Physical limit to concentration sensing amid spurious ligands

Thierry Mora[1]

[1]*Laboratoire de physique statistique, École normale supérieure, CNRS and UPMC, 24 rue Lhomond, 75005 Paris, France*

(Dated: July 23, 2015)

To adapt their behaviour in changing environments, cells sense concentrations by binding external ligands to their receptors. However, incorrect ligands may bind nonspecifically to receptors, and when their concentration is large, this binding activity may interfere with the sensing of the ligand of interest. Here, I derive analytically the physical limit to the accuracy of concentration sensing amid a large number of interfering ligands. A scaling transition is found when the mean bound time of correct ligands is twice that of incorrect ligands. I discuss how the physical bound can be approached by a cascade of receptor states generalizing kinetic proof-reading schemes.

Because of their small sizes, biological systems typically operate with only a few copies of the molecules they sense and communicate with. In their pioneering work, Berg and Purcell derived the fundamental bound that the noise arising from these small numbers sets on the accuracy of concentration sensing [1]. Experimental progress in the characterization of single-cell variability [2] and sensing precision [3] has fueled a renewed interest in small-number noise and its implications for information processing [4–6]. General or refined bounds on sensing accuracy have been recently derived for single receptors [7–9], and extended to spatial [10–14] or temporal [15] gradient sensing, while the metabolic cost and trade-offs of sensing accuracy have been explored [16–25]. Much of this past work has assumed perfect specificity between the biological receptors and their cognate ligands. In realistic biological contexts, large numbers of spurious ligands may bind receptors nonspecifically, interfering with the ligand of interest [26]. This is the case in the problem of antigen recognition by T-cell receptors, where cells must react to a small number of specific foreign peptides among a large number of nonspecific self-peptides [27]. Biochemical network architectures based on kinetic proofreading [28, 29] have been shown to provide a solution to the discrimination problem, and have been studied in depth theoretically [26, 30–32]. However, no fundamental bound has been derived against which to compare the performance of these solutions, save for Ref. [33] where concepts of statistical decision theory were used to derive the minimal decision time to detect cognate ligands. In this paper I derive the fundamental limit on concentration sensing accuracy and ligand detection error in the presence of a large number of spurious ligands. The maximum likelihood estimate achieving the bound can be implemented biologically by simple networks based on push-pull reactions.

Consider a mixture of two ligands, only one of which the biological system wishes to sense. The ligand of interest (hereafter referred to as the correct ligand) is present in concentration c, while the interfering or spurious ligand (called the incorrect ligand) is present in concentration c'. The biological unit can sense ligands through N identical receptors, which can be bound by either ligand with a common rate $k = 4Da$, where D is the molecule diffusivity and a the effective receptor size. Receptors can distinguish between the two molecules thanks to their higher affinity to the correct ligand. Physically, this means that the unbinding rate r of the correct ligand is smaller than that of the incorrect ligand $r' > r$.

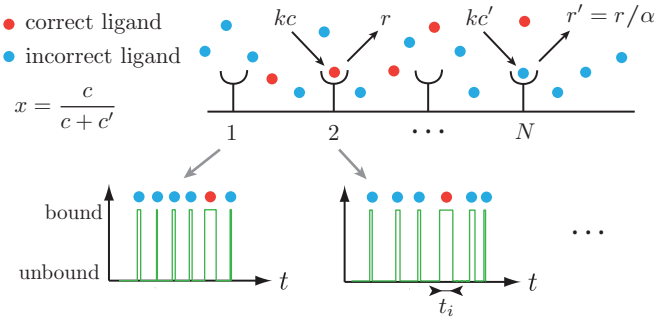

FIG. 1: **Reading concentrations off trajectories of receptor occupancies.** Cognate (correct) and spurious (incorrect) ligands may bind N receptors, typically presented on the cell surface, with rate kc and kc'. The incorrect ligands are in excess, $c' > c$, but lead to shorter binding events, $r' > r$. The information the cell can theoretically use is contained in the time traces of occupancy of all receptors (green curves). The maximum likelihood estimate fully exploits these traces to optimally infer the input concentrations c and c'.

The occupancy of each receptor,

$$p = \frac{kcr^{-1} + kc'r'^{-1}}{1 + kcr^{-1} + kc'r'^{-1}}, \tag{1}$$

depends on both concentrations, and cannot be used alone to determine c. The interchangeability of the ratios c/r and c'/r' in this expression emphasizes the ambiguity between many incorrect ligands and a few correct ones. To discern these two effects, one must use the full temporal record of occupancy of each receptor. The probability distribution for the binding and unbinding events at all receptors during a time interval T reads:

$$P = e^{-kc_{\rm tot}T_{\rm u}} \prod_{i=1}^{n} \left(kcre^{-rt_i} + kc'r'e^{-r't_i} \right), \tag{2}$$

with $c_{\rm tot} = c + c'$ is the total concentration of ligands, T_u is the total unbound time accrued over all receptors, and $t_1, \ldots, t_n$ the durations of the n binding events occurring at all N receptors during T. The log-likelihod $\mathcal{L} = \ln P$ can be rewritten as a sum of three independent contributions, $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_1 + \mathcal{L}_2$, where $\mathcal{L}_0$ depends on neither c or c', and where $\mathcal{L}_1$ and $\mathcal{L}_2$ pertain to the unbound and bound intervals respectively:

$$\mathcal{L}_1(c_{\rm tot}) = n \ln c_{\rm tot} - k c_{\rm tot} T_{\rm u} \tag{3}$$

$$\mathcal{L}_2(x) = \sum_{i=1}^{n} \ln\left(1 - x + x\alpha e^{(1-\alpha) r' t_i}\right), \tag{4}$$

where $x = c/c_{\rm tot}$ is the fraction of correct ligands, and $\alpha = r/r' < 1$ is the binding constant ratio. As can be seen in the respective dependencies of $\mathcal{L}_1$ and $\mathcal{L}_2$ upon $c_{\rm tot}$ and x, unbound intervals are informative of the total concentration, while the bound intervals are informative of the fractions of ligands.

The maximum likelihood estimate for the total concentration is obtained by the condition $\partial \mathcal{L}_1 / \partial c_{\rm tot} = 0$, which gives $c^*_{\rm tot} = k T_u / n$. The error made by this estimate is given in the large time limit by the Cramér-Rao bound, which sets the best possible performance of any estimator [34]:

$$\langle \delta c_{\rm tot}^2 \rangle \approx -\left(\frac{\partial^2 \mathcal{L}_1}{\partial c_{\rm tot}^2}\right)^{-1} = \frac{c_{\rm tot}^2}{n} \approx \frac{c_{\rm tot}}{4Da(1-p)NT}, \tag{5}$$

with $\delta c_{\rm tot} = c^*_{\rm tot} - c_{\rm tot}$. This result is that obtained in [8] for a single ligand, where the maximum-likelihood error was shown to be half as small as the classical Berg and Purcell bound [1] based on the average receptor occupancy. The reason for this difference is that the maximum likelihood estimate is not affected by the noise due to the stochastic nature of receptor unbinding, as evident in Eq. (3). In the case of a mixture, the receptor occupancy Eq. (1) depends on x as well as $c_{\rm tot}$, and does not even suffice to determine the total concentration.

The fraction x of correct ligands can be estimated by maximum likelihood as well, by solving:

$$\left.\frac{\partial \mathcal{L}_2}{\partial x}\right|_{x^*} = \sum_{i=1}^{n} \frac{\alpha e^{(1-\alpha) r' t_i} - 1}{1 - x^* + x^* \alpha e^{(1-\alpha) r' t_i}} = 0. \tag{6}$$

The error can be estimated from the Cramér-Rao bound (App. A.1):

$$\langle \delta x^2 \rangle \approx -\left(\frac{\partial^2 \mathcal{L}_2}{\partial x^2}\right)^{-1} \approx \frac{f(x,\alpha)}{n}$$
$$\text{with} \quad f(x,\alpha)^{-1} = \int_0^{+\infty} du\, e^{-u} \frac{(\alpha e^{(1-\alpha)u} - 1)^2}{1 - x + x\alpha e^{(1-\alpha)u}}. \tag{7}$$

The total error in the concentration of the correct ligand $c = x c_{\rm tot}$ is then the sum of the (independent) errors in $c_{\rm tot}$ and x from Eqs. (5) and (7): $\langle \delta c^2 \rangle \approx c_{\rm tot}^2 (x^2 + f(x,\alpha))/n$.

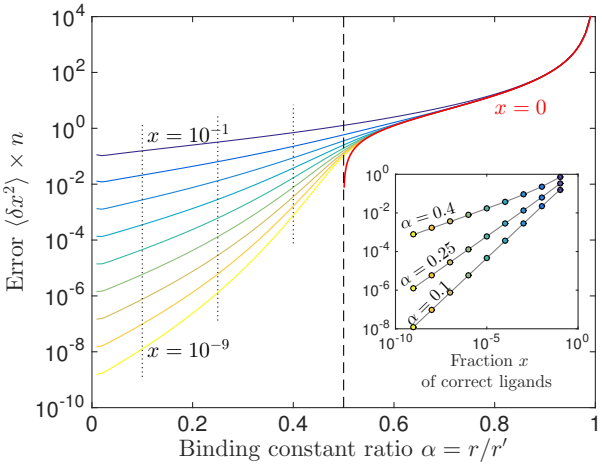

FIG. 2: **Physical bound on concentration sensing error.** The fundamental bound on the relative error in the fraction of correct ligands $x = c/c_{\rm tot}$ scales with the inverse of the number of binding events n. Here the rescaled error $f(x,\alpha) = n\langle \delta x^2 \rangle$ [Eq. (7)], is represented as a function of the binding constant ratio α for various values of x. There are two distinct scaling regimes [Eq. (8)]. For $\alpha > 1/2$, the rescaled error depends weakly on x, while for $\alpha < 1/2$ it scales as x^β, with $\beta = 1 - \alpha/(1-\alpha)$, as illustrated in the inset for three example values of α.

It is interesting to consider the limit where the correct ligands are rare, $x \ll 1$, as in the case of immune recognition. Two scaling regimes, illustrated in Fig. 2, are found depending on the value of the ratio α between the two binding constants:

$$f(x,\alpha) \approx \begin{cases} g(\alpha) & \alpha > 1/2, \\ h(\alpha) x^\beta, \ \beta = 1 - \frac{\alpha}{1-\alpha} & \alpha < 1/2, \end{cases} \tag{8}$$

with $g(\alpha) = (2\alpha - 1)/(1-\alpha)^2$ and $h(\alpha) = (1-\alpha)\alpha^{-\frac{1}{1-\alpha}} \sin(\pi\alpha/(1-\alpha))/\pi$, and $0 < \beta < 1$. Since $x^2 \ll f(x,\alpha)$, the error in c reduces to:

$$\begin{aligned} \langle \delta c^2 \rangle &\approx g(\alpha) \frac{c_{\rm tot}}{4Da(1-p)NT} \quad \alpha > 1/2, \\ \langle \delta c^2 \rangle &\approx h(\alpha) \frac{c^\beta c_{\rm tot}^{1-\beta}}{4Da(1-p)NT} \quad \alpha < 1/2. \end{aligned} \tag{9}$$

In the hard discrimination regime ($\alpha > 1/2$), incorrect ligands dominate the error, which is governed by $c_{\rm tot}$ as in Eq. (5). The prefactor $g(\alpha)$ diverges at $\alpha = 1$, as expected when the two ligands have the same binding constant and are thus indistinguishable. By contrast, in the easy discrimination regime ($\alpha < 1/2$) the error is governed by a weighted geometric mean between c and $c_{\rm tot}$. In the limit of very small α, corresponding to clearly distinguishable ligands, the error is $\langle \delta c^2 \rangle \approx c/[4Da(1-p)NT)]$—precisely the error when no interfering ligand is present [8]. For $x = 0$, the maximum-likelihood estimate may infer a small $x^* = \delta x > 0$. However, the second derivative of the likelihood diverges at

$x = 0$ for $\alpha < 1/2$, indicating that the Cramér-Rao bound (7) fails to give a correct estimate of this error, which instead scales anomalously with the number of events: $\delta x \sim n^{\alpha-1}$, hence $\delta c \sim c_{\rm tot}^{\alpha}[4Da(1-p)NT]^{\alpha-1}$ (App. A.2).

In many situations, it is more useful for the system to determine the presence of the correct ligand rather than its precise concentration, as in the recognition of foreign pathogens by immune receptors. This decision can be made optimally (in the Bayesian sense) by comparing the likelihoods of the two competing hypotheses: presence *versus* absence of the correct ligand at fraction x. The presence of the correct ligand is detected when $\ln[P(\{t_i\}|x)/P(\{t_i\}|0)] = \mathcal{L}_2(x) > \theta$, where θ is an adjustable parameter controlling the balance between the false-positive and false-negative error rates FP and FN. These errors decay exponentially fast with large numbers n of binding events, and can be estimated in that limit using a saddle-point approximation (App. B1):

$$FP \approx \frac{\exp\left[n\phi(\lambda) - \lambda\theta\right]}{\lambda\sqrt{2\pi n|\phi''(\lambda)|}}, \quad FN \approx \frac{\lambda}{1-\lambda}e^{\theta}FP, \qquad (10)$$

where $\phi(\lambda) = \ln\left[\int_0^{+\infty} du\, e^{-u}[1 - x + \alpha x e^{(1-\alpha)u}]^{\lambda}\right]$, and where λ satisfies the saddle-point condition $\theta = n\phi'(\lambda)$. The receiver operating characteristics (ROC) giving the dependency between FP and FN can thus be estimated parametrically by varying λ. This saddle-point approximation is well verified by numerical simulations (Fig. 3).

As in the case of concentration sensing error, a scaling transition is found in the limit of scarce correct ligands, $x \ll 1$. When $\alpha > 1/2$, one obtains

$$FP \approx \frac{e^{-\frac{1}{2}\lambda^2 nx^2/g(\alpha)}}{\sqrt{2\pi\lambda^2 nx^2/g(\alpha)}}, \; FN \approx \frac{e^{-\frac{1}{2}(1-\lambda)^2 nx^2/g(\alpha)}}{\sqrt{2\pi(1-\lambda)^2 nx^2/g(\alpha)}}, \qquad (11)$$

while when $\alpha < 1/2$ both error rates decay as $\sim (nx^{\gamma})^{-1/2}\exp[-Cnx^{\gamma}]$, with $\gamma = (1-\alpha)^{-1}$, $1 < \gamma < 2$ and C a function of α and λ (App. B.2). The time T necessary to make a reliable decision scales as $[4Da(1-p)N]^{-1}c_{\rm tot}c^{-2}$ for $\alpha > 1/2$, and as $[4Da(1-p)N]^{-1}c_{\rm tot}^{\gamma-1}c^{-\gamma}$ for $\alpha < 1/2$. Equivalent scaling laws were obtained in [33] for minimal on-the-fly detection times.

Can biological systems approach the physical bound on concentration sensing given by Eq. (7)? To gain insight into this question, one can expand Eq. (6) at first order in x to get an approximation to the maximum likelihood estimate when $\alpha > 1/2$ (for $\alpha < 1/2$ this expansion gives quantities with diverging means and cannot be used):

$$x^* \approx \frac{2\alpha - 1}{(1-\alpha)^2}\frac{1}{n}\sum_{i=1}^{n}\left(\alpha e^{(1-\alpha)r' t_i} - 1\right). \qquad (12)$$

This estimator, which is subject to the same asymptotic error as in Eq. (8), suggests a simple strategy, where each receptor signals "positively" with a rate that depends on how long it has been bound, $\alpha(1-\alpha)r'e^{(1-\alpha)r't}$, and "negatively" (*i.e.* with an opposite effect on the readout, see below) through a fixed burst $(\alpha-1)\delta(t)$ upon binding, so that the net effect of each binding event i on the readout molecule concentration is

$$\int_0^{t_i} dt\left[\alpha(1-\alpha)r'e^{(1-\alpha)r't}\right] + \alpha - 1 = \alpha e^{(1-\alpha)r't_i} - 1, \qquad (13)$$

i.e. exactly the argument of the sum in Eq. (12).

This idea can be implemented biologically by a cascade of receptor conformational states triggered by binding, and proceeding irreversibly from states 1 to m, each transition to the next state occurring with rate s (Fig. 4). The ligand is free to detach from the receptor at any time, bringing the receptor back to the unbound state 0. The receptors signal through the production or activation of two molecules B and D with opposite effects on a push-pull network governing the state of a molecule X, which provides the final readout for x through its modified state X^*. If one requires that the equilibration of B and D are fast, and that X and X^* are always in excess in the Michaelis-Mentens reactions, then

$$\frac{dX^*}{dt} = X_0\sum_{j=1}^{N}(b_{\mu(j)} - d_{\mu(j)}), \quad X_0 = \text{const}, \qquad (14)$$

where $\mu(j)$ is the state of the $j^{\rm th}$ receptor, and $b_0 = d_0 = 0$. For the purpose of this discussion, the internal molecules B, D and X are assumed to be unaffected by biochemical noise, restricting the source of noise to the input alone. In this design X^* increases indefinitely to mimick the sum in Eq. (12) over all events at all receptors. A more a realistic but equivalent scheme would

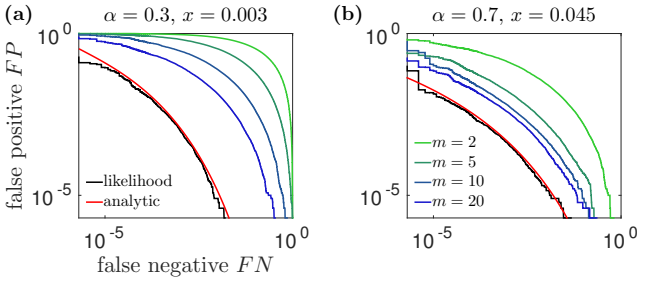

FIG. 3: **Error in the detection of the correct ligand.** Numerical and analytical estimate of the rate of false positive (FP) versus false negative (FN) errors in the detection of a small fraction x of correct ligands, for $n = 10^5$ and (a) $\alpha = 0.3$ and $x = 0.003$ and (b) $\alpha = 0.7$ and $x = 0.045$. The black curve is the result of a numerical experiment, repeated $5 \cdot 10^5$ times in presence of the correct ligand, and $5 \cdot 10^5$ without, and where a likelihood ratio test $\mathcal{L}_2(x) > \theta$ was used with a varying threshold θ. The red curve is the analytical prediction from Eq. (10). The green to blue curves show the performance of optimized networks schematized in Fig. 4, for various numbers of receptor states m.

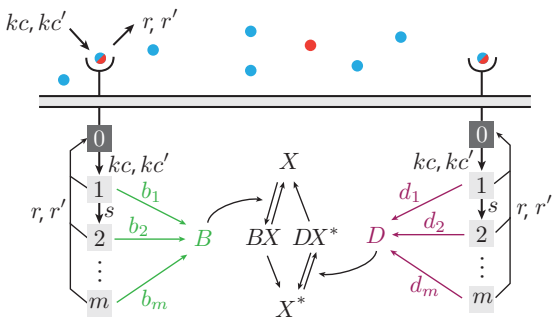

FIG. 4: **Network for sensing the concentration of correct ligands.** Upon binding, each receptor enters a cascade of m states along which it proceeds with rate s. The ligand can detach at any moment with rate r or r' depending on its identity (correct or incorrect), bringing the receptor back to the unbound state 0. While in conformational state $\mu = 1, \ldots, m$, the receptor activates two enzymes B and D with rates b_μ and d_μ, each catalyzing two oppposite Michaelis-Mentens reactions of a pull-push network. B and D are assumed to equilibrate fast and to be always limiting in the reactions they catalyze, so that $dX^*/dt \propto b_\mu - d_\mu$.

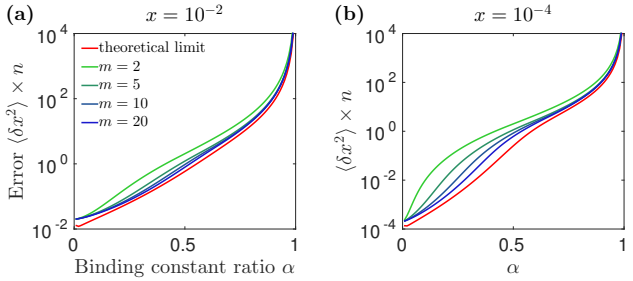

FIG. 5: **Network performance.** Error made by optimized networks with a finite number of receptor states m (green to blue curves), compared to the theoretical bound (red curve), for (a) $x = 10^{-2}$ and (b) $x = 10^{-4}$.

involve a running sum over an effective time T, obtained by relaxing X^* to X with rate $\sim T^{-1}$ [15].

When the number of states m is large and the transitions between them are rapid, X^* can track Eq. (12) with arbitrary precision when $\alpha > 1/2$. In that case, the receptor state μ provides an approximation to the time since binding, $\mu \approx st$. Then, for example, receptors signaling positively with rate $b_\mu \propto \alpha(1-\alpha)r'e^{(1-\alpha)r'\mu/s} \approx \alpha(1-\alpha)r'e^{(1-\alpha)r't}$, and negatively with rate $d_\mu \propto (1-\alpha)(s/\mu_0)e^{-\mu/\mu_0} \approx (1-\alpha)\delta(t)$ (with μ_0 an adjustable parameter) would exactly realize Eq. (13) and thus the estimator of Eq. (12) in the limit $m \gg s/r \gg \mu_0 \gg 1$.

Although such optimal performance is only reached for large m and $\alpha > 1/2$, this network design may still perform well in more general situations. One can optimize the expected error produced by this network over the net signaling rates $(b_\mu - d_\mu)$, with the constraint that the mean effect of binding incorrect ligands on X^* be zero, so that $\Delta X^* \propto c$ on average (App. C). Fig. (5) shows how the performance of such optimized networks approaches the theoretical bound as the number of states m increases. The convergence is significantly worse for $\alpha < 1/2$ at small x. In that regime, the estimator of (12) is not valid, suggesting that this network design may not achieve the optimal bound even with an infinite number of states. The output of these networks can also be used to detect ligands. Their performance in doing so is compared to the optimal discrimination errors of Eq. (10) in Fig. (3).

The principle of maximum likelihood not only yields the fundamental bound on the accuracy of discerning cognate ligands from spurious ones, but also suggests biochemical solutions to approach this optimal bound. Such maximum-likelihood inspired designs have been previously proposed in the case of a single ligand [15, 19]. The network structure proposed in this study (Fig. 4) is reminiscent of kinetic proofreading schemes and their generalizations, which provide a well-known solution to the ligand discrimination problem [26, 28, 29, 32, 35]. An important difference is that here signaling occurs during all steps, albeit at various, fine-tuned rates, and with potentially negative contributions, the role of which is to buffer the effect of wrong ligands. Consistent with this prediction, it was shown that a negative interaction through a diffusible molecule between kinetic-proofreading receptors could mitigate the effects of large numbers of incorrect ligands in a discrimination task [26].

The present results are relevant beyond the particular case of sensing by receptors, and apply to any kind of biochemical signaling in presence of competing ligands or "cross-talk." This is the case for example in the context of gene regulation, where competing transcription factors may bind regulatory sites unspecifically, a problem particularly acute in metazoans [36].

The scaling transition occurring at the binding constant ratio $\alpha = 1/2$ suggests that different strategies should be employed depending on how hard the discrimination task is. In particular, the approximate but biologically implementable estimator of Eq. (12) curiously breaks down in the easy discrimination regime, $\alpha < 1/2$. In that regime, the optimal bound is harder to achieve because it is dominated by rare, long binding events that are hard to encode by biochemical solutions. The example of immune recognition falls precisely into that regime, with a binding constant ratio α between agonist and nonagonist ligands ranging from one fifth to one third [27]. More elaborate network designs, probably with feedback, may be needed to achieve the theoretical bound Eq. (7) in that case. Finally, this study has assumed throughout that the unbinding rates r and r' are priorly known to the system. Complex mixtures of ligands with unknown binding constants would make for interesting generalizations.

I thank A. Walczak for her helpful comments on the manuscript. While this article was under review, a paper treating a similar topic was submitted to the arXiv [37].

Appendix A: Cramér-Rao bound

1. The Cramér-Rao bound is tight: a physicist's proof

In general the Cramér-Rao bound is a lower bound on the error made by any unbiased estimator, but it is not always certain whether this bound can be achieved. Here the maximum likelihood estimate is shown to approach the Cramér-Rao bound in the limit of large samples.

Assume that the likelihood of the data factorizes over independent datapoints,

$$\mathcal{L} = \ln P = \sum_{i=1}^{n} \ell(x, t_i), \tag{A1}$$

where x is the model parameter to be estimated, and $(t_1, \ldots, t_n)$ the series of datapoints. In the specific case of receptors binding to two types of ligands, x is the fraction of correct ligands, t_i the duration of binding event i, and

$$\ell(x, t_i) = \ln r' - r' t_i + \ln\left(1 - x + x\alpha e^{(1-\alpha) r' t_i}\right). \tag{A2}$$

The derivative of ℓ with respect to x is denoted by $\ell'(x, t_i) = \partial \ell(x, t_i)/\partial x$. The maximum likelihood estimate x^* satisfies:

$$\sum_{i=1}^{n} \ell'(x^*, t_i) = 0. \tag{A3}$$

This estimator is unbiased: if $\tilde{x}$ denotes the true parameter with which the data was generated, then x^* should give back $\tilde{x}$ on average. Equivalently,

$$\begin{aligned}\left\langle \frac{\partial \mathcal{L}(\tilde{x})}{\partial x} \right\rangle_{\tilde{x}} &= \sum_{i=1}^{n} \langle \ell'(\tilde{x}, t_i) \rangle_{\tilde{x}} = n \int_0^{+\infty} dt\, e^{\ell(\tilde{x},t)} \ell'(\tilde{x}, t_i) \\ &= n \frac{\partial}{\partial x} \int_0^{+\infty} dt\, e^{\ell(x,t)} \Big|_{\tilde{x}} = 0,\end{aligned} \tag{A4}$$

(the last integral is just 1 because of normalization), where $\langle \cdot \rangle_{\tilde{x}}$ denote averages over data generated with the true parameter $\tilde{x}$. In other words, the maximum of $\mathcal{L}$ is reached at $\tilde{x}$ on average. The probability that this maximum x^* be larger than a certain value $x > \tilde{x}$ is:

$$\mathbb{P}(x^* > x) = \mathbb{P}\left(\frac{\partial \mathcal{L}(x)}{\partial x} > 0 \right) = \left\langle \Theta\left(\frac{\partial \mathcal{L}}{\partial x} \right) \right\rangle_{\tilde{x}}. \tag{A5}$$

The Heaviside function Θ can be replaced by its Fourier representation:

$$\Theta(x) = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi\omega} e^{i\omega x} = \int_{-i\infty}^{+i\infty} \frac{d\lambda}{2\pi i \lambda} e^{\lambda x}, \tag{A6}$$

allowing for factorization over datapoints:

$$\begin{aligned}\mathbb{P}(x^* > x) &= \int_{-i\infty}^{+i\infty} \frac{d\lambda}{2\pi i \lambda} \prod_{i=1}^{n} \int_0^{+\infty} dt_i e^{\ell(\tilde{x}, t_i) + \lambda \ell'(x, t_i)} \\ &= \int_{-i\infty}^{+i\infty} \frac{d\lambda}{2\pi i \lambda} \exp\left[n \ln \int_0^{+\infty} dt e^{\ell(\tilde{x},t) + \lambda \ell'(x,t)} \right].\end{aligned} \tag{A7}$$

and

$$\begin{aligned}P(x) &= -\frac{d\mathbb{P}(x^* > x)}{dx} = \\ &= -\int_{-i\infty}^{+i\infty} \frac{d\lambda}{2\pi i} \left[\int_0^{+\infty} dt \ell''(x,t) e^{\ell(\tilde{x},t) + \lambda \ell'(x,t)} \right] \\ &\quad \times \exp\left[(n-1) \ln \int_0^{+\infty} dt e^{\ell(\tilde{x},t) + \lambda \ell'(x,t)} \right].\end{aligned} \tag{A8}$$

This integral can be evaluated by a saddle-point approximation in the large n limit:

$$\int d\lambda G(\lambda) e^{nF(\lambda)} \approx G(\lambda^*) \sqrt{\frac{2\pi}{n|F''(\lambda^*)|}} e^{nF(\lambda^*)}, \tag{A9}$$

with

$$G(\lambda) = \frac{\int_0^{+\infty} dt \ell''(x,t) e^{\ell(\tilde{x},t) + \lambda \ell'(x,t)}}{\int_0^{+\infty} dt e^{\ell(\tilde{x},t) + \lambda \ell'(x,t)}}, \tag{A10}$$

$$F(\lambda) = \ln \int_0^{+\infty} dt e^{\ell(\tilde{x},t) + \lambda \ell'(x,t)}. \tag{A11}$$

The saddle λ^* is given by the condition that the derivative of the argument of the exponential with respect to λ be zero:

$$\int_0^{+\infty} dt \ell'(x,t) e^{\ell(\tilde{x},t) + \lambda \ell'(x,t)} = 0. \tag{A12}$$

At $x = \tilde{x}$, this condition is satisfied for $\lambda = 0$. In the limit of large samples, $x^* - \tilde{x}$ is small and so should the corresponding λ. One can expand at small $x - \tilde{x} > 0$ and λ:

$$\ell'(x,t) \approx \ell'(\tilde{x},t) + (x - \tilde{x}) \ell''(\tilde{x},t), \tag{A13}$$

and

$$\int_0^{+\infty} dt \left[(x - \tilde{x}) \ell''(\tilde{x},t) + \lambda \ell'(\tilde{x},t)^2 \right] e^{\ell(\tilde{x},t)} = 0, \tag{A14}$$

yielding $\lambda = x - \tilde{x}$ and:

$$P(x) \approx \frac{\sqrt{nH}}{\sqrt{2\pi}} \exp\left[-\frac{n}{2} (x - \tilde{x})^2 H \right], \tag{A15}$$

with

$$H = -\int_0^{+\infty} dt \ell''(\tilde{x},t) e^{\ell(\tilde{x},t)} = \int_0^{+\infty} dt \ell'(\tilde{x},t)^2 e^{\ell(\tilde{x},t)}. \tag{A16}$$

A symmetric argument gives the same result for $x < \tilde{x}$. The resulting distribution of x^* is Gaussian, with mean $\tilde{x}$ and variance

$$\langle \delta x^2 \rangle = \frac{1}{nH} = -\left\langle \frac{\partial^2 \mathcal{L}(\tilde{x})}{\partial x^2} \right\rangle_{\tilde{x}}^{-1} = \left\langle \left(\frac{\partial \mathcal{L}(\tilde{x})}{\partial x} \right)^2 \right\rangle_{\tilde{x}}^{-1}. \tag{A17}$$

In the specific case of Eq. (A2),

$$H = \int_0^{+\infty} dt r' e^{-r't} \frac{(\alpha e^{(1-\alpha)r't} - 1)^2}{1 + x(\alpha e^{(1-\alpha)r't} - 1)}. \tag{A18}$$

Performing the change of variable $u = r't$ yields the result of the main text:

$$\langle \delta x^2 \rangle \approx \frac{f(x,\alpha)}{n}$$
$$\text{with} \quad f(x,\alpha)^{-1} = \int_0^{+\infty} du\, e^{-u} \frac{(\alpha e^{(1-\alpha)u} - 1)^2}{1 - x + x\alpha e^{(1-\alpha)u}}. \tag{A19}$$

2. Small x limit

For $\alpha > 1/2$ and small x, the integral in Eq. (A19) can be approximated by:

$$f(x,\alpha)^{-1} \approx \int_0^{+\infty} du\, e^{-u} (\alpha e^{(1-\alpha)u} - 1)^2 = \frac{(1-\alpha)^2}{2\alpha - 1}. \tag{A20}$$

For $\alpha < 1/2$, the function $e^{-u}(\alpha e^{(1-\alpha)u} - 1)^2$ is not integrable, and the denominator of Eq. (A19) is necessary to ensure integrability at large t, however small x is. Thus, the values of u governing the behavior of the integral satisfy $x\alpha e^{(1-\alpha)u} = O(1)$. This observation suggests the change of variable $y = x\alpha e^{(1-\alpha)u}$:

$$f(x,\alpha)^{-1} = x^{-\beta} \alpha^{\frac{1}{1-\alpha}} \frac{1}{1-\alpha} \int_{\alpha x}^{+\infty} dy\, y^{-\frac{2-\alpha}{1-\alpha}} \frac{(y-x)^2}{1+y-x}, \tag{A21}$$

with $\beta = 1 - \alpha/(1-\alpha)$. Expanding $(y-x)^2$ gives three terms scaling as $x^{-\beta} y^{-\alpha/(1-\alpha)}$, $x^{1-\beta} y^{-1/(1-\alpha)}$ and $x^{2-\beta} y^{-\frac{2-\alpha}{1-\alpha}}$ at small y, respectively. The last two give diverging integrals as $x \to 0$ for all α, yielding terms of order 1. Only when $\alpha > 1/2$ does the first term give a diverging integral, and thus a term of order 1 in x; in that case, the sum of all three terms gives back the result of Eq. (A20). If $\alpha < 1/2$ however, the first term is integrable and thus dominates the expression for $x \ll 1$, yielding:

$$f(x,\alpha)^{-1} = x^{-\beta} \alpha^{\frac{1}{1-\alpha}} \frac{1}{1-\alpha} \int_0^{+\infty} dy\, \frac{y^{-\frac{\alpha}{1-\alpha}}}{1+y} + O(1), \tag{A22}$$

where $O(1)$ denotes a term of order 1 at small x. The integral can be calculated:

$$\int_0^{+\infty} dy\, \frac{y^{-\frac{\alpha}{1-\alpha}}}{1+y} = \frac{\pi}{\sin\left(\pi \frac{\alpha}{1-\alpha}\right)}, \tag{A23}$$

to finally obtain:

$$f(x,\alpha)^{-1} \approx x^{-\beta} \alpha^{\frac{1}{1-\alpha}} \frac{1}{1-\alpha} \frac{\pi}{\sin\left(\pi \frac{\alpha}{1-\alpha}\right)}. \tag{A24}$$

In the intermediate case $\alpha = 1/2$, the three terms in the integral of Eq. (A21) are of order y^{-1}, xy^{-2} and $x^2 y^{-3}$. Again the last two terms diverge in the integral and give contributions of order 1. The first term also diverges, but its contribution reads:

$$\alpha^{\frac{1}{1-\alpha}} \frac{1}{1-\alpha} (-\ln(\alpha x)) = \frac{1}{2} |\ln(x/2)|, \tag{A25}$$

so that:

$$f(x,\alpha)^{-1} = \frac{1}{2} |\ln(x/2)| + O(1). \tag{A26}$$

When $x = 0$ and $\alpha \leq 1/2$, $f(x,\alpha) = \infty$, as the large deviation function of x^* becomes nonanalytic. The expansion of ℓ' in Eq. (A13) is no longer integrable when done around $\tilde{x} = 0$, and needs revisiting. The integral in Eq. (A7) reads:

$$\int_0^{+\infty} du e^{-u} \exp\left[\frac{\lambda(\alpha e^{(1-\alpha)u} - 1)}{1 + x(\alpha e^{(1-\alpha)u} - 1)}\right] = 1 + \int_0^{+\infty} du e^{-u} \left\{ \exp\left[\frac{\lambda(\alpha e^{(1-\alpha)u} - 1)}{1 + x(\alpha e^{(1-\alpha)u} - 1)}\right] - 1 - \lambda(\alpha e^{(1-\alpha)u} - 1) \right\} \tag{A27}$$

where the same change of variable $u = r't$ has been done. Doing a further change of variable to $y = \alpha x e^{(1-\alpha)u}$ yields:

$$1 + \frac{(\alpha x)^{\frac{1}{1-\alpha}}}{1-\alpha} \int_{\alpha x}^{+\infty} y^{-\frac{2-\alpha}{1-\alpha}} \left\{ \exp\left[\frac{(\lambda/x)(y-x)}{1+y-x}\right] - 1 - (\lambda/x)(y-x) \right\} \tag{A28}$$

The term is the brackets is of order $(y-x)^2$, as was the case in Eq. (A19). Hence, terms in y^2 are integrable and

dominate the expression, which becomes at leading order in x, λ:

$$1+\frac{(\alpha x)^{\frac{1}{1-\alpha}}}{1-\alpha}\int_0^{+\infty} y^{-\frac{2-\alpha}{1-\alpha}}\left\{\exp\left[\frac{(\lambda/x)y}{1+y}\right]-1-(\lambda/x)y\right\} \tag{A29}$$

With $\tilde{\lambda}=\lambda x$, the saddle-point condition becomes:

$$\psi'(\tilde{\lambda},\alpha)=0, \tag{A30}$$

with

$$\psi(\tilde{\lambda},\alpha)=-\frac{\alpha^{\frac{1}{1-\alpha}}}{1-\alpha}\int_0^{+\infty} y^{-\frac{2-\alpha}{1-\alpha}}\left(e^{\frac{\tilde{\lambda}y}{1+y}}-1-\tilde{\lambda}y\right) \tag{A31}$$

and the cumulative probability distribution is:

$$\mathbb{P}(x^*>x)\approx\frac{1}{\sqrt{2\pi n x^{\frac{1}{1-\alpha}}\psi''(\tilde{\lambda},\alpha)\tilde{\lambda}}}\exp\left[-nx^{\frac{1}{1-\alpha}}\psi(\tilde{\lambda},\alpha)\right], \tag{A32}$$

where $\psi'=\partial\psi/\partial\tilde{\lambda}$ and $\psi''=\partial^2\psi/\partial\tilde{\lambda}^2$. Fluctuation of x^* are thus of order $n^{\alpha-1}\ll 1/\sqrt{n}$.

When $\alpha=1/2$, the term of order $(y-x)^2$ in the brackets of Eq. (A28) dominates and diverges, so that this expression reduces at leading order to:

$$1+\frac{x^2}{2}|\ln(x/2)|\left(\frac{\tilde{\lambda}^2}{2}-\tilde{\lambda}\right). \tag{A33}$$

The saddle point condition gives $\tilde{\lambda}=1$ and one obtains:

$$\mathbb{P}(x^*>x)\approx\frac{1}{\sqrt{\pi n x^2|\ln(x/2)|}}\exp\left[-\frac{nx^2|\ln(x/2)|}{4}\right], \tag{A34}$$

which implies fluctuations of order $\delta x\sim(n\ln n)^{-1/2}$.

Appendix B: Probability of discrimination error

1. General case

The discrimination between two competing hypotheses—presence *versus* absence of the correct ligand in fraction x—can be performed by a likelihood ratio test:

$$\ln\frac{P(t_1,\ldots,t_n|x)}{P(t_1,\ldots,t_n|x=0)}=\sum_{i=1}^{n}\left[\ell(x,t_i)-\ell(0,t)\right]>\theta, \tag{B1}$$

where θ is an adjustable parameter. The false-positive and false-negative error rates are defined as the probabilities of detecting the presence of a ligand that is in fact absent, and of missing it where it is there:

$$\begin{aligned}FP&=\int\prod_{i=1}^{n}[dt_i e^{\ell(0,t_i)}]\Theta\left(\sum_{i=1}^{n}\left[\ell(x,t_i)-\ell(0,t)\right]-\theta\right),\\ FN&=\int\prod_{i=1}^{n}[dt_i e^{\ell(x,t_i)}]\Theta\left(\sum_{i=1}^{n}\left[\ell(0,t_i)-\ell(x,t)\right]+\theta\right).\end{aligned} \tag{B2}$$

The integral representation of the Heaviside function, Eq. (A6) can be used again to obtain:

$$\begin{aligned}FP&=\int_{-i\infty}^{+i\infty}\frac{d\lambda}{2\pi i\lambda}e^{-\lambda\theta}\left[\int dt e^{(1-\lambda)\ell(0,t)+\lambda\ell(x,t)}\right]^n,\\ FN&=\int_{-i\infty}^{+i\infty}\frac{d\lambda}{2\pi i\lambda}e^{\lambda\theta}\left[\int dt e^{(1-\lambda)\ell(x,t)+\lambda\ell(0,t)}\right]^n.\end{aligned} \tag{B3}$$

Substituting $\lambda\to(1-\lambda)$ in the second equation gives an expression for FN that looks very similar to FP:

$$\int\frac{d\lambda}{2\pi i(1-\lambda)}e^{(1-\lambda)\theta}\left[\int dt e^{(1-\lambda)\ell(0,t)+\lambda.\ell(x,t)}\right]^n. \tag{B4}$$

In summary:

$$\begin{aligned}FP&=\int_{-i\infty}^{+i\infty}\frac{d\lambda}{2\pi i\lambda}e^{n\phi(\lambda)-\lambda\theta}\\ FN&=\int_{-i\infty}^{+i\infty}\frac{d\lambda}{2\pi i(1-\lambda)}e^{n\phi(\lambda)+(1-\lambda)\theta},\end{aligned} \tag{B5}$$

with:

$$\begin{aligned}\phi(\lambda)&=\ln\int dt e^{(1-\lambda)\ell(0,t)+\lambda\ell(x,t)}\\ &=\ln\int du\, e^{-u}\left[1+x(\alpha e^{(1-\alpha)u}-1)\right]^{\lambda}.\end{aligned} \tag{B6}$$

These two expressions can be evaluated in the large n limit using a saddle-point approximation, with the same saddle-point condition $\theta=n\phi'(\lambda)$ for both, yielding:

$$\begin{aligned}FP&\approx\frac{1}{\lambda\sqrt{2\pi n|\phi''(\lambda)|}}\exp\left[n\phi(\lambda)-\lambda\theta\right],\\ FN&\approx\frac{1}{(1-\lambda)\sqrt{2\pi n|\phi''(\lambda)|}}\exp\left[n\phi(\lambda)+(1-\lambda)\theta\right].\end{aligned} \tag{B7}$$

2. Small x limit

Again two regimes emerge in the $x\ll 1$ limit, depending on whether α is smaller or greater than $1/2$. When $\alpha>1/2$, $\phi(\lambda)$ can be expand at small x:

$$\phi(\lambda)\approx-\frac{1}{2}\frac{(1-\alpha)^2}{2\alpha-1}\lambda(1-\lambda)x^2. \tag{B8}$$

This implies:

$$\begin{aligned}FP&\approx\frac{e^{-\frac{1}{2}\lambda^2 nx^2/g(\alpha)}}{\sqrt{2\pi\lambda^2 nx^2/g(\alpha)}},\\ FN&\approx\frac{e^{-\frac{1}{2}(1-\lambda)^2 nx^2/g(\alpha)}}{\sqrt{2\pi(1-\lambda)^2 nx^2/g(\alpha)}},\end{aligned} \tag{B9}$$

with

$$g(\alpha)=\frac{2\alpha-1}{(1-\alpha)^2}. \tag{B10}$$

When $\alpha < 1/2$, one can do the same change of variable as before, $y = x\alpha e^{(1-\alpha)u}$, to obtain at leading order:

$$\phi(\lambda) = \frac{(\alpha x)^{\frac{1}{1-\alpha}}}{1-\alpha} \int_0^{+\infty} dy\, y^{-\frac{2-\alpha}{1-\alpha}} \left[(1+y)^\lambda - 1 - \lambda y\right]. \tag{B11}$$

The integrand is of order $y^{-\alpha/(1-\alpha)}$ at small y, and therefore is integrable. The error rates are then given by:

$$\begin{aligned} FP &\approx \frac{\exp\left[nx^{\frac{1}{1-\alpha}}\left(\chi(\lambda) - \lambda\chi'(\lambda)\right)\right]}{\lambda\sqrt{2\pi n x^{\frac{1}{1-\alpha}}|\chi''(\lambda)|}} \\ FN &\approx \frac{\exp\left[nx^{\frac{1}{1-\alpha}}\left(\chi(\lambda) + (1-\lambda)\chi'(\lambda)\right)\right]}{(1-\lambda)\sqrt{2\pi n x^{\frac{1}{1-\alpha}}|\chi''(\lambda)|}}. \end{aligned} \tag{B12}$$

where

$$\chi(\lambda) = \frac{\alpha^{\frac{1}{1-\alpha}}}{1-\alpha} \int_0^{+\infty} dy\, y^{-\frac{2-\alpha}{1-\alpha}} \left[(1+y)^\lambda - 1 - \lambda y\right]. \tag{B13}$$

The intermediate case $\alpha = 1/2$ is treated similarly as before, by noting that the integral defining $\phi(\lambda)$ is dominated by the (diverging) term of order y^{-1}. This gives:

$$\phi(\lambda) \approx -\frac{x^2}{4} |\ln(x/2)| \lambda(1-\lambda). \tag{B14}$$

and therefore:

$$\begin{aligned} FP &\approx \frac{e^{-\frac{1}{4}\lambda^2 n x^2 |\ln(x/2)|}}{\sqrt{\pi\lambda^2 n x^2 |\ln(x/2)|}}, \\ FN &\approx \frac{e^{-\frac{1}{4}(1-\lambda)^2 n x^2 |\ln(x/2)|}}{\sqrt{\pi(1-\lambda)^2 n x^2 |\ln(x/2)|}}. \end{aligned} \tag{B15}$$

As a result, the number of binding events n necessary to a make reliable decision scales as x^{-2} for $\alpha > 1/2$, $x^{-2}|\ln(x)|^{-1}$ for $\alpha = 1/2$ and $x^{-\gamma}$ for $\alpha < 1/2$, with $\gamma = (1-\alpha)^{-1}$. Replacing $n \approx 4Dac_{\rm tot}(1-p)NT$ gives the scaling for the minimal detection time:

$$T \sim \frac{1}{4Da(1-p)N} \times \begin{cases} c_{\rm tot} c^{-2} & \alpha > 1/2, \\ c_{\rm tot} c^{-2} |\ln(c/c_{\rm tot})|^{-1} & \alpha = 1/2, \\ c_{\rm tot}^{\gamma-1} c^{-\gamma} & \alpha < 1/2. \end{cases} \tag{B16}$$

Appendix C: Optimization of the signaling rates in the receptor cascade

Each receptor goes through a cascade of states $\mu = 1, \ldots, m$ upon binding. At any moment, the receptor can become unbound with rate r or r'. In the following some expressions will be given in terms of the unbinding rate of the correct ligand r, but the same expressions hold for the incorrect ligand after substitution by r'.

The probability of reaching state μ is $[s/(s+r)]^{\mu-1}$. Assuming it has reached state μ, the time t_μ spent in that state is distributed according to $(s+r)e^{-(s+r)t_\mu}$. In summary t_μ is distributed as follows:

$$P_r(t_\mu) = \frac{s^{\mu-1}}{(s+r)^{\mu-2}} e^{-(s+r)t_\mu} + \left[1 - \left(\frac{s}{s+r}\right)^{\mu-1}\right] \delta(t_\mu), \tag{C1}$$

where $\delta(x)$ is Dirac's delta function. Its first and second moments are:

$$\langle t_\mu \rangle_r = \frac{s^{\mu-1}}{(s+r)^\mu}, \tag{C2}$$

$$\langle \delta t_\mu^2 \rangle_r = \frac{s^{\mu-1}}{(s+r)^{\mu+1}} \left[2 - \frac{s^{\mu-1}}{(s+r)^{\mu-1}}\right]. \tag{C3}$$

The output of the network is given by:

$$\frac{dX^*}{dt} = X_0 \sum_{j=1}^{N} (b_{\mu(j)} - d_{\mu(j)}), \tag{C4}$$

where $\mu(j)$ is the state of the $j^{\rm th}$ receptor and $b_0 - d_0 = 0$, so that the net effect of one binding event is

$$\Delta X^* = X_0 \sum_{\mu=1}^{m} (b_\mu - d_\mu) t_\mu. \tag{C5}$$

On average, binding a wrong ligand will cause a change

$$\langle \Delta X^* \rangle_{r'} = X_0 \sum_{\mu=1}^{m} (b_\mu - d_\mu) \frac{s^{\mu-1}}{(s+r')^\mu}. \tag{C6}$$

When optimizing over the net rates $(b_\mu - d_\mu)$, this quantity is set to zero, to ensure that only the correct ligand changes X^* on average. This way, X^* is proportional to c in the limit of long times:

$$\langle X^*(T) \rangle \approx 4Da(1-p)NTc \langle \Delta X^* \rangle_r, \tag{C7}$$

Although the mean of $X^*(T)$ is not affected by incorrect binding events, its variance is, and reads:

$$\begin{aligned} &\langle X^*(T)^2 \rangle - \langle X^*(T) \rangle^2 \approx 4Da(1-p)NTc_{\rm tot} \\ &\quad \times \left[x\langle (\Delta X^*)^2 \rangle_r + (1-x)\langle (\Delta X^*)^2 \rangle_{r'} - x^2 \langle \Delta X^* \rangle_r^2\right], \end{aligned} \tag{C8}$$

where

$$\langle (\Delta X^*)^2 \rangle_r = \langle \Delta X^* \rangle_r^2 + X_0^2 \sum_{\mu=1}^{m} (b_\mu - d_\mu)^2 \langle \delta t_\mu^2 \rangle_r. \tag{C9}$$

and the same for r'.

For a given m, the signal-to-noise ratio

$$SNR = \frac{\langle X^*(T) \rangle^2}{\langle X^*(T)^2 \rangle - \langle X^*(T) \rangle^2} \tag{C10}$$

is maximized over the rates $b_\mu - d_\mu$. The procedure gives the optimized networks discussed in the main text.

[1] Berg HC, Purcell EM (1977) Physics of chemoreception. *Biophys. J.* 20:193–219.

[2] Elowitz MB, Levine AJ, Siggia ED, Swain PS (2002) Stochastic gene expression in a single cell. *Science* 297:1183–1186.

[3] Gregor T, Tank DW, Wieschaus EF, Bialek W (2007) Probing the Limits to Positional Information. *Cell* 130:153–164.

[4] Tkačik G, Walczak AM (2011) Information transmission in genetic regulatory networks: a review. *J. Phys. Condens. Matter* 23:153102.

[5] Bowsher CG, Swain PS (2014) Environmental sensing, information transfer, and cellular decision-making. *Curr. Opin. Biotechnol.* 28:149–155.

[6] Tkačik G, Bialek W (2014) Information processing in living systems. *arXiv* p 1412.8752.

[7] Bialek W, Setayeshgar S (2005) Physical limits to biochemical signaling. *Proc. Natl. Acad. Sci. U. S. A.* 102:10040–10045.

[8] Endres RG, Wingreen NS (2009) Maximum likelihood and the single receptor. *Phys. Rev. Lett.* 103:158101.

[9] Kaizu K, et al. (2014) The Berg-Purcell limit revisited. *Biophys. J.* 106:976–85.

[10] Endres RG, Wingreen NS (2008) Accuracy of direct gradient sensing by single cells. *Proc. Natl. Acad. Sci.* 105:15749–15754.

[11] Rappel WJ, Levine H (2008) Receptor noise limitations on chemotactic sensing. *Proc. Natl. Acad. Sci. U. S. A.* 105:19270–19275.

[12] Rappel WJ, Levine H (2008) Receptor noise and directional sensing in eukaryotic chemotaxis. *Phys. Rev. Lett.* 100:228101.

[13] Endres RG, Wingreen NS (2009) Accuracy of direct gradient sensing by cell-surface receptors. *Prog. Biophys. Mol. Biol.* 100:33–39.

[14] Hu B, Chen W, Rappel WJ, Levine H (2010) Physical limits on cellular sensing of spatial gradients. *Phys. Rev. Lett.* 105:1–4.

[15] Mora T, Wingreen NS (2010) Limits of sensing temporal concentration changes by single cells. *Phys. Rev. Lett.* 104:1–11.

[16] Mehta P, Schwab DJ (2012) Energetic costs of cellular computation. *Proc. Natl. Acad. Sci.* 109:17978–17982.

[17] Lan G, Sartori P, Neumann S, Sourjik V, Tu Y (2012) The energyspeedaccuracy trade-off in sensoryadaptation. *Nat. Phys.* 8:422–428.

[18] Becker NB, Mugler A, ten Wolde PR (2013) Prediction and Dissipation in Biochemical Sensing. *arXiv* p 1312.5625.

[19] Lang AH, Fisher CK, Mora T, Mehta P (2014) Thermodynamics of statistical inference by cells. *Phys. Rev. Lett.* 113:148103.

[20] Govern CC, ten Wolde PR (2014) Optimal resource allocation in cellular sensing systems. *Proc. Natl. Acad. Sci.* 111:17486–17491.

[21] Govern CC, ten Wolde PR (2014) Energy Dissipation and Noise Correlations in Biochemical Sensing. *Phys. Rev. Lett.* 113:1–5.

[22] Barato AC, Hartich D, Seifert U (2014) Efficiency of cellular information processing. *New J. Phys.* 16:103024.

[23] Mancini F, Marsili M, Walczak A (2015) Trade-offs in delayed information transmission in biochemical networks. *arXiv* p 1504.03637.

[24] Barato AC, Seifert U (2015) Thermodynamic Uncertainty Relation for Biomolecular Processes. *Phys. Rev. Lett.* 114:158101.

[25] Hartich D, Barato AC, Seifert U (2015) Nonequilibrium sensing and its analogy to kinetic proofreading. *arXiv* p 1502.02594.

[26] Lalanne JB, François P (2015) Chemodetection in fluctuating environments: Receptor coupling, buffering, and antagonism. *Proc. Natl. Acad. Sci. U. S. A.* 112:1898–903.

[27] Feinerman O, Germain RN, Altan-Bonnet G (2008) Quantitative challenges in understanding ligand discrimination by $\alpha\beta$ T cells. *Mol. Immunol.* 45:619–631.

[28] Hopfield JJ (1974) Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc. Natl. Acad. Sci. U. S. A.* 71:4135–4139.

[29] Ninio J (1975) Kinetic amplification of enzyme discrimination. *Biochimie* 57:587–595.

[30] McKeithan TW (1995) Kinetic proofreading in T-cell receptor signal transduction. *Proc. Natl. Acad. Sci. U. S. A.* 92:5042–5046.

[31] François P, Voisinne G, Siggia ED, Altan-Bonnet G, Vergassola M (2013) Phenotypic model for early T-cell activation displaying sensitivity, specificity, and antagonism. *Proc. Natl. Acad. Sci. U. S. A.* 110:E888–97.

[32] Lalanne JB, François P (2013) Principles of adaptive sorting revealed by in silico evolution. *Phys. Rev. Lett.* 110:218102.

[33] Siggia ED, Vergassola M (2013) Decisions on the fly in cellular sensory systems. *Proc. Natl. Acad. Sci. U. S. A.* 110:E3704–12.

[34] Kay SM (2001) *Fundamentals Of Statistical Signal Processing: Estimation Theory* (Prentice Hall PTR, Upper Saddle River, NJ).

[35] Murugan A, Huse DA, Leibler S (2012) Speed, dissipation, and error in kinetic proofreading. *Proc. Natl. Acad. Sci.* 109:12034–12039.

[36] Cepeda-Humerez SA, Rieckh G, Tkačik G (2015) Stochastic proofreading mechanism alleviates crosstalk in transcriptional regulation. *ArXiv e-prints*.

[37] Singh V, Nemenman I (2015) Accurate sensing of multiple ligands with a single receptor. *ArXiv e-prints*.