. Devreotes, "The directional response of chemotactic cells depends on a balance between cytoskeletal architecture and the external gradient," *Cell Rep.*, vol. 9, no. 3, pp. 1110–1121, Nov. 2014.
- [48] C. Shi, C.-H. Huang, P. N. Devreotes, and P. A. Iglesias, "Interaction of motility, directional sensing, and polarity modules recreates the behaviors of chemotaxing cells," *PLoS Comput. Biol.*, vol. 9, no. 7, 2013, Art. no. e1003122.
- [49] R. A. Gatenby and B. R. Frieden, "Information theory in living systems, methods, applications, and challenges," *Bull. Math. Biol.*, vol. 69, no. 2, pp. 635–657, Feb. 2007.
- [50] H. Qian, "Thermodynamic and kinetic analysis of sensitivity amplification in biological signal transduction," *Biophys. Chem.*, vol. 105, nos. 2–3, pp. 585–593, Sep. 2003.
- [51] P. Mehta and D. J. Schwab, "Energetic costs of cellular computation," *Proc. Nat. Acad. Sci. USA*, vol. 109, no. 44, pp. 17978–17982, Oct. 2012.
- [52] C. E. Shannon, "An algebra for theoretical genetics," Ph.D. dissertation, Dept. Math., Massachusetts Inst. Technol., Cambridge, MA, USA, 1940.



Pablo A. Iglesias was born in Caracas, Venezuela, in 1964. He received the B.A.Sc. degree in engineering science from the University of Toronto in 1987, and the Ph.D. degree in control engineering from Cambridge University in 1991. Since then he has been on the faculty of the Johns Hopkins University, where he is currently the Edward J. Schaefer Professor of electrical engineering. He also holds appointments with the Department of Biomedical Engineering and Department of Applied Mathematics and Statistics. He has had visiting appointments with Lund University (Automatic Control), the Weizmann Institute of Science (Mathematics), the California Institute of Technology (Control and Dynamical Systems), the Johns Hopkins School of Medicine (Cell Biology), and the Max-Planck Institute for the Physics of Complex Systems, Dresden. He has authored over 100 research articles, and two books entitled *Minimum Entropy Control for Time-Varying Systems* (Birkhäuser) and *Control Theory and Systems Biology* (MIT Press). He was a recipient of a number of awards for research such as the Charles E. Ives Best Paper Award for the *Journal of Imaging Technology* and for teaching such as the George E. Owen Teaching Award at Johns Hopkins University. He was also named a Distinguished Lecturer by the IEEE Control Systems Society. His research focuses on the use of control and information theory to study biological signal transduction pathways, understanding how cells interpret directional cues to guide cell motion, the regulatory mechanisms that control cell division, and the sensing and actuation that enable cells to maintain lipid homeostasis.