Physical Limit to Concentration Sensing in a Changing Environment

Thierry Mora and Ilya Nemenman

Phys. Rev. Lett. **123**, 198101 — Published 5 November 2019

DOI: 10.1103/PhysRevLett.123.198101

Physical limit to concentration sensing in a changing environment

Thierry Mora[1,*] and Ilya Nemenman[2]

[1]*Laboratoire de physique de École normale supérieure (PSL University), CNRS, Sorbonne University, Université de Paris, 24 rue Lhomond, 75005 Paris, France*
[2]*Department of Physics, Department of Biology, and Initiative in Theory and Modeling of Living Systems, Emory University, Atlanta, GA 30322, USA*

Cells adapt to changing environments by sensing ligand concentrations using specific receptors. The accuracy of sensing is ultimately limited by the finite number of ligand molecules bound by receptors. Previously derived physical limits to sensing accuracy largely have assumed that the concentration was constant and ignored its temporal fluctuations. We formulate the problem of concentration sensing in a strongly fluctuating environment as a non-linear field-theoretic problem, for which we find an excellent approximate Gaussian solution. We derive a new physical bound on the relative error in concentration c which scales as $\delta c/c \sim (Dac\tau)^{-1/4}$ with ligand diffusivity D, receptor cross-section a, and characteristic fluctuation time scale τ, in stark contrast with the usual Berg and Purcell bound $\delta c/c \sim (DacT)^{-1/2}$ for a perfect receptor sensing concentration during time T. We show how the bound can be achieved by a biochemical network downstream the receptor that adapts the kinetics of signaling as a function of the square root of the sensed concentration.

Cells must respond to extracellular signals to guide their actions. The signals typically come in the form of changing concentrations of various molecular ligands, which are conveyed to the cell through ligand binding to cell surface receptors. A lot of ink has been expended on deriving the fundamental limits to the precision with which a cell can measure the concentrations from the activity of its receptors, constrained by the stochasticity of ligand binding and unbinding [1–4]. In particular, it has become clear that the temporal sequence of binding-unbinding events carries more information about the underlying ligand concentration than just the mean receptor occupancy, typically used in deterministic chemical kinetics models [5]. In particular, such precise temporal information allows cells to estimate the concentration of a cognate ligand even in a sea of weak spurious ligands [6–8], as well as to estimate concentrations of multiple ligands from fewer receptor types [9, 10], and molecular network motifs able to perform such complex estimation exist in the real world, even potentially taking advantage of cross-talk between receptor-ligand pairs [11].

Importantly, concentrations of ligands are worth measuring only when they are *a priori* unknown; or, in other words, if they change with time, allowing for instance cells to adapt their behaviour accordingly and maximize their long-term growth [12]. Chemotacting microorganisms may experience sudden and unpredictable changes in the concentration of attractants and repellents within seconds as they navigate through complex environments [13] shaped by microbial communities [14]. Likewise, during fly development, cells choose their fate within minutes by sensing time-varying maternal gradients [15]. However, with the exception of a recent study on the optimal resource allocation for sensing time-varying concentrations [16], all previous analyses have focused on the regime with a clear time scale separation, where the concentration is constant or constantly changing [17] during the period over which it is estimated. In this article, we calculate the accuracy with which a temporally varying ligand concentration may be estimated from a sequence of binding and unbinding events. This requires assumptions about the time scale over which significant changes of the concentration are possible. In our formulation, the optimal sensor performs a Bayesian computation, formalized mathematically as a stochastic field theory. Crucially, we show how simple biochemical circuits can perform the relevant complex computations.

Field theory of concentration sensing. We associate to the ligand concentration $c(t)$ a field $\varphi(t)$ through $c(t) = c_0 e^{-\varphi(t)}$, where c_0 is an irrelevant reference concentration. Ligand concentration controls the ligand-receptor binding rate $r(t) = 4Dac(t) = 4Dac_0 e^{-\varphi(t)} \equiv r_0 e^{-\varphi(t)}$, where $4Da$ is the diffusion-limited binding rate per molecule of the ligand to its target receptor, modeled as a circle of diameter a on the cell's surface, and D is the ligand diffusivity. This binding rate can be readily generalized to N receptors by using instead $r(t) = 4NDac(t)$. All our results will then hold with this additional N factor. We assume that the concentration follows a geometric random walk, with characteristic time scale τ: $d\varphi = \tau^{-1/2} dW$, with W a Wiener process. This choice is justified by the fact that in many biological contexts, such as bacterial chemotaxis, concentrations may vary over many orders of magnitude [18, 19].

The probability of the temporal evolution of the concentration over the time interval $[0, T]$ is given by

$$P_{\rm prior}(\{\varphi(t)\}) = \frac{1}{Z_{\rm prior}} \exp\left[-\frac{\tau}{2}\int_0^T dt \left(\frac{d\varphi}{dt}\right)^2\right]. \quad (1)$$

The receptor sees binding events at times $t_1, t_2, \ldots, t_n$, each occuring with rate $4Dac(t_i) = r_0 e^{-\varphi(t_i)}$. To simplify, let us assume that unbinding is instantaneous (generalization to finite binding times is discussed later). The posterior distribution of the concentration profile then

follows Bayes' rule (App. A [20]):

$$P(\{\varphi(t)\}) = \frac{P(t_1,\dots,t_n|\{\varphi(t)\})P_{\rm prior}(\{\varphi(t)\})}{P(t_1,\dots,t_n)}$$
$$= \frac{1}{Z}\exp\left\{-\int_0^T dt\,\left[\frac{\tau}{2}\left(\frac{d\varphi}{dt}\right)^2 + r_0 e^{-\varphi(t)}\right] - \sum_{i=1}^n \varphi(t_i)\right\}, \quad (2)$$

where Z is a normalization constant independent of φ. The term $r_0 e^{-\varphi} dt$ in the integral reflects the probability $\exp(-r_0 e^{-\varphi} dt)$ of not binding a ligand between t and $t+dt$ (except at times t_i). The binding events at $t = t_i$ are generated by the *true* temporal trace of ligand concentration, $c^*(t) = c_0 e^{-\varphi^*(t)}$. In the following the true trace $\varphi^*(t)$ will be distinguished from the field φ, which refers to our observation-based belief.

The one-dimensional field-theoretic problem (2) is a particular case of Bayesian filtering [21]. When collecting information from binding events, cells do not have access to the future and cannot use the full span $[0,T]$ of observations to infer the concentration at time t. Instead, they must infer it solely based on past observation in the interval $[0,t]$, which distinguishes our problem from the mathematically similar inference of a continuous probability density [22–26]. This inference can be performed recursively by the rules of Bayesian sequential forecasting, similar to the transfer matrix technique, and also known as the forward algorithm [21]. To do this recursion, we first define:

$$Z(\varphi,t) = \int \mathcal{D}\varphi(t)\,\delta(\varphi(t)-\varphi)\exp\left[-\frac{\tau}{2}\int_0^t dt'\left(\frac{d\varphi}{dt}\right)^2 - \int_0^t dt'\left(r_0 e^{-\varphi(t')} + \varphi(t')\sum_{i=1}^n \delta(t'-t_i)\right)\right]. \quad (3)$$

Considering past observations during the interval $[0,t]$, the posterior distribution of φ at time t reads:

$$P(\varphi,t) = \frac{Z(\varphi,t)}{Z(t)}, \quad \text{with} \quad Z(t) = \int_{-\infty}^{\infty} d\varphi' Z(\varphi',t). \quad (4)$$

When considering periods during which no binding event was observed, we can write a recursion for $Z(\varphi,t)$ between t and $t+dt$. Taking the $\delta t \to 0$ limit yields, for $t \neq t_i$ (App. A [20]):

$$\frac{\partial P(\varphi,t)}{\partial t} = -r_0(e^{-\varphi} - \langle e^{-\varphi}\rangle)P(\varphi,t) + \frac{1}{2\tau}\frac{\partial^2 P}{\partial \varphi^2}, \quad (5)$$

where $\langle\cdot\rangle$ denotes an average over $P(\varphi)$. When a binding even does occur at time t_i, the posterior distribution is updated using Bayes' rule:

$$P(\varphi, t_i^+) = \frac{e^{-\varphi}P(\varphi,t_i^-)}{\langle e^{-\varphi}\rangle}, \quad (6)$$

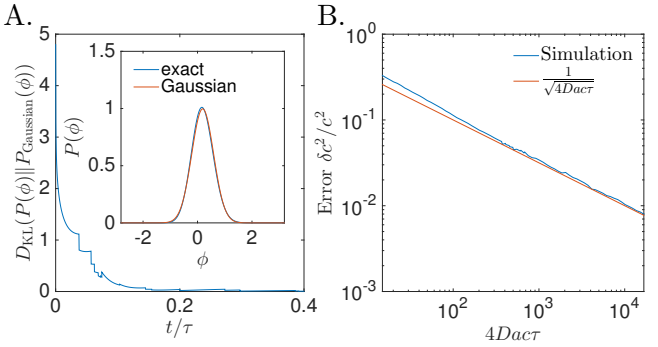

FIG. 1: **Numerical validations of analytical results. A.** The Gaussian Ansatz (7)-(8) is validated by simulating the general equations for Bayesian filtering (5)-(6). The numerical solution approaches the Gaussian solution rapidly, as indicated by the decay of the Kullback-Leibler divergence $D_{\rm KL}(P(\varphi)\|P_{\rm Gaussian}(\varphi)) = \int d\varphi\, P(\varphi)\ln(P(\varphi)/P_{\rm Gaussian}(\varphi))$. We used $r\tau = Dac\tau = 50$. **B.** Concentration sensing error as a function of concentration. The error estimated from simulations follows closely the prediction from (13), which is expected to be valid for $4Dac\tau \gg 1$.

where $t_i^\pm$ refer to the values right before and after the observation. The partition function $Z(t)$ can be similarly calculated (App. A [20]) and could in principle be used to infer the correct timescale τ by maximizing $P(\tau|\{t_1,\dots,t_N\}) \propto Z$ (App. C [20]).

Gaussian solution. Because of the $P(\varphi)$ dependence in $\langle e^{-\varphi}\rangle$, the equations for the evolution of the posterior probability (5)-(6) are nonlinear. However, assuming a Gaussian Ansatz $P(\varphi,t) = (2\pi\sigma(t)^2)^{-1/2}\exp[-(\varphi-\hat\varphi(t))^2/2\sigma(t)^2]$, which is accurate in the limit of long measurement times (see below), gives a closed-form solution (App. B [20]), with:

$$\frac{d\hat\varphi}{dt} = \sigma^2\left[r_0 e^{-\hat\varphi+\sigma^2/2} - \sum_{i=1}^n \delta(t-t_i)\right], \quad (7)$$
$$\frac{d\sigma^2}{dt} = \frac{1}{\tau} - \sigma^4 r_0 e^{-\hat\varphi+\sigma^2/2}. \quad (8)$$

The maximum a posterior estimator for the concentration is then simply given by $\hat c(t) = c_0 e^{-\hat\varphi(t)}$, while $\sigma(t)^2$ defines the Bayesian uncertainty on the estimator.

To check the validity of the Gaussian solution, we simulated (5)-(6) numerically, starting from a uniform distribution ($P(\varphi,0) = 1/2$ for $\varphi \in [-1,1]$ and 0 otherwise), with $r_0\tau = 50$ and a true $\varphi^*(t)$ starting at $\varphi^*(0) = 0$. The numerical solution quickly approaches the Gaussian solution given by (7)-(8) starting with $\hat\varphi(0) = \langle\varphi\rangle_{t=0}$ and $\sigma(0)^2 = {\rm Var}(\varphi)_{t=0}$. The Kullback-Leibler divergence between the numerical and analytical solutions falls rapidly (Fig. 1A) and the numerical solution approaches the predicted Gaussian very closely (Fig. 1A, inset). Thus, the Gaussian solution provides an excellent approximation.

Error estimate. To study the typical behaviour of (7)-

(8), we now assume that the rate of binding events is large compared to the rate of change of the concentration, $4Dac\tau = r\tau \gg 1$. This regime is the biologically relevant one: to sense concentration, cells need to record many binding events over the time scale on which the concentration fluctuates. In that limit the estimator $\hat{\varphi}$ is close to the true value φ^*, and the Bayesian uncertainty σ^2 is small, allowing for two simplifications. First, (8) relaxes over time scale $r(t)^{-1}$ to a quasi-steady state value $\sigma^2 \approx 1/\sqrt{r_0 e^{-\hat{\varphi}}\tau} \ll 1$. Second, we can make a small noise approximation for binding events: over some time interval Δt, with $r^*(t)^{-1} \ll \Delta t \ll \tau$, the number of binding events has both mean and variance equal to $r^*(t)\Delta t$, allowing us to replace discrete jumps in (7) by:

$$d\left(\sum_{i=1}^{n}\delta(t-t_i)\right) \approx r_0 e^{-\varphi^*}dt + (r_0 e^{-\varphi^*})^{1/2}dW', \quad (9)$$

where W' is a Wiener process. As a result, the estimator $\hat{\varphi}$ tracks the true value φ^* according to:

$$d\hat{\varphi} \approx (r_0 e^{-\hat{\varphi}}/\tau)^{1/2}(\varphi^* - \hat{\varphi}) + \tau^{-1/2}dW', \quad (10)$$

where we have expanded at first order in $\hat{\varphi} - \varphi^*$. In the general case, the true field may evolve according to a different characteristic time scale, τ^*, than the one assumed by the Bayesian filter, τ, so that $d\varphi^* = (\tau^*)^{-1/2}dW$. The estimation error $\epsilon = \hat{\varphi} - \varphi^*$ then evolves according to:

$$d\epsilon = -(r/\tau)^{1/2}\epsilon\, dt + \tau^{-1/2}dW' - (\tau^*)^{-1/2}dW. \quad (11)$$

Intrigingy, the noises dW' and dW have very different interpretations, one being due to the random arrival of binding events, and the other to the geometric diffusion of the concentration. Yet they come in the same form in this equation. Relying again on the assumption that $r\tau \gg 1$, we get an estimate of the error:

$$\langle\epsilon^2\rangle = \frac{1}{2\sqrt{r}}\left(\frac{1}{\sqrt{\tau}} + \frac{\sqrt{\tau}}{\tau^*}\right), \quad (12)$$

which has a minimum as a function of τ, reached for the true value of the characteristic fluctuation time $\tau = \tau^*$:

$$\frac{\langle(\hat{c}-c^*)^2\rangle}{c^2} \approx \langle\epsilon^2\rangle = \frac{1}{\sqrt{r\tau}} = \frac{1}{\sqrt{4Dac\tau}}. \quad (13)$$

This error is equal to the Bayesian uncertainty $\sigma^2 = 1/\sqrt{r_0\tau e^{-\hat{\varphi}}} \approx 1/\sqrt{4Dac\tau}$ and is consistent with the error found using the saddle-point approximation in the related problem of probability density estimate [22].

We checked the validity of our small-noise approximation by comparing the prediction from (12) with the results of a numerical simulation of (7)-(8), in which we averaged the error $\langle(\hat{c}-c^*)^2\rangle$ as a function of c for many realization of the process. The agreement is excellent, and gets better as $r\tau = 4Dac\tau$ becomes larger (Fig. 1B).

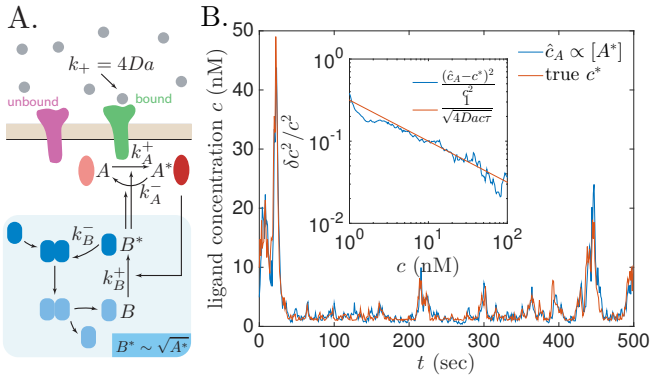

FIG. 2: **Performance of adaptive biochemical network in fluctuating ligand concentration.** A. Schematic of the biochemical network implementing optimal Bayesian filtering. The receptor-induced activation of the readout molecule A^*, as well as its deactivation are regulated by a second molecule B^*, which is made to scale like $\sqrt{A^*}$ using a mechanism of deactivation by dimerisation (shaded box). **B.** Simulation of the network readout $c_A(t) \propto A^*(t)$ in response to stochastic binding events in a fluctuating concentration field $c^*(t)$. The relative estimation error $\langle(\hat{c}_A - c^*)^2\rangle/c^2$ behaves according to the theoretical bound $1/\sqrt{4Dac\tau}$ (inset).

The error in (13) sets a fundamental physical limit on any concentration sensing device, biological or artificial, in a concentration profile that follows a geometric random walk. This bound is radically different from that obtained Berg and Purcell for the concentration sensing by a single receptor integrating over time T [1, 5]:

$$\frac{\delta c^2}{c^2} = \frac{1}{4DacT} \quad (14)$$

(in the limit where binding events are short so that the receptor is always free).

The major difference is that Berg and Purcell, as well as most of the literature on concentration sensing, assume that the sensed concentration does not change with time. Our result can be reconciled with Berg and Purcell by defining an effective measurement time $T_{\text{eff}} \sim \sqrt{\tau/4Dac}$ — the geometric mean between the mean time between binding events and the time scale of the concentration variation — which can be read off from the relaxation rate in (10), $(r_0 e^{-\varphi}/\tau)^{1/2} = T_{\text{eff}}^{-1}$. This T_{eff} realizes the optimal tradeoff between the requirement to integrate many binding events, $T_{\text{eff}} \gg 1/(4Dac)$, but over a relatively constant concentration, $T_{\text{eff}} \ll \tau$, as can also be anticipated from a semi-quantitative argument using Fourier analysis [27, 28] (see App. F [20] for a derivation). A similar trade-off was reported in a more detailed chemical kinetics model of concentration sensing [16].

Plausible biological implementation. Can cells implement the optimal Bayesian filtering scheme and reach the bound set by (13)? To gain intuition, it is useful to rewrite (7)-(8) in term of the concentration estimator $\hat{c}$,

in the limit $4Dac\tau \gg 1$ where σ^2 can be eliminated:

$$\frac{d\hat{c}}{dt} = \sqrt{4Da\hat{c}/\tau}\left(\frac{1}{4Da}\sum_{i=1}^{n}\delta(t-t_i) - \hat{c}\right). \quad (15)$$

Each binding event should lead to an increment of $\hat{c}$, followed by a continuous decay with a rate given by $T^{-1} = \sqrt{4Da\hat{c}/\tau}$.

This scheme can be implemented by a simple biochemical network schematized in Fig. 2A. The concentration readout $\hat{c}_A$ may be represented by the "active" (for instance phosphorylated) form A^* of a chemical species. Binding events cause the receptor to activate A into A^*, which gets subsequently deactivated. Both the activation and deactivation of A are catylized by a second chemical species in its active form, B^*. Thus, upon a binding event, the concentration of active A^* is increased by:

$$\Delta[A^*] = k_A^+[A][B^*], \quad (16)$$

and it decays between binding events according to:

$$\frac{d[A^*]}{dt} = -k_A^-[B^*][A^*], \quad (17)$$

where $k_A^\pm$ are biochemical parameters.

To implement (15), the concentration of B^* must be controled by the square root of A^*. This dependence can be achieved by assuming that B is activated into B^* through the catalityc activity of A^*, and that B^* gets deactivated cooperatively as a dimer:

$$\frac{d[B^*]}{dt} = k_B^+[B][A^*] - k_B^-[B^*]^2, \quad (18)$$

where $k_B^\pm$ are biochemical reaction rates.

Assuming that the kinetics of B are fast compared to A, we obtain $B^* = (Bk_B^+/k_B^-)^{1/2}\sqrt{A^*}$ and

$$\frac{d[A^*]}{dt} = \alpha\sqrt{[A^*]}\left(\beta\sum_{i=1}\delta(t-t_i) - [A^*]\right). \quad (19)$$

with $\alpha = k_A^-([B]k_B^+/k_B^-)^{1/2}$ and $\beta = (k_A^+[A]/k_A^-)$. If A and B are in excess, and thus approximately constant, then this biochemical network exactly implements (15), with $4Da\hat{c}_A \equiv k_A^-[A^*]/k_A^+[A]$, and $\tau = \tau_{\rm net} \equiv 1/(\alpha^2\beta) = k_B^-/(k_B^+k_A^+k_A^-[A][B])$.

Interestingly, the amount of inactive ($\approx$ total) B controls the time scale of concentration fluctuations, and could be tuned through gene regulation to adapt to different speeds of environmental fluctuations. A biochemical network might be able to find the optimal τ and then adjust $[B]$ accordingly by empirically measuring the fold-change of $r(t)$ (which can be done by biochemical networks, see e.g. [29]) but with a delay, $\langle r(t+\Delta t)/r(t)\rangle = e^{\Delta t/2\tau}$, and then inverting the relationship to extract τ.

We tested the performance of the biochemical network for sensing concentration by simulating (16)-(18) with a fluctuating ligand concentration $c(t)$ with characteristic time scale τ. For concreteness, we set $c^*(0) = 10$nM, $\tau^* = 10$s, $k_A^+[A] = 0.01$, $k_A^- = k_B^+ = k_B^+ = 1\mu\text{M}^{-1}\text{s}^{-1}$ and $[B] = 10\mu$M, so that $\tau_{\rm net} = \tau^*$. Fig. 2B shows the network estimate $\hat{c}_A(t)$ along with the true value $c^*(t)$. The empirical error $\langle(\hat{c}_A - c^*)^2\rangle$ as a function of c^* averaged over 10^4s (Fig. 2B, inset), again shows an excellent agreement with the theoretical bound $1/\sqrt{4Dac\tau}$.

Discussion. For the sake of clarity our analysis made simplifying assumptions which can be easily relaxed. Our proposed biochemical implementation assumed a constant burst of activity following each binding event, consistent with the optimal estimation strategy. However, in real receptors, stochasticity in the bound time is known to double the variance in the estimate [5] (App. D [20]). Treating this effect simply adds a factor $\sqrt{2}$ in the noise term of (9) as well as in (13), $\langle\delta c^2\rangle/c^2 \approx 1/\sqrt{2Dac\tau}$. We also ignored periods during which the receptor was bound. During that time the receptor is blind to the external world, and the posterior evolves according to the prior: $\partial_t P = (1/2\tau)\partial_\varphi^2 P$, $\partial_t\hat{\varphi} = 0$ and $\partial_t\sigma^2 = 1/\tau$. In our results, these "down times" renormalize the effective observation time by the fraction of time the receptor is free, $p_{\rm free} = (1+4Dacu)^{-1}$, where u is the average bound time, $\langle\delta c^2\rangle/c^2 \approx 1/\sqrt{4Dacp_{\rm free}\tau}$ (App. D [20]). Combining the two effects (stochasticity in bound time and receptor availability) would yield $\langle\delta c^2\rangle/c^2 \approx 1/\sqrt{2Dacp_{\rm free}\tau}$. Our network analysis also ignores noise in the readout molecules, as we focused exclusively on the sensing noise itself. For a thorough discussion of tradeoffs between difference noise sources, see Ref. [30], and Ref. [16] in the context of time-varying signals.

The field theory of (2) is mathematically similar to the problem of estimating a density function from a small sample set with a smoothing prior [22–24, 26]. The main difference lies in the domain of observations. In density estimation the whole function $\{\varphi(t)\}_{t\in[0,T]}$ is infered together on the whole domain of t, while sensors can only learn from past observations, *i.e.* the $t' < t$ half-plane. However, our solution can easily be generalized to deal with the entire time domain using the forward-backward algorithm (App. E [20]). Eqs. (5)-(6) and (7)-(8) can be solved both forward(from 0 to t) and backward (from T to t, with time reversal) in time, giving $P_\rightarrow(\varphi)$, $\hat{\varphi}_\rightarrow$, $\sigma^2_\rightarrow$ for the forward solution (the one treated in this article), and $P_\leftarrow(\varphi)$, $\hat{\varphi}_\leftarrow$, $\sigma^2_\leftarrow$ for the backward solution. The Bayesian posterior at any given time is then given by $\propto P_\rightarrow(\varphi)P_\leftarrow(\varphi)$, of mean $(\sigma^2_\leftarrow\hat{\varphi}_\rightarrow + \sigma^2_\rightarrow\hat{\varphi}_\leftarrow)/(\sigma^2_\leftarrow + \sigma^2_\rightarrow)$ and variance $\sigma^2_\leftarrow\sigma^2_\rightarrow/(\sigma^2_\leftarrow + \sigma^2_\rightarrow)$ in the Gaussian approximation. While this situation is not relevant for concentration sensing, our general solution should be applicable to problems of density estimation. The saddle-point approximation usually made in that context [22–24] is expected to work in the same limit as our Gaussian Ansatz;

however, recent work has emphasized the importance of non-Gaussian fluctuations for small datasets [26].

The biological implementation we propose is speculative. An interesting direction would be to identify square-root or similar control of receptor signaling in real biological systems, and interpret them in terms of optimal Bayesian filtering. Another experimental test of our theory could be to measure the accuracy of the chemotactic response in a fluctuating ligand environment, for various values of the mean concentration and fluctuation time scale. Signaling pathways dealing with concentration changes over several orders of magnitude, such as bacterial chemotaxis, typically use adaptation mechanisms to increase the dynamic range of sensing [19]—a feature that is absent from our approach as we neglect noise in the signaling output. Combining adaptation design with ideas from Bayesian estimation could help us gain insight into the fundamental bounds and resource allocation tradeoffs that limit biological information processing.

Acknowledgments. We are grateful to W. Bialek for his insightful comments. We thank the Casa Matemática Oaxaca from the Banff International Research Station where this work was initiated. TM was partially supported by Agence National pour la Recherche (ANR) grant No. ANR-17-ERC2-0025-01 "IRREVERSIBLE" and IN by NSF Grants No. PHY-1410978 and IOS-1822677.

* Corresponding author: thierry.mora@ens.fr

[1] H. C. Berg and E. M. Purcell, Biophysical journal **20**, 193 (1977).
[2] W. Bialek and S. Setayeshgar, Proceedings of the National Academy of Sciences of the United States of America **102**, 10040 (2005).
[3] K. Kaizu, W. De Ronde, J. Paijmans, K. Takahashi, F. Tostevin, P. R. T. Wolde, and P. R. ten Wolde, Biophysical journal **106**, 976 (2014).
[4] G. Aquino, N. S. Wingreen, and R. G. Endres, Journal of Statistical Physics **162**, 1353 (2016).
[5] R. G. Endres and N. S. Wingreen, Physical Review Letters **103**, 158101 (2009).
[6] E. D. Siggia and M. Vergassola, Proceedings of the National Academy of Sciences of the United States of America **110**, E3704 (2013).
[7] J.-B. Lalanne and P. François, Proceedings of the National Academy of Sciences **112**, 1898 (2015).
[8] T. Mora, Physical Review Letters **115**, 038102 (2015).
[9] V. Singh and I. Nemenman, PLoS Computational Biology **13**, 1 (2017).
[10] V. Singh and I. Nemenman, arXiv:1906.08881 (2019).
[11] M. Carballo-Pacheco, J. Desponds, T. Gavrilchenko, A. Mayer, R. Prizak, G. Reddy, I. Nemenman, and T. Mora, Physical Review E **99** (2019).
[12] E. Kussell and S. Leibler, Science (New York, N.Y.) **309**, 2075 (2005).
[13] A. Celani and M. Vergassola, Proceedings of the National Academy of Sciences of the United States of America **107**, 1391 (2010).
[14] X. Shao, A. Mugler, J. Kim, H. J. Jeong, B. R. Levin, and I. Nemenman, PLoS Computational Biology **13**, 1 (2017).
[15] T. Lucas, H. Tran, C. A. Perez Romero, A. Guillou, C. Fradin, M. Coppey, A. M. Walczak, and N. Dostatni, PLoS Genetics **14**, 1 (2018).
[16] G. Malaguti and P. R. ten Wolde, arXiv:1902.09332 (2019).
[17] T. Mora and N. S. Wingreen, Physical Review Letters **104**, 1 (2010).
[18] H. Berg, *E. coli in Motion* (Springer, New York, 2004).
[19] M. D. Lazova, T. Ahmed, D. Bellomo, R. Stocker, and T. S. Shimizu, Proceedings of the National Academy of Sciences of the United States of America **108**, 13870 (2011).
[20] *See Supplemental Material for detailed derivations.*
[21] Z. H. E. Chen, Statistics **182**, 1 (2003).
[22] W. Bialek, C. Callan, and S. Strong, Physical review letters **77**, 4693 (1996).
[23] I. Nemenman and W. Bialek, Physical Review E **65**, 2 (2002).
[24] J. B. Kinney, Physical Review E - Statistical, Nonlinear, and Soft Matter Physics **90** (2014).
[25] J. B. Kinney, Physical Review E - Statistical, Nonlinear, and Soft Matter Physics **92** (2015).
[26] W. C. Chen, A. Tareen, and J. B. Kinney, Physical Review Letters **121**, 160605 (2018).
[27] A. Kolmogoroff, C. R. Acad. Sci. Paris **208**, 2043 (1939).
[28] M. Potters and W. Bialek, Journal de physique. I **4**, 1755 (1994).
[29] L. Goentoro, O. Shoval, M. W. Kirschner, and U. Alon, Molecular Cell **36**, 894 (2009).
[30] C. C. Govern and P. R. ten Wolde, Proceedings of the National Academy of Sciences of the United States of America **2014**, 17486 (2014).