

# Information-theoretic Analysis of Signal Transduction Pathways

Burton W. Andrews, Joshua R. Porter, Pablo A. Iglesias\*

Johns Hopkins University, Baltimore, MD, US

\*E-mail: pi@jhu.edu

## Introduction

The proper functioning of many cellular systems requires accurate decision-making in the presence of limited information about the environment. Random effects that are inherent in the cellular world such as diffusion and stochastic reaction dynamics disrupt signals of interest and make errorless response mechanisms impossible. Here we introduce rate distortion, a branch of information theory from engineering, as a tool to study the ability of cellular decision-making systems to respond to random stimuli. This approach quantifies the cost associated with improved performance in terms of mutual information and identifies the mechanisms that cells use to balance this tradeoff.

We demonstrate the utility of rate distortion theory in binary decision-making systems such as the pheromone response pathway of yeast. Random fluctuations in pheromone concentration hinder the ability of these cells to accurately make cell-fate decisions. Using rate distortion theory, we show that, depending on the prior knowledge that a cell has about its environment, different optimal response strategies exist. Furthermore, we show that experimentally observed characteristics such as high sensitivity and bistability correspond to these various optimally efficient decision-making processes.

## Setup

Consider a single cell that must decide between two different cell fates based on measurements of some external chemical cue. In the case of the yeast *Saccharomyces cerevisiae*, the cue is a pheromone, and the cell must decide whether to continue normal cell growth or, if the pheromone is large enough, to induce cell cycle arrest – a process necessary for mating. We thus consider pheromone concentration ( $X$ ) to be the input to the cell-fate decision system and the output ( $Z$ ) to be the decision of arrest ( $Z = 1$ ) or not arrest ( $Z = 0$ ). The signal transduction pathway that implements this decision is modeled as the conditional probability distribution  $p_{Z|X}(z|x)$ .

We characterize the quality of the cell's decision with the distortion function

$$d(x, z) = \begin{cases} 0 & \text{if } z = h(x) \\ 1 & \text{if } z \neq h(x) \end{cases} \quad (1)$$

where  $h(x)$  indicates whether the pheromone concentration exceeds some desired critical amount  $x_c^*$ , and  $d(x, z)$  denotes the accuracy of the decision  $z$  with respect to this indicator. The distortion is zero if the decision is correct and one if incorrect. The desired critical concentration  $x_c^*$  represents a pheromone concentration large enough to warrant the induction of mating behavior (e.g., if a mating partner is nearby).

## Optimal response

Pheromone concentration is assumed to be distributed exponentially:  $p_X(x) = (e^{-\lambda x}) / \int e^{-\lambda x}$ , so that the cell is most likely to experience zero pheromone concentration, and the probability of experiencing greater concentrations decreases with  $x$  at a rate  $\lambda$ . Using this source distribution and  $d(x, z)$  given above, we computed the rate distortion function, which gives the minimum amount of mutual information ( $I(X; Z)$ ) about the pheromone concentration that the signaling network ( $p_{Z|X}(z|x)$ ) must obtain in order to successfully make cell-fate decisions ( $E[d] \leq D$ ). The rate distortion function is a decreasing function of  $D$  (Fig. 1A); less information about the input is required to achieve a larger or more relaxed distortion constraint.

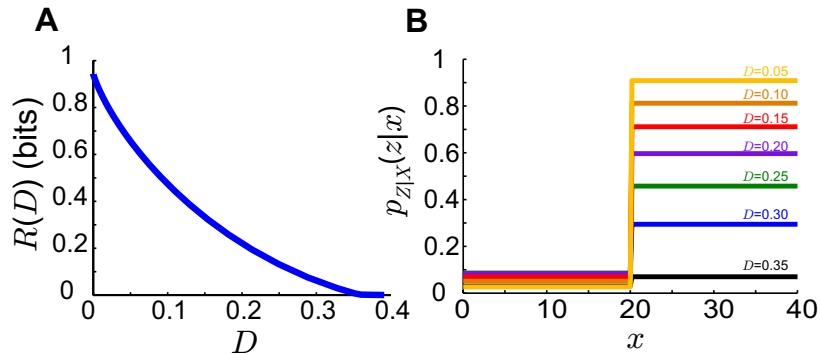


Figure 1: Rate distortion function and optimal response.

We next computed the optimal  $p_Z(z|x)$  that achieves  $R(D)$  for varying  $D$  (Figure 1B). The  $R(D)$ -achieving distribution is a switch function that assumes one of two output probabilities depending on the observed pheromone concentration relative to the critical concentration. This optimal decision function is not a perfect switch, however, because the probabilities of the lower and upper plateaus are not zero and one. As the allowable distortion decreases, this optimal map approaches that of a perfect switch.

## Comparison to experimental data

Experimental evidence suggests that the probability of a particular response or cell fate of *S. cerevisiae* is highly sensitive to the pheromone input concentration. The average response is not to arrest for pheromone concentrations much less than some critical concentration and to always arrest for concentrations much greater than the critical. Both types of response behavior are observed at pheromone concentrations close to the critical so that the overall response has the characteristic of a switch, much like the optimal map that achieves the rate distortion function.

To assess the performance of signaling networks in real cells further, we use a simple model of pheromone response based on experimental data and previously published models. The model incorporates stochastic pheromone-receptor binding and cooperative amplification via a Hill coefficient  $h$ . We computed the expected distortion  $E[d(X, Z)]$  and mutual information  $I(X; Z)$  for this model with varying Hill coefficients  $h$ . We find that, as  $h$  increases,

$I(X; Z)$  increases and  $E[d(X, Z)]$  decreases, implying that steeper Hill coefficients produce better decision functions. However this improved performance comes at the cost of mutual information as well as any energetic costs associated with amplification. A perfect switch ( $h = \infty$ ) gives the most information and the least amount of distortion.

## Bistability

The above analysis suggests that the high sensitivity found in the switch-like response of the signaling pathway of *S. cerevisiae* is nearly optimal in the context of rate distortion. However, some models and experimental data indicate that the response is not a simple monostable switch, but rather has a bistable characteristic. We next show how the above rate distortion theory analysis may be extended to explain this bistability as an optimal response to varying environmental conditions.

Consider the following the distortion function:

$$d(x, z) = \begin{cases} 0 & \text{if } z = h(x) \\ \min\{1, m|x - x_c^*|\} & \text{if } z \neq h(x). \end{cases} \quad (2)$$

This distortion function resembles the previous one with the exception that there is a gradual increase in distortion around the switch point as opposed to a discontinuity. This less-abrupt transition allows some tolerance for mistakes that are made when the pheromone concentration is close to  $x_c^*$ . When  $m = \infty$ , Equation (1) is recovered.

Computation of  $R(D)$  and the corresponding optimal map for the exponential source distribution with both positive and negative  $\lambda$  reveals a bistable optimal response (Figure 2). The  $\lambda > 0$  case is well representative of a cell with some *a priori* knowledge that the pheromone concentration is, on average, at low levels. An exponential source distribution with  $\lambda < 0$ , on the other hand, is representative of a cell's *a priori* knowledge of the pheromone existing at relatively high concentrations and has an  $R(D)$  that lies below that for  $\lambda > 0$ .

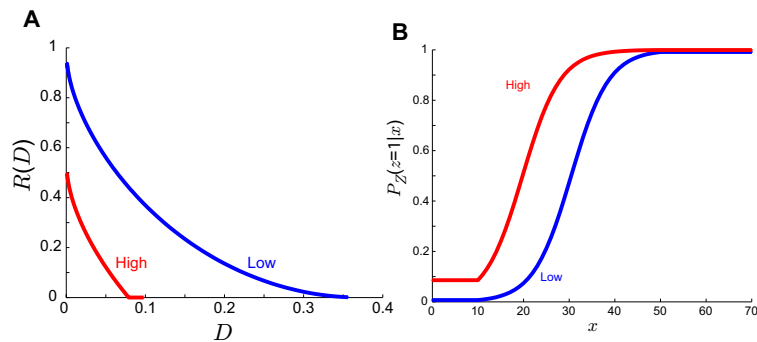


Figure 2:  $R(D)$  and optimal, bistable response.

These results closely resemble the bistable pheromone response observed experimentally in *S. cerevisiae* cells, and they suggest that the cellular response should reflect *a priori* knowledge of the environment. Biasing of the response based on an *a priori* presumption of the input is reminiscent of the different responses in unpolarized and polarized *D. discoideum* cells observed in previous work.