

Receptor crosstalk improves concentration sensing of multiple ligandsMartín Carballo-Pacheco,^{1,*} Jonathan Desponds,^{2,*} Tatyana Gavrilchenko,^{3,*} Andreas Mayer,^{4,*} Roshan Prizak,^{5,*} Gautam Reddy,^{2,*} Ilya Nemenman,^{6,†} and Thierry Mora^{7,†}¹*School of Physics and Astronomy, University of Edinburgh, Edinburgh, EH9 3FD, United Kingdom*²*Department of Physics, University of California San Diego, La Jolla, CA 92093, USA*³*Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA*⁴*Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA*⁵*Institute of Science and Technology Austria, Am Campus 1, A-3400, Klosterneuburg, Austria*⁶*Department of Physics, Department of Biology, and Initiative in Theory and Modeling of Living Systems, Emory University, Atlanta, GA 30322, USA*⁷*Laboratoire de physique de l'École normale supérieure (PSL university), CNRS, Sorbonne University, and University Paris-Diderot, 75005 Paris, France*

(Received 10 October 2018; published 26 February 2019)

Cells need to reliably sense external ligand concentrations to achieve various biological functions such as chemotaxis or signaling. The molecular recognition of ligands by surface receptors is degenerate in many systems, leading to crosstalk between ligand-receptor pairs. Crosstalk is often thought of as a deviation from optimal specific recognition, as the binding of noncognate ligands can interfere with the detection of the receptor's cognate ligand, possibly leading to a false triggering of a downstream signaling pathway. Here we quantify the optimal precision of sensing the concentrations of multiple ligands by a collection of promiscuous receptors. We demonstrate that crosstalk can improve precision in concentration sensing and discrimination tasks. To achieve superior precision, the additional information about ligand concentrations contained in short binding events of the noncognate ligand should be exploited. We present a proofreading scheme to realize an approximate estimation of multiple ligand concentrations that reaches a precision close to the derived optimal bounds. Our results help rationalize the observed ubiquity of receptor crosstalk in molecular sensing.

DOI: [10.1103/PhysRevE.99.022423](https://doi.org/10.1103/PhysRevE.99.022423)**I. INTRODUCTION**

Living cells need to collect information with high precision to respond and adapt to their environment [1]. For example, chemotactic swimming bacteria can react to changes in concentrations of nutrients and toxins [2], and cells from the innate immune system can recognize distinct microbial components and initiate immune responses [3]. Presence or concentrations of key ligands are measured via receptor proteins, which are usually located in the cell surface, and later processed by complex downstream signaling networks to trigger cellular responses. The accuracy of these measurements suffer from multiple sources of noise, including the transport of ligands by diffusion, the binding of the ligands to the receptors after they have arrived to the surface, and the communication between components of the signaling network.

In recent years, the fundamental limits to cellular sensing has received thorough theoretical consideration [4–7]. Berg and Purcell [8] were the first to study the problem and found that the sensing error can be minimized by increasing the number of receptors or the number of measurements per receptor. More recently, Endres and Wingreen showed [9],

using a maximum likelihood estimation, that the accuracy can be increased by a factor of 2 by solely taking into account the unoccupied time intervals. Further theoretical work has concentrated on understanding the limits in cellular sensing for single receptors with spatial [10–14] and temporal gradients [15,16], for multiple receptors [17–19], and even for cells that can communicate [20–22]. The thermodynamic cost [23–27] and the trade-offs between different resources for sensing [28–30] has also been explored at large.

Most of the aforementioned models assume that receptors sense individual ligands. However, cognate ligands usually reside among other spurious ligands, and receptors must identify them accurately [16,31–34]. Recently, Mora [35] derived the fundamental limit to measuring concentrations among spurious ligands and devised a signaling network, based on a kinetic proofreading scheme [36], to approximately reach this limit. Similarly, Singh and Nemenman [37] found that a single receptor is capable of correctly measuring two different ligands with disparate binding affinities using a similar network. These studies focused on understanding the optimal sensing capacity of one receptor with a cognate and one (or many) noncognate ligands. In reality, receptors with different cognate ligands may communicate with each other through their downstream signaling networks, thereby increasing the efficiency of the measurement of the concentration of all ligands. This *crosstalk* between ligand-receptors pairs, to which we will refer to simply as “crosstalk”

*These authors equally to this work.

†Corresponding authors: ilya.nemenman@emory.edu, thierry.mora@ens.fr

throughout the paper, is distinct from the much-studied crosstalk between signaling pathways downstream of receptors). Ligand-receptor crosstalk is likely to occur in systems with a larger number of ligands than receptors. For example, in the bone morphogenetic protein signaling pathway, more than 30 different ligands interact with only seven different receptors [38,39]. Similar circumstances arise for Toll-like receptors in the innate immune system [40] and T-cell receptors in the adaptive immune system [41,42]. While the impact of crosstalk on adaptation dynamics has been studied in the context *Escherichia coli* chemotaxis [34], the fundamental limits to how different receptors combine information through crosstalk is currently unknown.

In this paper, we place physical limits on concentration sensing with crosstalking receptors. We focus on two types of receptors, each with their own cognate ligand, and compare specific (noncrossreactive) to crossreactive binding. To gain intuition, we first consider specific receptors which never bind to noncognate ligands. We then analyze the more realistic case where both receptors bind to both ligands with different binding strengths in a background with other ligands. In the latter case, we call the receptors specific if the off-rates of the noncognate ligand are as high as those for the background ligands. In both cases, we demonstrate that crosstalk outperforms specific binding of ligands in some parameter regimes as measured by relative errors on concentration estimation of the two ligands. We also discuss a related problem: detection of the presence of a ligand over a given concentration threshold using sequential probability ratio tests.

Our eventual goal is to understand systems with many ligands and receptors (many inputs–many outputs or MIMO). Nonetheless, our treatment of the two ligands–two receptors (2×2) scenario is an important special case that must be analyzed as the first step, and it is already a large departure from systems previously considered. Indeed, experimental and computational analysis of such systems usually focuses on deterministic chemical kinetics. However, in this regime, crosstalk is always detrimental for information transmission. Thus the discussion usually revolves around how to avoid the seemingly obvious disadvantages of crosstalk [43] and how crosstalk allows one to calculate various lower-dimensional functions of the involved ligands concentrations (e.g., their ratios), even when estimation of all ligand concentrations is impossible [39]. Here we argue that crosstalk in MIMO systems is advantageous, potentially suggesting an explanation for the ubiquity of MIMO systems in nature.

II. MAXIMUM LIKELIHOOD ANALYSIS OF TWO CROSTALKING RECEPTORS

We consider a situation of crosstalk in the simplified case of two receptors, labeled A and B, and two ligands, labeled 1 and 2. The binding rates of the two receptors k_A and k_B are assumed to be independent of the identity of the ligand, as in the case of diffusion-limited binding, in which case $k_A = 4Ds_A$, $k_B = 4Ds_B$, where s_A , s_B are the sizes of idealized circular receptors located on the cell surface and D is the diffusivity of the ligand molecules. The distinction between the two ligands appears in the distinct rates of unbinding, which are denoted by $r_{A,1}$, $r_{A,2}$, $r_{B,1}$, $r_{B,2}$. In the presence of both ligands,

receptors alternate between bound and unbound states with exponential waiting times associated with the off and on rates, respectively. In general, the mean occupancy of each receptor contains useful information about the concentrations of each ligand. However, the temporal information contained in the sequence of bound and unbound times is lost; a maximum likelihood estimation based off the binding and unbinding events yields an estimator that is unbiased and asymptotically achieves the least variance, as given by the Cramér-Rao bound [9,35,44]. For independent receptors, we may split the log-likelihood \mathcal{L} into contributions from binding events at each receptor, $\mathcal{L} = \mathcal{L}_A + \mathcal{L}_B$. For ligands with concentrations c_1 and c_2 (with $c_{\text{tot}} = c_1 + c_2$), the probability of a sequence of bound and unbound times on receptor A is written as

$$P(\{u_i^A, b_i^A\}) \equiv e^{\mathcal{L}_A} = \prod_{i=1}^{n_A} e^{-k_A c_{\text{tot}} u_i^A} [k_A c_1 r_{A,1} e^{-r_{A,1} b_i^A} + k_A c_2 r_{A,2} e^{-r_{A,2} b_i^A}], \quad (1)$$

where $\{u_i^A, b_i^A\}$ denotes the sequence of unbound and bound times and the index i runs from one to the total number of binding events n_A in a fixed time interval T , which is taken to be much longer than the typical binding and unbinding times. A similar expression can be written for receptor B. The log-likelihood now reads

$$\mathcal{L} = \sum_{R=A,B} \left\{ -k_R c_{\text{tot}} T_u^R + n_R \log k_R c_{\text{tot}} + \sum_{i=1}^{n_R} \log [x r_{R,1} e^{-r_{R,1} b_i^R} + (1-x) r_{R,2} e^{-r_{R,2} b_i^R}] \right\}, \quad (2)$$

with $x = c_1/c_{\text{tot}}$ and where T_u^R is the total unbound time of the receptor R. Maximum likelihood (ML) estimates of x and c_{tot} are obtained from the conditions $\frac{\partial \mathcal{L}}{\partial x}|_{x^*} = 0$ and $\frac{\partial \mathcal{L}}{\partial c_{\text{tot}}}|_{c_{\text{tot}}^*} = 0$. The ML estimate of the total concentration, c_{tot}^* , is given by

$$c_{\text{tot}}^* = \frac{n_A + n_B}{k_A T_u^A + k_B T_u^B}. \quad (3)$$

The ML estimate x^* satisfies the equation

$$\sum_{i=1}^{n_A} \frac{\alpha_A e^{-(\alpha_A-1)r_{A,2}b_i^A} - 1}{x^* \alpha_A e^{-(\alpha_A-1)r_{A,2}b_i^A} + (1-x^*)} - \sum_{i=1}^{n_B} \frac{\alpha_B e^{-(\alpha_B-1)r_{B,1}b_i^B} - 1}{(1-x^*) \alpha_B e^{-(\alpha_B-1)r_{B,1}b_i^B} + x^*} = 0. \quad (4)$$

Here we have defined $\alpha_A = r_{A,1}/r_{A,2}$ and $\alpha_B = r_{B,2}/r_{B,1}$.

III. PRECISION IN CONCENTRATION SENSING

Since c_{tot} and x are involved in separate terms of the log-likelihood (2), the variances of their optimal estimators are given, in the limit of large numbers of binding events, by the inverse of their respective Fisher information, $\delta x^2 = \langle \frac{\partial^2 \mathcal{L}}{\partial x^2} \rangle^{-1}$ and $\delta c_{\text{tot}}^2 = \langle \frac{\partial^2 \mathcal{L}}{\partial c_{\text{tot}}^2} \rangle^{-1}$, where the angled brackets denote an expectation over the distribution parametrized by the true parameter values. A similar approach is taken in Ref. [35];

here we simply write the final form for the variance in the estimates of c_{tot} and x :

$$\delta c_{\text{tot}}^2 = \frac{c_{\text{tot}}^2}{\langle n_A \rangle + \langle n_B \rangle}, \quad (5)$$

$$\delta x^2 = \left[\frac{\langle n_A \rangle}{f(x, \alpha_A)} + \frac{\langle n_B \rangle}{f(1-x, \alpha_B)} \right]^{-1}, \quad \text{where} \quad (6)$$

$$f(x, \alpha)^{-1} = \int_0^\infty dt e^{-t} \frac{[\alpha e^{-(\alpha-1)t} - 1]^2}{x \alpha e^{-(\alpha-1)t} + (1-x)}, \quad (7)$$

where averages and variances are taken over realizations of sequences of bound and unbound times (or, equivalently, averaged over one sequence in the limit of large T) and $\langle n_i \rangle = T / [(k_i c_{\text{tot}})^{-1} + x r_{i,1}^{-1} + (1-x) r_{i,2}^{-1}]$. Before proceeding further, we make a few simplifying assumptions of symmetric binding of the two ligands on the two receptors (Fig. 1). Particularly, we assume $r_{A,1} = r_{B,2} = r$, $r_{A,2} = r_{B,1} = r'$ and $k_A = k_B = k$, which give $\alpha_A = \alpha_B \equiv \alpha = r/r'$. For $\alpha < 1$, the conditions imply that ligand 1 acts as a cognate ligand for receptor A and a noncognate ligand for receptor B, whereas the reciprocal relationship is true for ligand 2. The discriminability of the two ligands is set by α , which measures the ratio of bound times of the cognate and noncognate ligands. It is convenient to nondimensionalize the concentrations of the two ligands as $\tilde{c}_1 = k c_1 / r$ and $\tilde{c}_2 = k c_2 / r$. The mean number of binding events on the two receptors is given by the total time divided by the mean time for each binding and unbinding cycle, which gives

$$\langle n_A \rangle = \frac{rT(\tilde{c}_1 + \tilde{c}_2)}{1 + \tilde{c}_1 + \alpha \tilde{c}_2}, \quad \langle n_B \rangle = \frac{rT(\tilde{c}_1 + \tilde{c}_2)}{1 + \alpha \tilde{c}_1 + \tilde{c}_2}. \quad (8)$$

In the case of specific receptors, i.e., when each ligand binds to only one receptor type, the minimum variance of the estimated concentration of each ligand can be derived in a similar manner from the log-likelihood (see Ref. [9]). If we suppose ligand 1 binds specifically to receptor A (with the same binding and unbinding rates as in the crosstalk case above), and ligand 2 binds only to receptor B, then the error in ML estimation is given by

$$\delta c_1^2 = c_1^2 / \langle n_A \rangle_{\text{spec}}, \quad \delta c_2^2 = c_2^2 / \langle n_B \rangle_{\text{spec}}, \quad \text{where} \quad (9)$$

$$\langle n_A \rangle_{\text{spec}} = \frac{rT \tilde{c}_1}{1 + \tilde{c}_1}, \quad \langle n_B \rangle_{\text{spec}} = \frac{rT \tilde{c}_2}{1 + \tilde{c}_2}, \quad (10)$$

where $\langle n_A \rangle_{\text{spec}}$, $\langle n_B \rangle_{\text{spec}}$ are the average number of binding events for the specific receptor case in the same interval T . Note that (10) does not correspond to the $r' \rightarrow \infty$ limit in Eqs. (5) and (6): When r' is very large, noncognate ligand bound times can be read easily because there is no cutoff for the readout of small bound times. This biological inconsistency can be removed by taking into account binding of nonspecific molecules (see Sec. V).

To make a comparison between the effectiveness of crosstalking and specific receptors, it is more pertinent to estimate relative errors, $\delta c_1 / c_1$, $\delta c_2 / c_2$, as concentrations can span many orders of magnitude. In the limit of long times,

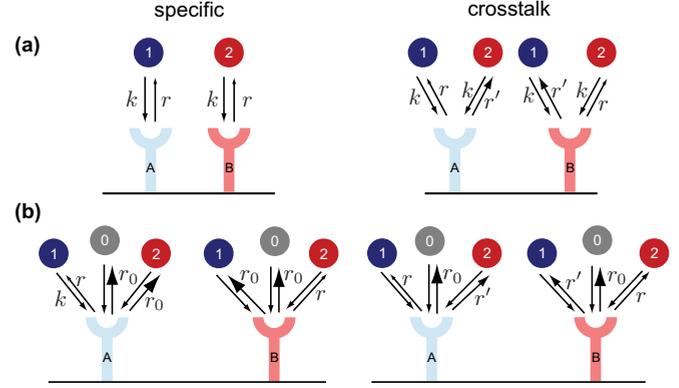


FIG. 1. Specific and crosstalk receptors for concentration sensing. (a) The receptors A and B bind to their cognate ligand 1 and 2, respectively, with the same rate k . Crosstalk is defined as the receptors additionally binding to another noncognate ligand. The noncognate ligand unbinds faster ($r' > r$), which allows for discrimination between the two types of binding events. (b) The two receptors bind to both ligand species, as well as to a pool of background ligands 0, all with the same binding rate k . The off rates of the noncognate ligand can either be identical to the background ligands (specific case, $r < r' = r_0$) or in between the background and the cognate ligand (crosstalk case, $r < r' < r_0$).

where errors are Gaussian distributed, the covariance matrix

$$\Sigma = \begin{bmatrix} \frac{\delta c_1^2}{c_1^2} & \frac{\delta c_1 \delta c_2}{c_1 c_2} \\ \frac{\delta c_1 \delta c_2}{c_1 c_2} & \frac{\delta c_2^2}{c_2^2} \end{bmatrix} \quad (11)$$

describes the magnitude and shape of relative estimation errors as an ellipse around the true value. Its determinant, $\Omega = \det(\Sigma)$, gives the volume of that ellipse and is used as a measure of error. Intuitively, the discriminability between two pairs of concentrations (c_1, c_2) and (c'_1, c'_2) depends on the overlap in the areas of the two ellipses centered around the pairs [45]. In the crosstalk case, from (5) we have

$$\Omega_{\text{CT}} = \frac{1}{x^2(1-x)^2} \frac{1}{\langle n_A \rangle + \langle n_B \rangle} \times \frac{1}{\langle n_A \rangle f(x, \alpha)^{-1} + \langle n_B \rangle f(1-x, \alpha)^{-1}}, \quad (12)$$

For the case of specific binding, the measurements of c_1 and c_2 are independent, and from (10), the determinant of the covariance matrix is

$$\Omega_S = \frac{1}{\langle n_A \rangle_{\text{spec}} \langle n_B \rangle_{\text{spec}}}. \quad (13)$$

In the limit of weak crosstalk, $\alpha \ll 1$, we find that crosstalk between receptors always improves sensing capacity over noncrosstalking receptors, regardless of the concentrations of each ligand. To see this, note $f(x, \alpha) \approx (1-x)$ and $f(1-x, \alpha) \approx x$ [from (7)], and so

$$\Omega_S / \Omega_{\text{CT}} \approx 2 \frac{[\mu + 2x(1-x)](\mu + 1/2)}{(\mu + x)(\mu + 1 - x)} > 1, \quad (14)$$

where $\mu = (\tilde{c}_1 + \tilde{c}_2)^{-1}$. In particular, when $c_1 \approx c_2$, the number of binding events for the crosstalking receptors is twice that for the specific receptors and, consequently, Ω_S is a factor

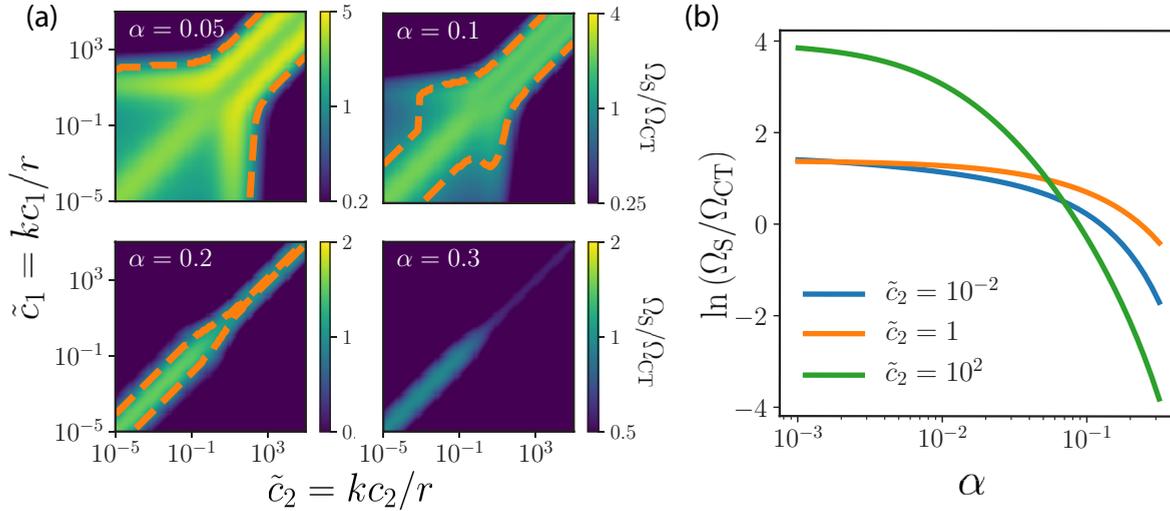


FIG. 2. Comparison of concentration sensing accuracy with specific and crosstalking receptors. (a) Ratio between the specific and crosstalk errors, Ω_S/Ω_{CT} , for four different values of α and different concentration pairs. The region where crosstalk shows greater precision $\Omega_{CT} < \Omega_S$ is delineated by the orange dashed line. For $\alpha > \alpha_c \approx 0.27$, crosstalk does not exhibit greater precision in any range of concentrations. (b) The logarithm of the same ratio for $\tilde{c}_1 = 1$ and three values of \tilde{c}_2 as a function of α .

2 greater than Ω_{CT} in this limit. In Fig. 2 we mark the regions in the $(\tilde{c}_1, \tilde{c}_2)$ plane where $\Omega_S > \Omega_{CT}$ for different values of α . We make two major observations. First, for $\alpha = 0$ i.e., for perfect discriminability, crosstalking receptors always show lower estimation error [Eq. (14)]. Second, there is a critical value of α , $\alpha_c \approx 0.27$, beyond which Ω_S is smaller than Ω_{CT} for all concentrations. Thus, the usefulness of crosstalk diminishes with increasing α , as in Ref. [37].

IV. PRECISION IN CONCENTRATION SENSING

Next, we consider the task of discriminating between two external environmental states or “hypotheses,” H_0 and H_1 , corresponding to ligand concentrations (c_1, c_2) and (c'_1, c'_2) , respectively. A Bayes-optimal decision-maker compares the likelihood ratio or, equivalently, the difference in log-likelihoods, $\mathcal{L}_0 - \mathcal{L}_1 \equiv \Delta\mathcal{L}$, for the two hypotheses, given the observed bound and unbound times, and makes a decision if $\Delta\mathcal{L}$ crosses a certain threshold θ in a given amount of time T . The false-positive and false-negative error rates depend on θ and the distribution of $\Delta\mathcal{L}$ under the hypotheses. To conveniently compare the discriminability for the crosstalk and specific cases, we use the discriminability (or “sensitivity index” d'), defined as

$$d' = \frac{\langle \Delta\mathcal{L} \rangle_c - \langle \Delta\mathcal{L} \rangle_{c'}}{\sqrt{\langle \delta^2 \Delta\mathcal{L} \rangle_c + \langle \delta^2 \Delta\mathcal{L} \rangle_{c'}}}, \tag{15}$$

where subscripts c and c' denote expectation values under concentrations (c_1, c_2) and (c'_1, c'_2) , respectively. If $\Delta\mathcal{L}$ is Gaussian distributed and its variance equal for both sets of concentrations, then d' and θ together uniquely determine the false-positive and false-negative error rates. For large T , the central limit theorem guarantees that the Gaussian approximation is a good one. Although the variances $\langle \delta^2 \Delta\mathcal{L} \rangle_c$ and $\langle \delta^2 \Delta\mathcal{L} \rangle_{c'}$ are not equal in general, d' is often used as a general measure of discriminability.

To calculate the mean and variance of $\langle \Delta\mathcal{L} \rangle_c$, we first observe that since the two receptors are independent, it is sufficient to compute them for a single receptor and sum them up. We show in the Appendix that for a single receptor with arbitrary unbound and bound time distributions, the cumulant generating function of $\langle \Delta\mathcal{L} \rangle_c$ has a simple form in the limit of large T , from which we derive explicit forms for the mean and variance. In Fig. 7, we validate these analytical expressions using numerical simulations.

The discriminabilities for the crosstalking and specific receptors, d'_{CT} and d'_S , are compared in Fig. 3. The orange line mark the region where crosstalk offers greater discriminability compared to the specific case. Large concentrations of the noncognate ligand can mask the accurate discrimination between different concentrations of the cognate ligand, as evidenced by the blue region. As $\alpha \rightarrow 0$, masking plays a limited role, since even though binding events are dominated by the noncognate ligand, the bound times are extremely short and easily distinguishable from the bound times of the cognate ligand.

V. CONCENTRATION ESTIMATES OF TWO LIGANDS IN A POOL OF NONSPECIFIC LIGANDS

Next, we consider a more realistic scenario the cell faces: the problem of concentration estimation of two ligands in a presence of a pool of background nonspecific ligands. This scenario also allows us to resolve the inconsistency of the limit of perfect specificity, $\alpha \rightarrow 0$. As we observed earlier, that limit did not reduce to the case of no crosstalk because of infinitely short but mathematically informative nonspecific binding events. Adding a background of nonspecific ligands removes the informative content of these spurious events by making them indistinguishable from background ligand binding. Here, as before, we have two receptors, A and B, and two cognate ligands, 1 and 2, whose concentrations c_1 and c_2 the cell needs to estimate. In addition, there exists a

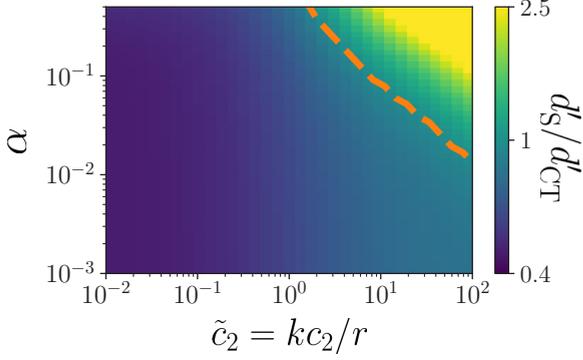


FIG. 3. Crosstalking receptors show improved discriminability relative to specific receptors. Ratio of sensitivity index d' for the crosstalk (d'_{CT}) and specific (d'_S) cases, as a function of the concentration of the second ligand c_2 , and the specificity ratio $\alpha = r/r'$, in the task of discriminating two values of the first ligand concentration, $kc_1/r = 0.5$ or $kc'_1/r = 2$. The orange line separates the regions in which each of the two strategy (crosstalk or specific) is optimal. Crosstalk is beneficial for small c_2 and α , i.e., when the ligands are easy to distinguish and do not saturate the receptor. The sensitivity index was computed using $rT = 10^4$.

pool of nonspecific ligands labeled 0, with concentration c_0 . The receptors now have to differentiate between three types of ligands to correctly estimate their concentrations. We assume diffusion-limited ligand-receptor binding as before. The off-rates for the nonspecific ligands are $r_{A,0} = r_{B,0} = r_0 \geq r' > r$ from both receptors. With total concentration given by $c_{tot} = c_0 + c_1 + c_2$, and relative fractions $x = c_1/c_{tot}$, $y = c_2/c_{tot}$,

the probability $P(\{u_i^R, b_i^R\})$ of a sequence of bound and unbound times can be written as a function of the on-rate, various off-rates, and concentrations. As before, the ML estimate of $\theta = [x, y, c_{tot}]$ can be obtained by maximizing the log-likelihood $\mathcal{L} = \log P$ by setting $\frac{\partial \mathcal{L}(\theta)}{\partial \theta} \Big|_{\theta^*} = 0$. Further, in the limit of large numbers of binding events, the Cramér-Rao bound guarantees that the covariance matrix of the estimator θ is given by the inverse of the Fisher information matrix:

$$\langle \delta \theta^T \delta \theta \rangle = - \left\langle \frac{\partial^2 \mathcal{L}}{\partial \theta^T \partial \theta} \right\rangle_{\theta^*}^{-1}. \quad (16)$$

The covariance matrix Σ_{NS} of the relative errors on concentrations, $\delta c_1/c_1$ and $\delta c_2/c_2$, can then be obtained by a change of variable from θ to (c_0, c_1, c_2) , from which the error volume $\Omega_{NS} = \det(\Sigma_{NS})$ is computed, analogous to Ω_S and Ω_{CT} defined earlier.

We are interested in how Ω_{NS} varies as a function of r and r' to understand whether crosstalk (defined now as $r' < r_0$ by contrast to background unspecific binding, $r' = r_0$) can result in relatively better concentration estimates. Up to scalar scaling, Ω_{NS} depends only on the ratios kc_{tot}/r_0 , r/r_0 , and r'/r_0 and the ligand concentration fractions x and y .

The estimation error quantified by Ω_{NS} attains a minimum [Fig. 4(a)] at some finite values of $r = r_{min}$ and $r' = r'_{min} < r_0$, meaning that finite crosstalk minimizes the estimation error, providing a substantial drop in errors relative to no crosstalk, $r' = r_0$ (as illustrated by the difference between the minimum and the rightmost point of the curve in the inset). At a fixed r , for $r' > r$, while Ω_{CT} monotonically decreases with increasing r' , Ω_{NS} is nonmonotonic with a local minimum at $r' = r'_{min} \in (r, r_0)$ [inset, Fig. 4(a)]. The existence of a minimum

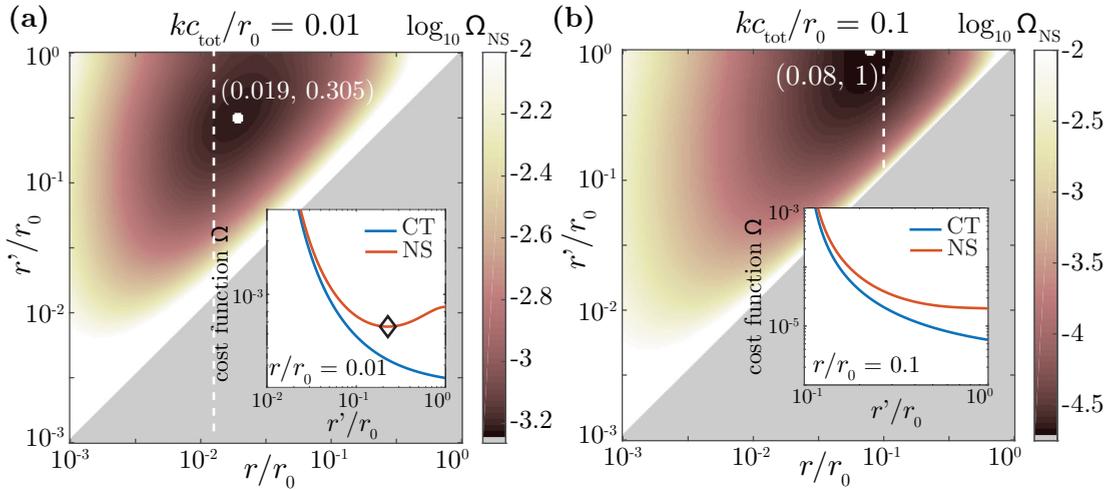


FIG. 4. Optimal unbinding rates in presence of background unspecific binding. Logarithm of the error, $\log \Omega_{NS}$ as a function of the unbinding rate of the cognate ligands, r , and the unbinding of noncognate ligands, r' , for the ML estimates in the crosstalk model with nonspecific ligands with $x = y = 0.4$. (a) Weak background binding ($kc_{tot}/r_0 = 0.01$). Ω_{NS} attains a local minimum at $r_{min} < r'_{min} < r_0$, meaning that crosstalk optimizes the accuracy of concentration sensing. In the inset, we plot Ω_{NS} (red) and Ω_{CT} (blue) as a function of r'/r_0 , at fixed $r/r_0 = 0.01$ (along the dotted white line in the main plot). While Ω_{CT} decreases monotonously as r' increases (and α decreases), Ω_{NS} attains a local minimum (black diamond). A good estimation of the cognate ligand concentrations requires all unbinding rates to be as dissimilar as possible, ensuring that the identity of the bound ligand can be faithfully inferred from the binding times. (b) Strong background binding ($kc_{tot}/r_0 = 0.1$). The minimum is reached at the boundary, ($r'_{min} = r_0$), meaning that the noncognate ligand is indistinguishable from background binding and thus treated as noise. This optimal solution thus reduces to the case of specific binding in the presence of background binding. Plots are obtained with $kc_1/r_0 = 0.4$, $kc_2/r_0 = 0.4$, $kc_3/r_0 = 0.2$, and $r_0T = 10^4$.

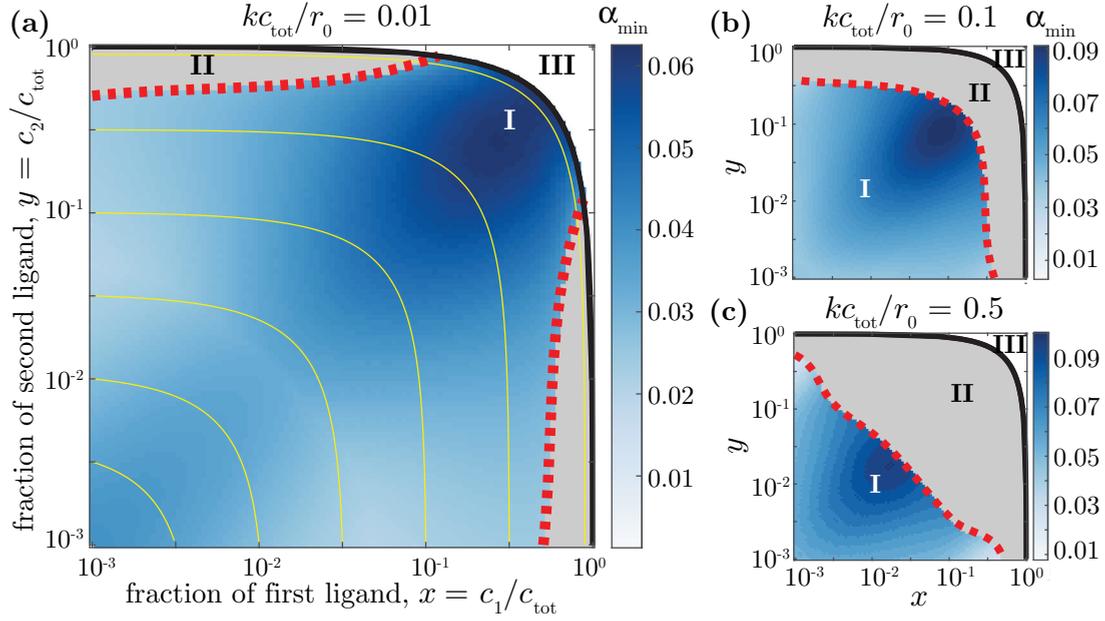


FIG. 5. Optimal strategy as a function of ligand fractions. Three phases are shown: I, crosstalking receptors are optimal ($r_{\min} < r'_{\min} < r_0$); II, specific receptors are optimal ($r_{\min} < r'_{\min} = r_0$); III, impossible region ($x + y > 1$). The heatmap shows the optimal value of the specificity ratio, $\alpha_{\min} = r_{\min}/r'_{\min}$. Yellow lines indicate a constant c_0 (or $x + y$). The phase diagram depends on the strength of background binding: (a) $kc_{\text{tot}}/r_0 = 0.01$, (b) $kc_{\text{tot}}/r_0 = 0.1$, and (c) $kc_{\text{tot}}/r_0 = 0.5$. As this strength increases, the region where specific binding is optimal (phase II) extends toward smaller values of x and y .

results from a trade-off between three conflicting effects. First, when r' is large, receptors cannot reliably distinguish between the noncognate and the nonspecific ligand, resulting in an increased estimation error for x and y . On the other hand, larger values of r' and r result in more binding events, thereby improving statistics and accuracy. Last, r and r' should be as far as possible from each other for each receptor to be able to distinguish cognate from noncognate ligands. These opposing forces result in a local minimum at (r_{\min}, r'_{\min}) , balancing the need for speed (high unbinding rates r, r') with that of specificity (low but different r and r'), ensuring that receptors can maximally distinguish between the three different kinds of ligands by looking at their binding times.

Distinct regimes emerge depending on the value of the dimensionless parameter kc_{tot}/r_0 , which can be viewed as an effective unspecific receptor occupancy quantifying the effect of background binding, as well as on the ligand fractions x and y . For small kc_{tot}/r_0 [Fig. 4(a)], a trade-off exists between speed and specificity ($r'_{\min} < r_0$), while for large kc_{tot}/r_0 [Fig. 4(b)], the low receptor availability caused by unspecific background binding makes the speed requirement dominate, resulting in absence of crosstalk in the optimal solution ($r'_{\min} = r_0$). Figure 5 shows the phase diagram as a function of the ligand fractions x and y for fixed values of $kc_{\text{tot}}/r_0 = 0.01$. Three phases emerge: crosstalk ($r_{\min} < r_0$, I), no crosstalk ($r_{\min} = r_0$, II), and impossibility region ($x + y > 1$, III).

When one ligand is present in very low concentrations and the other in high concentrations (low x , high y , or vice versa), introducing crosstalk would cause the abundant ligand to saturate both receptors, keeping the receptor that is cognate to the sparse ligand from sensing its concentration. In that regime, crosstalk is not optimal [region II of Fig. 5(a)]. As kc_{tot}/r_0

increases [Fig. 5(b)], this region spreads toward smaller values of x and y . Yet even when the effect of background ligand is felt strongly, crosstalk is still advantageous when the cognate ligands are sparse ($x + y \ll 1$).

VI. A BIOCHEMICAL NETWORK SCHEME THAT REACHES CLOSE TO THE OPTIMAL BOUNDS

In this section, we present a simple kinetic proofreading-based scheme that implements an approximate maximum likelihood estimation with a precision close to the derived optimal bounds. In the following analysis, for simplicity, we revert to the former case of one cognate and one noncognate ligand. We note that the ML estimate for the total concentration $\hat{c}_{\text{tot}} = (n_A + n_B)/(k_A T_u^A + k_B T_u^B)$ has a simple expression; the terms in the numerator and denominator can be measured biochemically and combined to form \hat{c}_{tot} . To estimate x , the scheme relies on a proofreading “classifier” associated to each receptor that distinguishes between bound times above and below a certain threshold. An unbiased estimate of x , which we denote by \hat{x}_R , can be formed for each receptor R ($R = A$ or B), based on the fraction \hat{f}_R of binding events where the ligand is bound longer than τ_R . Defining $y_R = e^{-r'\tau_R}$, $h_R = e^{-r\tau_R} - e^{-r'\tau_R} = y_R^\alpha - y_R$, this estimate reads:

$$\hat{x}_A = \frac{\hat{f}_A - y_A}{h_A}, \quad (17)$$

$$\hat{x}_B = \frac{y_B^\alpha - \hat{f}_B}{h_B}. \quad (18)$$

The error in each estimate is $\delta x_R^2 = f_R(1 - f_R)/(h_R^2 n_R)$, where n_R is the mean total number of binding events in time T and $f_R \equiv \langle \hat{f}_R \rangle$. Because the estimates from receptors A and B stem

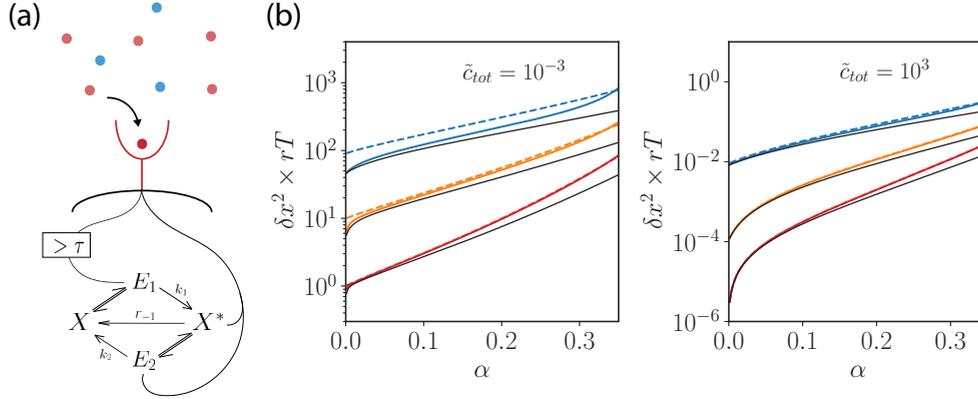


FIG. 6. (a) Network scheme to estimate $x = c_1/c_{\text{tot}}$ using the fraction of binding events that last longer than some threshold τ . An estimate of x is read off from the fraction of X molecules in the active state (see text). (b) The error in estimating x for $x = 10^{-1}, 10^{-3}, 10^{-5}$ (blue, orange, red respectively) using (19) (solid, colored lines) and the crude estimate of using only one receptor (dashed lines) is compared with the optimal ML error (solid, black lines).

from independent binding events, the best way to combine them is through a weighted average of the estimates from each receptor:

$$\hat{x} = \beta \hat{x}_A + (1 - \beta) \hat{x}_B, \quad (19)$$

with $\beta = 1/(1 + \delta x_A^2/\delta x_B^2)$.

The value of the threshold τ_A can be optimized to yield the most precision in \hat{x}_A . In Appendix B, we show that the optimum is reached for $r'\tau_A^* = (1 - \alpha)^{-1} \log(\frac{1-2\alpha}{\alpha x})$, valid for $\alpha < 1/2$. We note here that even though τ_A^* depends on x , the dependence is logarithmic. To implement adaptive thresholds that depend on $\log x$, we can imagine a collection of “gating” networks which apply different thresholds to the length of the binding events. An independent crude estimate of x is sufficient to choose the network that has a threshold closest to τ_A^* . The optimal value for τ_B has a similar form with x replaced by $1 - x$.

The cell can easily compute \hat{x}_A and \hat{x}_B using a proofreading motif followed by a downstream push-pull network as shown in Fig. 6(a). For concreteness, consider the case of receptor A. Suppose the proofreading motif produces a molecule of enzyme E_1 for each binding event longer than the threshold τ_A . Further, every binding event generates a single protein X and an enzyme molecule E_2 . The enzymes E_1 and E_2 catalyze the conversion of X to its active state X^* and vice versa, respectively. Assuming that X and X^* are in excess in the enzymatic reaction, the rate at which X and X^* are interconverted is directly proportional to the enzyme numbers with catalytic rates denoted by k_1 and k_2 . Suppose also that X^* reverts to X at a finite rate r_{-1} . At steady state, $dX^*/dt = 0$, and we have $r_{-1}X^* = k_1E_1 - k_2E_2$. Then, setting $k_1/r_{-1} = 1/h_A$ and $k_2/r_{-1} = \gamma_A/h_A$, the fraction of X molecules in the active state tracks \hat{x}_A .

Combining \hat{x}_A and \hat{x}_B as in Eq. (19) using a biochemical network can be done at fixed β . For instance, when $x \sim 1/2$, the errors from both receptors are about the same and we may weight the estimate from each receptor equally, $\beta = 1/2$. However, tuning β to reflect its dependency of the concentrations requires additional adaptive mechanisms. In the regime $x \ll 1$, receptor A has the highest precision in

estimating x , and \hat{x}_A can be taken as a crude estimate of x (and symmetrically for B when $1 - x \ll 1$). In Fig. 6(b), we compare the error from this crude estimate against the optimal crosstalk error and the error from using the optimal network weights (19). This comparison shows that although the approximate biochemical solutions are not optimal, they stand reasonably close to the ML estimate.

VII. DISCUSSION

The key result of our paper is that crosstalk is generically the optimal strategy for sensing multiple ligand concentration using multiple receptors. The theory predicts an optimal level of crosstalk, which balances the opposing requirements of maximizing the number of binding events with the receptor’s ability to distinguish the ligands involved in those events. For this discrimination to be possible, the difference between cognate and noncognate binding affinities is maintained.

While we have restricted our analysis to two ligands sensed by two receptors (2×2), our framework can be readily generalized to more of each to deal with MIMO ligand-receptor crosstalking systems. We expect crosstalk to be also beneficial in this more general setting. In practice, crosstalk may be sparsely distributed and composed of many subsystems with limited number of binding partners as in our 2×2 system. Alternatively, in a related context of steady-state information processing, it was argued [46] that the joint probability distribution of activities of units in a large network can often be decomposed into contributions from just pairwise interactions, with an error from discarding the higher-order contributions becoming asymptotically small as the system grows. Thus, even though larger ligand-receptor networks may harbor new phenomena, necessitating future studies, the analysis presented here is likely to be useful in these scenarios as well.

In the simplest theory that we presented (Secs. II and III), higher unbinding rates are always advantageous because they increase receptor availability without hurting discriminability. However, in reality there are several reasons why high unbinding rates are not optimal. First, the molecular machinery required to process the information of the receptor binding state operates with its own incompressible timescale, which

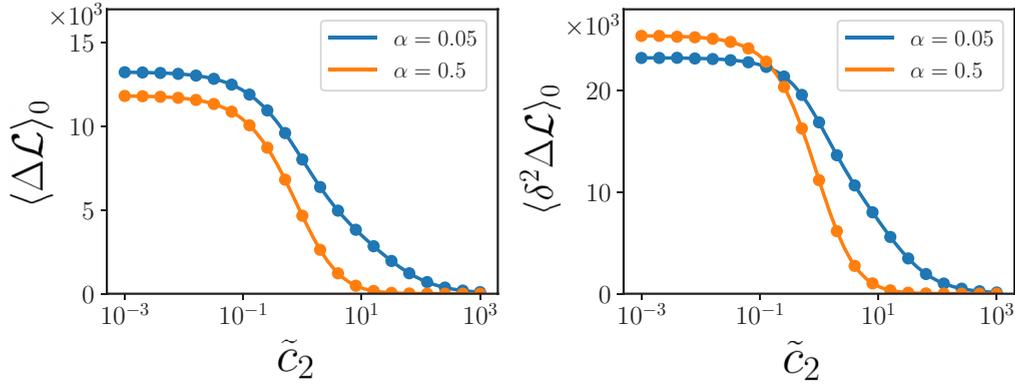


FIG. 7. The values from expressions for the mean [(A14), solid lines, left panel] and variance [(A16), solid lines, right panel] of the log-likelihood differences, $\Delta\mathcal{L}$, align with those obtained from numerical simulations (circles). Parameters are the same as those in Fig. 3.

sets a lower bound on the duration of binding events that can be detected. Second, it is not realistic to assume that unbinding rate can be increased arbitrarily without affect binding rates as well. Very unspecific ligands that unbind very quickly are also less likely to bind in the first place, and the assumption of diffusion-limited binding rate is no longer valid. Last, for high-enough unbinding rates, cognate and crosstalk binding events become indistinguishable from completely unspecific binding with other generic molecules. Modeling that situation as we did in Sec. V allowed us to find well-defined optimal unbinding rates.

We focused the application of our theory to a purely symmetric case, with equal proportions of the two types of receptors, but more general situations could be considered. For instance, in the case of an imbalance between the two ligands, $c_1 > c_2$, it makes sense to have more receptors of type A than of type B. Other interesting transitions may occur as we vary parameters, such as receptors of type B vanishing from the optimal solution as c_1 becomes more and more predominant, into a regime where a single receptor type A can sense both ligands as in Ref. [37]. We also assumed that binding rates were constant (since limited by diffusion only), because we were interested in the fundamental limits of sensing. More realistic situations where binding rates decrease with specificity would obviously limit the benefit of crosstalk.

In addition to maximizing the use of each receptor to gain information about each ligand, crosstalk has the additional advantage of expanding the dynamic range of concentrations over which ligands can be sensed, as extensively discussed in the context of chemotaxis [34]. For instance, when a ligand is present at high concentration, its cognate receptors will be fully saturated, making it difficult to reliably read off the concentration from the receptor's activity. Lower-affinity binding to a second receptor can then allow for more accurate sensing, as long as that receptor is not itself saturated by other ligands. In our language, when receptor A is saturated by ligand 1, $c_1 \gg r/k \equiv K_d$, then receptor B is still sensitive to the concentration of ligand 1 in the regime $c_1 \sim r'/k = \alpha^{-1}K_d$. More generally in the presence of multiple ligands and multiple receptors, a good strategy could be to organize specificities (i.e., unbinding rates) so that, for each ligand, dissociation constants collectively tile the sensory space. For such strategies to work, the concentration space must be

sparse, meaning that only one or a few of the ligands of interest are present in large concentrations at the same time. Receptors could also avoid saturation by only interacting with other receptors locally instead of globally [34], leading to accurate sensing of multiple ligands.

Crosstalk, also known as promiscuous binding, cross-reactivity, or multiplexing, depending on the context, is widespread in biology. It is an important feature of the Bone Morphogenetic Proteins (BMP) [38], Notch, Wnt [47], and JAK-STAT [48] pathways, as well as the Eph-ephrin system for cell positioning [49], T- and B-cell receptor antigen recognition [42,50], and olfactory receptors [51]. An often cited benefit of promiscuity is that it confers the ability to design combinatorial codes. In the context of olfaction [52], such design can be advantageous in the presence of sparse odors [53]. Combinatorial codes also allow for flexible computations in signaling pathways [39]. In the adaptive immune system, cross-reactivity is necessary to cover the large space of possible antigens with a limited number of receptors [41,42]. Our results suggest another advantage of crosstalk: sensing accuracy.

An open question is whether some of the biological systems that exhibit crosstalk make use of that benefit and whether they are organized in a way that approaches the optimal solution. While it is not clear which system (among those mentioned above or others) would be a good candidate to test our hypothesis, our theory makes a number of concrete predictions which could be tested experimentally. Specifically, our results suggest that receptors benefiting from crosstalk should signal through two or more different pathways, depending on the duration of the binding event [Eqs. (17) and (18)]. Long binding events should contribute to the signaling pathway that encodes the concentration of the ligand for which the receptor is most specific, while short binding events should activate the pathway corresponding to the secondary ligand and primarily activated by the other receptor type. This splitting of pathways as a function of binding time is possible in principle and underlies the difference between the recognition of foreign antigens and tonic stimulation of T-cell receptors in the adaptive immune system [54]. Incidentally, our design also predicts the existence of signaling crosstalk [Eq. (19)] as a consequence of crosstalk at the level of receptor binding, which could also be tested experimentally. For example,

quality of activation in response to ligand combinations can be measured on wild-type cells, as well as in cells with signaling crosstalk knocked out. In summary, our theoretical predictions could guide future experiments to focus on the predicted features and to quantify their action on the efficiency of concentration sensing.

ACKNOWLEDGMENTS

This work was initiated as a student project during the 2017 Cargèse school on biophysics held at the Institut d'Études Scientifique de Cargèse (Corsica, France), with support from the Centre National de la Recherche Scientifique through its training program and the *Groupement De Recherche Internationale* "Evolution, Regulation and Signaling," PSL University; the National Science Foundation (NSF) Grant No. PHY-1740578; the Institute for Complex Adaptive Matter through NSF Grant No. DMR-1411344; and the Collectivité Corse. I.N. was additionally partially supported by NSF Grant No. PHY-1410978.

APPENDIX A: CUMULANTS OF THE LOG-LIKELIHOOD FOR DISCRIMINATION

Consider a receptor switching between two states, bound and unbound, where the bound times u and unbound times

b are drawn from general distributions that depend on the external ligand environment. We will consider the problem of discriminating two possible environmental states, labeled H_0 and H_1 , given the time series of receptor states for a fixed time T , where T is much larger than the typical bound and unbound times. Our interest is in calculating the cumulants of the log-likelihood difference between the two hypotheses, $\mathcal{L}_0 - \mathcal{L}_1 \equiv \Delta\mathcal{L}$, under the hypothesis that either H_0 or H_1 is the true external state. We shall use $P_i(u)$ and $Q_i(b)$ to denote the distributions of unbound and bound times, respectively, for H_i .

The moment generating function $M(\lambda)$ of $\Delta\mathcal{L}$ under H_i is

$$M(\lambda) = \langle e^{\lambda\Delta\mathcal{L}} \rangle_i = \int e^{\lambda\Delta\mathcal{L} + \mathcal{L}_i} \mathcal{D}u \mathcal{D}b, \quad (\text{A1})$$

where the measure denotes an integral over all possible bound and unbound times in the interval T . The range of the integral can be split over distinct regions with a particular number of binding events n in time T , i.e., $\sum_{j=0}^n u_j + b_j > T$ but $\sum_{j=0}^{n-1} u_j + b_j < T$, where u_j and b_j denote the unbound and bound time at the j th binding event. Conditioning over n using the Heaviside function, we have

$$M(\lambda) = \sum_{n=1}^{\infty} \int e^{\lambda\Delta\mathcal{L} + \mathcal{L}_i} \left[\Theta\left(\sum_{j=0}^n u_j + b_j - T\right) - \Theta\left(\sum_{j=0}^{n-1} u_j + b_j - T\right) \right] \prod_{j=1}^n du_j db_j. \quad (\text{A2})$$

Note that \mathcal{L}_i is a sum of $2n$ independent contributions: $\mathcal{L}_i = \sum_{j=1}^n p_i(u_j) + \sum_{j=1}^n q_i(b_j)$, where $p_i \equiv \log P_i$ and $q_i \equiv \log Q_i$. The Heaviside function Θ can be replaced by its integral form

$$\Theta(x) = \int_{-\infty}^{i\infty} \frac{d\sigma}{2\pi i\sigma} e^{\sigma x}, \quad (\text{A3})$$

where the pole at $\sigma = 0$ should be taken to be in the left half of the complex plane. Applying this expression, we have

$$\Theta\left(\sum_j^n u_j + b_j - T\right) = \int_{-\infty}^{i\infty} \frac{d\sigma}{2\pi i\sigma} e^{-\sigma T} \prod_j^n e^{\sigma u_j + \sigma b_j}. \quad (\text{A4})$$

Using the same representation for the other Θ function, we have

$$M(\lambda) = \sum_{n=1}^{\infty} \int_{-\infty}^{i\infty} \frac{d\sigma}{2\pi i\sigma} e^{-\sigma T} \int \left[\prod_{j=1}^n e^{\lambda\Delta p(u_j) + p_i(u_j) + \lambda\Delta q(b_j) + q_i(b_j)} \right] \left(\prod_{j=1}^n e^{\sigma u_j + \sigma b_j} - \prod_{j=1}^{n-1} e^{\sigma u_j + \sigma b_j} \right) \prod_{j=1}^n du_j db_j, \quad (\text{A5})$$

where $\Delta p(u_j) = p_0(u_j) - p_1(u_j)$ and similarly for Δq . Simplifying, we have

$$M(\lambda) = \sum_{n=1}^{\infty} \int_{-\infty}^{i\infty} \frac{d\sigma}{2\pi i\sigma} e^{-\sigma T} [f(\sigma, \lambda)^n - f(0, \lambda)f(\sigma, \lambda)^{n-1}], \quad (\text{A6})$$

where we have defined

$$f(\sigma, \lambda) = \int_0^{\infty} e^{\lambda\Delta p(u) + p_i(u) + \sigma u} du \int_0^{\infty} e^{\lambda\Delta q(b) + q_i(b) + \sigma b} db. \quad (\text{A7})$$

Summing over n , we get

$$M(\lambda) = \int_{-\infty}^{i\infty} \frac{d\sigma}{2\pi i\sigma} e^{-\sigma T} \frac{f(\sigma, \lambda) - f(0, \lambda)}{1 - f(\sigma, \lambda)}. \quad (\text{A8})$$

From Cauchy's residue theorem, for large T , this integral is dominated by the residue from the smallest, positive pole σ such that $f(\sigma, \lambda) = 1$, which we call $\sigma_s(\lambda)$. Since we are interested in the cumulants of $\Delta\mathcal{L}$, which are generated by the Taylor series of $\log M(\lambda)$ as $\lambda \rightarrow 0$, it is sufficient to consider only small λ . To see that σ_s exists and is positive, we expand $f(\sigma, \lambda)$ around $f(0, 0)$ up to first order in a Taylor series,

$$f(\sigma, \lambda) = f(0, 0) + \sigma \frac{\partial f}{\partial \sigma} + \lambda \frac{\partial f}{\partial \lambda} + \dots, \quad (\text{A9})$$

where all derivatives here and below are evaluated at $(\sigma, \lambda) = (0, 0)$. Since $f(0, 0) = 1$ and $f(\sigma_s(\lambda), \lambda) = 1$, we have

$$\sigma_s(\lambda) = -\lambda \frac{\partial f / \partial \lambda}{\partial f / \partial \sigma} + O(\lambda^2). \quad (\text{A10})$$

To keep the moment-generating function well defined, for $i = 0$, we take $\lambda \rightarrow 0^-$ and for $i = 1$, we take $\lambda \rightarrow 0^+$, which together ensure that for both cases σ_s is a positive pole of the integrand in Eq. (A8). We circumvent the problem of calculating σ_s for each P_i and Q_i by noticing that for large T , the cumulant generating function $\log M$ can be simply written as

$$\log M(\lambda) = -T\sigma_s(\lambda) + o(T). \quad (\text{A11})$$

The m th-order cumulants are then obtained by taking the m th-order derivatives of $\sigma_s(\lambda)$ at $\lambda = 0$. Here we will derive the expressions for the mean and the variance; higher-order cumulants can be obtained by taking further derivatives. As noted above, $f(\sigma_s(\lambda), \lambda) = 1$. Differentiating both sides by λ , we obtain

$$\frac{d\sigma_s}{d\lambda} \frac{\partial f}{\partial \sigma} + \frac{\partial f}{\partial \lambda} = 0. \quad (\text{A12})$$

From here, we get

$$\langle \Delta\mathcal{L} \rangle_i = -T \frac{d\sigma_s}{d\lambda} = T \frac{\partial f / \partial \lambda}{\partial f / \partial \sigma}. \quad (\text{A13})$$

Taking the partial derivatives of f and evaluating at $(0, 0)$ gives:

$$\frac{\langle \Delta\mathcal{L} \rangle_i}{\langle n \rangle_i} = \langle \mathcal{P} + \mathcal{Q} \rangle_i, \quad (\text{A14})$$

where we have defined $\mathcal{P} = \log \frac{P_0}{P_1}$, $\mathcal{Q} = \log \frac{Q_0}{Q_1}$ and $\langle n \rangle_i = \frac{T}{\langle t_u + t_b \rangle_i}$ is the total number of binding events on the receptor in time T . For instance, at $(0, 0)$, we have $\partial f / \partial \lambda = \langle \log \frac{P_0}{P_1} + \log \frac{Q_0}{Q_1} \rangle_i$ and $\partial f / \partial \sigma = \langle u + b \rangle_i$. The variance can similarly be obtained by applying another derivative with respect to λ on (A12). Finally, we get

$$\langle \delta^2 \Delta\mathcal{L} \rangle_i = -T \frac{d^2 \sigma_s}{d\lambda^2} = T \frac{\left(\frac{d\sigma_s}{d\lambda} \right)^2 \frac{\partial^2 f}{\partial \sigma^2} + 2 \frac{d\sigma_s}{d\lambda} \frac{\partial^2 f}{\partial \sigma \partial \lambda} + \frac{\partial^2 f}{\partial \lambda^2}}{\frac{\partial f}{\partial \sigma}}. \quad (\text{A15})$$

The expression above can be evaluated as in the examples above to obtain:

$$\begin{aligned} \frac{\langle \delta^2 \Delta\mathcal{L} \rangle_i}{\langle n \rangle_i} &= \langle \mathcal{P} + \mathcal{Q} \rangle_i^2 \frac{\langle (t_u + t_b)^2 \rangle_i}{\langle t_u + t_b \rangle_i^2} \\ &\quad - 2 \langle \mathcal{P} + \mathcal{Q} \rangle_i \frac{\langle (t_u + t_b)(\mathcal{P} + \mathcal{Q}) \rangle_i}{\langle t_u + t_b \rangle_i} + \langle (\mathcal{P} + \mathcal{Q})^2 \rangle_i. \end{aligned} \quad (\text{A16})$$

APPENDIX B: OPTIMIZING THE PROOFREADING SCHEME

The variance of \hat{x}_A is

$$\delta x_A^2 = \frac{f_A(1 - f_A)}{n_A(e^{-\alpha r' \tau_A} - e^{-r' \tau_A})^2}, \quad (\text{B1})$$

where $f_A = x e^{-r \tau_A} + (1 - x) e^{-r' \tau_A}$. We define $y_A = e^{-r' \tau_A}$ and write the above equation as

$$\delta x_A^2 \times n_A = \frac{y_A(1 - y_A)}{(y_A^\alpha - y_A)^2} + x \frac{1 - 2y_A}{y_A^\alpha - y_A} - x^2, \quad (\text{B2})$$

To make the right-hand side of the above expression tractable for optimization, we observe that we are imposing a cutoff that discriminates between samples drawn from exponential distributions of means with ratio α^{-1} . As $\alpha \rightarrow 0$, the ratio gets larger and we expect the threshold to be placed much greater than mean bound time of noncognate binding events, $\sim r'^{-1}$. Accordingly, we propose an ansatz that $y_A^\alpha = O(1)$ and $y_A \ll 1$ as $\alpha \rightarrow 0$. Using the ansatz, we simplify the equation above to get

$$\delta x_A^2 \times n_A \approx y_A^{1-2\alpha} + x y_A^{-\alpha}. \quad (\text{B3})$$

Optimizing the right-hand side of the above equation with respect to y_A by taking the derivative in y_A and equating to zero, we get

$$r' \tau_A^* \approx (1 - \alpha)^{-1} \log \left(\frac{1 - 2\alpha}{\alpha x} \right), \quad (\text{B4})$$

which is reproduced in the main text and used in Figure 6. We verify then that as $\alpha \rightarrow 0$, $e^{-r' \tau_A^*} \ll 1$ and since $\alpha^\alpha \rightarrow 1$, we have $e^{-\alpha r' \tau_A^*} = O(1)$. A similar expression can be derived for τ_B^* with x replaced by $1 - x$ in Eq. (B4).

APPENDIX C: NONSPECIFIC LIGAND POOL

With total concentration given by $c_{\text{tot}} = c_1 + c_2 + c_0$, the probability of a sequence of bound and unbound times $\{u_i^R, b_i^R\}$ can be written as

$$\begin{aligned} P(\{u_i^R, b_i^R\}) &= \prod_{R=A,B} \prod_{i=1}^{n_R} e^{-k_R c_{\text{tot}} u_i^R} (k_R c_1 r_{R,1} e^{-r_{R,1} b_i^R} \\ &\quad + k_R c_2 r_{R,2} e^{-r_{R,2} b_i^R} + k_R c_0 r_{R,0} e^{-r_{R,0} b_i^R}), \end{aligned} \quad (\text{C1})$$

with index R running over the receptors A and B, and index i running over the binding events n_R in a fixed interval T . Now, log-likelihood $\mathcal{L}(x, y, c_{\text{tot}})$ as a function of the fractions of the cognate ligands, x and y , and the total concentration, c_{tot} , is given by

$$\begin{aligned} \mathcal{L} &= \sum_{R=A,B} \left\{ -k_R c_{\text{tot}} T u^R + n_R \log k_R r_{R,0} c_{\text{tot}} - r_{R,0} T b^R \right. \\ &\quad + \sum_{i=1}^{n_R} \log [x \lambda_{R,1} e^{(1-\lambda_{R,1}) r_{R,0} b_i^R} + y \lambda_{R,2} e^{(1-\lambda_{R,2}) r_{R,0} b_i^R} + 1 \\ &\quad \left. - x - y \right\}, \end{aligned} \quad (\text{C2})$$

where R runs over the receptors, i runs over the binding events (a total of n_R) of each receptor, b_i^R denotes the bound time of the i th binding event of receptor R , and T_u^R and T_b^R denote the total unbound and bound time, respectively, of receptor R . We also define $\lambda_{R,j} \equiv \frac{r_{R,j}}{r_{R,0}}$ as the ratio of the unbinding rate of ligand j to the unbinding rate of the nonspecific ligand from receptor R . By setting $\frac{\partial \mathcal{L}(\theta)}{\partial \theta} |_{\theta^*} = 0$, we obtain the ML estimates of $\theta = [x, y, c_{\text{tot}}]$. The ML estimate of total concentration, c_{tot}^* , is

$$c_{\text{tot}}^* = \frac{n_A + n_B}{k_A T_u^A + k_B T_u^B}. \quad (\text{C3})$$

The ML estimates of the fractions of cognate ligands, x^* and y^* , satisfy the following equations:

$$\sum_{R=A,B} \sum_{i=1}^{n_R} \frac{\lambda_{R,1} e^{(1-\lambda_{R,1})r_{R,0}b_i^R} - 1}{x^* \lambda_{R,1} e^{(1-\lambda_{R,1})r_{R,0}b_i^R} + y^* \lambda_{R,2} e^{(1-\lambda_{R,2})r_{R,0}b_i^R} + 1 - x^* - y^*} = 0, \quad (\text{C4})$$

$$\sum_{R=A,B} \sum_{i=1}^{n_R} \frac{\lambda_{R,2} e^{(1-\lambda_{R,2})r_{R,0}b_i^R} - 1}{x^* \lambda_{R,1} e^{(1-\lambda_{R,1})r_{R,0}b_i^R} + y^* \lambda_{R,2} e^{(1-\lambda_{R,2})r_{R,0}b_i^R} + 1 - x^* - y^*} = 0. \quad (\text{C5})$$

Next, by taking expectation over the distribution of the sequence of bound and unbound times, we obtain $\langle \frac{\partial \mathcal{L}}{\partial \theta} \rangle$, which informs us whether the ML estimates are unbiased, and $\langle \frac{\partial^2 \mathcal{L}}{\partial \theta^2} \rangle$, which help us obtain the Fisher information matrix, $I(\theta)$. We note that this probability P factorizes and can be written as

$$P(\{u_i^R, b_i^R\}) = \prod_{R=A,B} \prod_{i=1}^{n_R} p(u_i^R) p(b_i^R). \quad (\text{C6})$$

It is trivial to show that $\langle \frac{\partial \mathcal{L}}{\partial \theta} \rangle = 0$ and hence the ML estimates θ^* are unbiased. Now, setting $u = r_{R,0} b_i^R$, we have

$$\left\langle \frac{\partial^2 \mathcal{L}}{\partial x^2} \right\rangle = - \sum_{R=A,B} \langle n_R \rangle \int_0^{+\infty} du e^{-u} \frac{[\lambda_{R,1} e^{(1-\lambda_{R,1})u} - 1]^2}{x \lambda_{R,1} e^{(1-\lambda_{R,1})u} + y \lambda_{R,2} e^{(1-\lambda_{R,2})u} + 1 - x - y}, \quad (\text{C7})$$

$$\left\langle \frac{\partial^2 \mathcal{L}}{\partial y^2} \right\rangle = - \sum_{R=A,B} \langle n_R \rangle \int_0^{+\infty} du e^{-u} \frac{[\lambda_{R,2} e^{(1-\lambda_{R,2})u} - 1]^2}{x \lambda_{R,1} e^{(1-\lambda_{R,1})u} + y \lambda_{R,2} e^{(1-\lambda_{R,2})u} + 1 - x - y}, \quad (\text{C8})$$

$$\left\langle \frac{\partial^2 \mathcal{L}}{\partial x \partial y} \right\rangle = - \sum_{R=A,B} \langle n_R \rangle \int_0^{+\infty} du e^{-u} \frac{[\lambda_{R,1} e^{(1-\lambda_{R,1})u} - 1][\lambda_{R,2} e^{(1-\lambda_{R,2})u} - 1]}{x \lambda_{R,1} e^{(1-\lambda_{R,1})u} + y \lambda_{R,2} e^{(1-\lambda_{R,2})u} + 1 - x - y}. \quad (\text{C9})$$

We also have

$$\left\langle \frac{\partial^2 \mathcal{L}}{\partial c_{\text{tot}}^2} \right\rangle = - \frac{\langle n_A \rangle + \langle n_B \rangle}{c_{\text{tot}}^2}, \quad (\text{C10})$$

$$\left\langle \frac{\partial^2 \mathcal{L}}{\partial x \partial c_{\text{tot}}} \right\rangle = 0, \quad (\text{C11})$$

$$\left\langle \frac{\partial^2 \mathcal{L}}{\partial y \partial c_{\text{tot}}} \right\rangle = 0. \quad (\text{C12})$$

Using Eq. (C7)–(C12), the Fisher information matrix $I(\theta)$ can be obtained as $-\langle \frac{\partial^2 \mathcal{L}}{\partial \theta^2} \rangle$. The Cramér-Rao bound, in the limit of a large number of binding events, which can be ensured by choosing T to be significantly longer than a typical binding or unbinding event, ensures that

$$\Sigma_{\theta} = I(\theta)^{-1}. \quad (\text{C13})$$

We are interested in the log concentrations, $l_i = \log c_i$. What is an appropriate cost function in this case? With $\theta' = [l_1, l_2]$, and $\mathbf{l}_{\text{all}} = [l_1, l_2, l_0] = [\theta', l_0]$, we obtain the covariance matrix $\Sigma_{\mathbf{l}_{\text{all}}}$ as

$$\Sigma_{\mathbf{l}_{\text{all}}} = J_{\text{all}} \Sigma_{\theta} J_{\text{all}}^T, \quad (\text{C14})$$

where J_{all} is the Jacobian given by

$$\begin{pmatrix} x^{-1} & 0 & c_{\text{tot}}^{-1} \\ 0 & y^{-1} & c_{\text{tot}}^{-1} \\ -z^{-1} & -z^{-1} & c_{\text{tot}}^{-1} \end{pmatrix}. \quad (\text{C15})$$

This describes an ellipsoid in I_{all} space. We are interested only in θ' . So we obtain the marginal distribution of l_1 and l_2 and obtain the covariance matrix, $\Sigma_{\theta'} = J \Sigma_{\theta} J^T$, and the cost function Ω_{NS} as the area of the ellipse centered at (l_1, l_2) as

$$\Omega_{\text{NS}} = \det(\Sigma_{\theta'}) = \det(J \Sigma_{\theta} J^T), \quad (\text{C16})$$

where the Jacobian J given by

$$\begin{pmatrix} x^{-1} & 0 & c_{\text{tot}}^{-1} \\ 0 & y^{-1} & c_{\text{tot}}^{-1} \end{pmatrix}. \quad (\text{C17})$$

-
- [1] C. G. Bowsher and P. S. Swain, Environmental sensing, information transfer, and cellular decision-making, *Curr. Opin. Biotechnol.* **28**, 149 (2014).
- [2] G. H. Wadhams and J. P. Armitage, Making sense of it all: Bacterial chemotaxis, *Nat. Rev. Mol. Cell Biol.* **5**, 1024 (2004).
- [3] S. Akira, S. Uematsu, and O. Takeuchi, Pathogen recognition and innate immunity, *Cell* **124**, 783 (2006).
- [4] P. R. ten Wolde, N. B. Becker, T. E. Ouldridge, and A. Mugler, Fundamental limits to cellular sensing, *J. Stat. Phys.* **162**, 1395 (2016).
- [5] W. Bialek and S. Setayeshgar, Physical limits to biochemical signaling, *Proc. Natl. Acad. Sci. USA* **102**, 10040 (2005).
- [6] K. Kaizu, W. de Ronde, J. Pajmans, K. Takahashi, F. Tostevin, and P. R. ten Wolde, The berg-purcell limit revisited, *Biophys. J.* **106**, 976 (2014).
- [7] G. Aquino, N. S. Wingreen, and R. G. Endres, Know the single-receptor sensing limit? think again, *J. Stat. Phys.* **162**, 1353 (2016).
- [8] H. C. Berg and E. M. Purcell, Physics of chemoreception, *Biophys. J.* **20**, 193 (1977).
- [9] R. G. Endres and N. S. Wingreen, Maximum Likelihood and the Single Receptor, *Phys. Rev. Lett.* **103**, 158101 (2009).
- [10] W.-J. Rappel and H. Levine, Receptor Noise and Directional Sensing in Eukaryotic Chemotaxis, *Phys. Rev. Lett.* **100**, 228101 (2008).
- [11] W.-J. Rappel and H. Levine, Receptor noise limitations on chemotactic sensing, *Proc. Natl. Acad. Sci. USA* **105**, 19270 (2008).
- [12] Bo Hu, W. Chen, W.-J. Rappel, and H. Levine, Physical Limits on Cellular Sensing of spatial Gradients, *Phys. Rev. Lett.* **105**, 048104 (2010).
- [13] R. G. Endres and N. S. Wingreen, Accuracy of direct gradient sensing by single cells, *Proc. Natl. Acad. Sci. USA* **105**, 15749 (2008).
- [14] T. Smith, S. Fancher, A. Levchenko, I. Nemenman, and A. Mugler, Role of spatial averaging in multicellular gradient sensing, *Phys. Biol.* **13**, 035004 (2016).
- [15] T. Mora and N. S. Wingreen, Limits of Sensing Temporal Concentration Changes by Single Cells, *Phys. Rev. Lett.* **104**, 248101 (2010).
- [16] J.-B. Lalanne and P. François, Chemodetection in fluctuating environments: Receptor coupling, buffering, and antagonism, *Proc. Natl. Acad. Sci. USA* **112**, 1898 (2015).
- [17] M. Skoge, Y. Meir, and N. S. Wingreen, Dynamics of Cooperativity in Chemical Sensing Among Cell-Surface Receptors, *Phys. Rev. Lett.* **107**, 178101 (2011).
- [18] E. Roob III, N. Trendel, P. Rein ten Wolde, and A. Mugler, Cooperative clustering digitizes biochemical signaling and enhances its fidelity, *Biophys. J.* **110**, 1661 (2016).
- [19] V. Singh, M. Tchernookov, and I. Nemenman, Effects of receptor correlations on molecular information transmission, *Phys. Rev. E* **94**, 022425 (2016).
- [20] S. Fancher and A. Mugler, Fundamental Limits to Collective Concentration Sensing in Cell Populations, *Phys. Rev. Lett.* **118**, 078101 (2017).
- [21] A. Mugler, A. Levchenko, and I. Nemenman, Limits to the precision of gradient sensing with spatial communication and temporal integration, *Proc. Natl. Acad. Sci. USA* **113**, E689 (2016).
- [22] J. Varennes, S. Fancher, B. Han, and A. Mugler, Emergent Versus Individual-Based Multicellular Chemotaxis, *Phys. Rev. Lett.* **119**, 188101 (2017).
- [23] A. C. Barato, D. Hartich, and U. Seifert, Efficiency of cellular information processing, *New J. Phys.* **16**, 103024 (2014).
- [24] A. H. Lang, C. K. Fisher, T. Mora, and P. Mehta, Thermodynamics of Statistical Inference by Cells, *Phys. Rev. Lett.* **113**, 148103 (2014).
- [25] C. C. Govern and P. R. ten Wolde, Energy Dissipation and Noise Correlations in Biochemical Sensing, *Phys. Rev. Lett.* **113**, 258102 (2014).
- [26] G. Lan, P. Sartori, S. Neumann, V. Sourjik, and Y. Tu, The energy–speed–accuracy trade-off in sensory adaptation, *Nat. Phys.* **8**, 422 (2012).
- [27] P. Mehta and D. J. Schwab, Energetic costs of cellular computation, *Proc. Natl. Acad. Sci. USA* **109**, 17978 (2012).
- [28] C. C. Govern and P. R. ten Wolde, Optimal resource allocation in cellular sensing systems, *Proc. Natl. Acad. Sci. USA* **111**, 17486 (2014).
- [29] N. B. Becker, A. Mugler, and P. R. ten Wolde, Optimal Prediction by Cellular Signaling Networks, *Phys. Rev. Lett.* **115**, 258103 (2015).
- [30] E. D. Siggia and M. Vergassola, Decisions on the fly in cellular sensory systems, *Proc. Natl. Acad. Sci. USA* **110**, E3704 (2013).
- [31] J.-B. Lalanne and P. François, Principles of Adaptive Sorting Revealed by in Xilico Evolution, *Phys. Rev. Lett.* **110**, 218102 (2013).
- [32] P. François, G. Voisinne, E. D. Siggia, G. Altan-Bonnet, and M. Vergassola, Phenotypic model for early t-cell activation displaying sensitivity, specificity, and antagonism, *Proc. Natl. Acad. Sci. USA* **110**, E888 (2013).
- [33] P. François, M. Hemery, K. A. Johnson, and L. N. Saunders, Phenotypic spandrel: Absolute discrimination and ligand antagonism, *Phys. Biol.* **13**, 066011 (2016).
- [34] G. Lan, S. Schulmeister, V. Sourjik, and Y. Tu, Adapt locally and act globally: Strategy to maintain high chemoreceptor

- sensitivity in complex environments, *Mol. Syst. Biol.* **7**, 475 (2011).
- [35] T. Mora, Physical Limit to Concentration Sensing Amid Spurious Ligands, *Phys. Rev. Lett.* **115**, 038102 (2015).
- [36] J. J. Hopfield, Kinetic proofreading: A new mechanism for reducing errors in biosynthetic processes requiring high specificity, *Proc. Natl. Acad. Sci. USA* **71**, 4135 (1974).
- [37] V. Singh and I. Nemenman, Simple biochemical networks allow accurate sensing of multiple ligands with a single receptor, *PLoS Comput. Biol.* **13**, 1 (2017).
- [38] T. D. Mueller and J. Nickel, Promiscuity and specificity in BMP receptor activation, *FEBS Lett.* **586**, 1846 (2012).
- [39] Y. E. Antebi, J. M. Linton, H. Klumpe, B. Bintu, M. Gong, C. Su, R. McCardell, and M. B. Elowitz, Combinatorial signal perception in the BMP pathway, *Cell* **170**, 1184 (2017).
- [40] T. Kawai and S. Akira, Toll-like receptors and their crosstalk with other innate receptors in infection and immunity, *Immunity* **34**, 637 (2011).
- [41] A. K. Sewell, Why must t cells be cross-reactive? *Nat. Rev. Immunol.* **12**, 669 (2012).
- [42] D. Mason, A very high level of crossreactivity is an essential feature of the t-cell receptor, *Immunol. Today* **19**, 395 (1998).
- [43] M. Laub and M. Goulian, Specificity in two-component signal transduction pathways, *Ann. Rev. Genet.* **41**, 121.
- [44] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory* (Prentice Hall, Upper Saddle River, NJ, 2001).
- [45] B. B. Averbeck, P. E. Latham, and A. Pouget, Neural correlations, population coding and computation, *Nat. Rev. Neurosci.* **7**, 358 (2006).
- [46] L. Merchan and I. Nemenman, On the sufficiency of pairwise interactions in maximum entropy models of networks, *J. Stat. Phys.* **162**, 1294 (2016).
- [47] A. Wodarz and R. Nusse, Mechanisms of WNT signaling in development, *Annu. Cell Dev. Biol.* **14**, 59 (1998).
- [48] P. J. Murray, The JAK-STAT signaling pathway: Input and output integration, *J. Immunol.* **178**, 2623 (2007).
- [49] C. Jørgensen, A. Sherman, G. I. Chen, A. Pasculescu, A. Poliakov, M. Hsiung, B. Larsen, D. G. Wilkinson, R. Linding, and T. Pawson, Cell-specific information processing in segregating populations of Eph receptor Ephrin-expressing cells, *Science* **326**, 1502 (2009).
- [50] A. Mayer, V. Balasubramanian, T. Mora, and A. M. Walczak, How a well-adapted immune system is organized, *Proc. Natl. Acad. Sci. USA* **112**, 5950 (2015).
- [51] E. A. Hallett and J. R. Carlson, Coding of odors by a receptor repertoire, *Cell* **125**, 143 (2006).
- [52] B. Malnic, J. Hirono, T. Sato, and L. B. Buck, Combinatorial receptor codes for odors, *Cell* **96**, 713 (1999).
- [53] K. Krishnamurthy, A. M. Hermundstad, T. Mora, A. M. Walczak, and V. Balasubramanian, Disorder and the neural representation of complex odors: Smelling in the real world, [bioRxiv](https://arxiv.org/abs/10.1101/160382), doi:10.1101/160382 (2017).
- [54] T. McKeithan, Kinetic proofreading in T-cell receptor signal transduction, *Proc. Natl. Acad. Sci. USA* **92**, 5042 (1995).