

Optimal resource allocation in cellular sensing systems

Christopher C. Govern and Pieter Rein ten Wolde¹

FOM Institute AMOLF, Science Park 104, 1098 XG Amsterdam, The Netherlands

Edited by Yuhai Tu, IBM T. J. Watson Research Center, Yorktown Heights, NY, and accepted by the Editorial Board October 22, 2014 (received for review June 23, 2014)

Living cells deploy many resources to sense their environments, including receptors, downstream signaling molecules, time, and fuel. However, it is not known which resources fundamentally limit the precision of sensing, like weak links in a chain, and which can compensate each other, leading to trade-offs between them. We present a theory for the optimal design of the large class of sensing systems in which a receptor drives a push–pull network. The theory identifies three classes of resources that are required for sensing: receptors and their integration time, readout molecules, and energy (fuel turnover). Each resource class sets a fundamental sensing limit, which means that the sensing precision is bounded by the limiting resource class and cannot be enhanced by increasing another class—the different classes cannot compensate each other. This result yields a previously unidentified design principle, namely that of optimal resource allocation in cellular sensing. It states that, in an optimally designed sensing system, each class of resources is equally limiting so that no resource is wasted. We apply our theory to what is arguably the best-characterized sensing system in biology, the chemotaxis network of *Escherichia coli*. Our analysis reveals that this system obeys the principle of optimal resource allocation, indicating a selective pressure for the efficient design of cellular sensing systems.

cell signaling | thermodynamics | design principles | chemotaxis | information transmission

Biochemical networks are the information-processing devices of life. Like any device, they require resources to be built and run. Components are needed to construct the network, space is required to accommodate the components, time is needed to process the information, and energy is required to make the components and operate the network. These resources constrain the design and performance of any biochemical network. However, it is not clear which resources are indispensable, thus fundamentally limiting the performance of the network, and which resources might trade-off against each other. Here, we consider the interplay among cellular resources, network design, and performance in a canonical biochemical function, namely sensing the environment.

Living cells can measure chemical concentrations with extraordinary precision (1–3), raising the question what sets the fundamental limit to the accuracy of chemical sensing (1). Cells measure chemical concentrations via receptors on their surface. These measurements are inevitably corrupted by noise that arises from the stochastic arrival of ligand molecules by diffusion and from the stochastic binding of the ligand to the receptor. Berg and Purcell pointed out that the sensing error is fundamentally bounded by this noise extrinsic to the cell, but that cells can reduce the error by taking multiple independent measurements (1). One way to increase the number of measurements is to add more receptors (1, 4). Another is to take more measurements per receptor over time; here, the cell infers the concentration not from the instantaneous number of ligand-bound receptors, but rather from the average receptor occupancy over an integration time T (1, 4–11).

This time integration has to be performed by the signaling networks that transmit the information from the surface of the cell to its interior (10). Although the work of Berg and Purcell and subsequent studies identify time and the number of receptors

as resources that limit the accuracy of sensing, the fundamental limits that have emerged ignore the cost of making and operating the signaling network. Making proteins is costly; producing proteins that confer no benefit to the cell can slow down bacterial growth (12). Moreover, many networks are driven out of thermodynamic equilibrium by the continuous turnover of fuel molecules such as ATP, leading to the dissipation of heat (13–17). In fact, one can estimate that the fuel needed to operate a sensory network is comparable to that to make new components after cell division (*SI Text*).

In this manuscript, we present a theory for the optimal design of sensing systems, which maximizes sensing precision given the available cellular resources. The theory applies to the large class of sensing systems in which a receptor drives a Goldbeter–Koshland push–pull network (18). These systems are ubiquitous in prokaryotic and eukaryotic cell signaling (19): examples include GTPase cycles, as in the Ras system, phosphorylation cycles, as in MAPK cascades, and two-component systems like the chemotaxis system of *Escherichia coli*.

We derive for this class of systems how the sensing accuracy depends on not only the number of receptors and their integration time, but also on the resources required to build and operate the downstream signaling network: the copies of signaling molecules and fuel. This allows us to address the following questions: How do the sensing limits set by the latter resources compare with the canonical limit of Berg and Purcell, which is set by the resources time and the number of receptors? How does the limit set by one resource depend on the levels of the other resources? Can resources compensate each other to achieve a desired sensing precision, leading to trade-offs between them, or are the limits set by the respective resources fundamental, i.e., independent of the levels of the other resources? In addition, what do these relationships imply for the optimal design of a system that maximizes sensing precision?

Significance

Cells continually have to sense their environments to make decisions—to stay put or move, to differentiate or proliferate, or even to live or die. However, they are thwarted by noise at the cellular scale. Cells use signaling networks to filter this noise as much as possible and sense accurately. To operate these networks, resources are required: time, protein copies, and energy. We present a theory for the optimal design of cellular sensing systems that maximize sensing precision given these resources. It reveals a new design principle, namely that of optimal resource allocation. It describes how these resources must be allocated so that none are wasted. We show that the chemotaxis network of *Escherichia coli* obeys this principle.

Author contributions: C.C.G. and P.R.t.W. designed research, performed research, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. Y.T. is a guest editor invited by the Editorial Board.

Freely available online through the PNAS open access option.

¹To whom correspondence should be addressed. Email: tenwolde@amolf.nl.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1411524111/-DCSupplemental.

We find that the resource limitations of these systems emerge naturally when the signaling networks are viewed as devices that discretely, rather than continuously, sample the receptor state via collisions of the signaling molecules with the receptor proteins. This analysis reveals that three classes of resources are required: (i) receptors and their integration time, (ii) copies of downstream molecules, and (iii) energy (fuel). Indeed, these classes cannot compensate each other: each imposes a sensing limit, and it is the limiting class that imposes the fundamental limit on the accuracy of sensing. However, there can be trade-offs within each class of resources. Receptors and integration time trade-off against each other in achieving a desired sensing accuracy, and power and response time trade-off against each other to meet the energy requirement for taking a measurement.

Our theory makes a strong prediction for the optimal design of sensing systems. Because the fundamental resource classes cannot compensate each other in achieving a desired sensing precision, any class that is not limiting the sensing precision is in excess and thus wasted. This naturally leads to a previously unidentified design principle, namely that of optimal resource allocation. It states that in an optimally designed sensing system each fundamental class of resources is equally limiting so that no resource is wasted. We test this prediction for the chemotaxis system of *Escherichia coli*, which is a specific example of the class of push-pull sensing systems (Fig. 1). Our analysis reveals that this network obeys the principle of optimal resource allocation. This indicates that there is a selective pressure on not only the topology of sensing networks that enhances robustness of adaptation (20–22), but also on the efficient allocation of cellular resources for precise sensing.

Results

Sensing at the Molecular Level. We consider a cell with R_T receptor proteins that independently bind ligand L , $R + L \rightleftharpoons RL$ (Fig. 2A). The receptor drives a push-pull network, which is a canonical nonequilibrium motif in prokaryotic and eukaryotic cell signaling

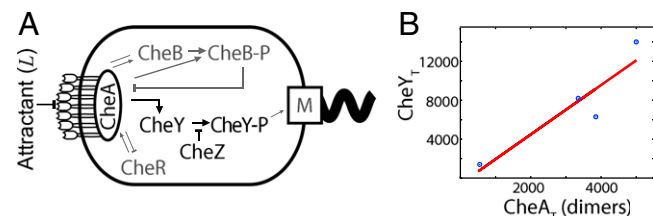


Fig. 1. The chemotaxis network of *E. coli* obeys the principle of optimal resource allocation, which states that in an optimally designed system each cellular resource is equally limiting. (A) Cartoon of the sensing system. The receptor is via the adaptor protein CheW associated with the kinase CheA. This complex, coarse-grained as R in our model, can bind extracellular ligand L and activate the intracellular messenger protein CheY (x in our model) by phosphorylating it; phosphorylated CheY controls the rotation direction of the motor. Deactivation, i.e., dephosphorylation, of CheY is catalyzed by the phosphatase CheZ; the effect of CheZ is coarse-grained into the deactivation rate. The proteins CheR and CheB, which implement adaptation, have been omitted, because we are interested in the lower bound on the accuracy of sensing in static environments. (B) The principle of optimal resource allocation, Eq. 5, predicts that the number of CheY proteins, X_T , scales linearly with the number of receptor–CheA complexes, R_T , with a slope given by the relaxation time of the signaling network, τ_r , over the correlation time of the receptor ligand-binding state, τ_c . Plotted are data from ref. 23 for two *E. coli* strains under two different growth conditions; the number of CheA dimers is a proxy for the number of receptor–CheA complexes. The line is a best fit to the data, having a slope of ≈ 3 . The resource allocation principle, Eq. 5, thus predicts that $\tau_r/\tau_c \approx 3$. This is on the same order of magnitude as that given by the relaxation time, $\tau_r \approx 100$ ms (24), and correlation time $\tau_c \approx 10$ ms, estimated from the measured receptor–ligand dissociation constant (25) and association rate (26).

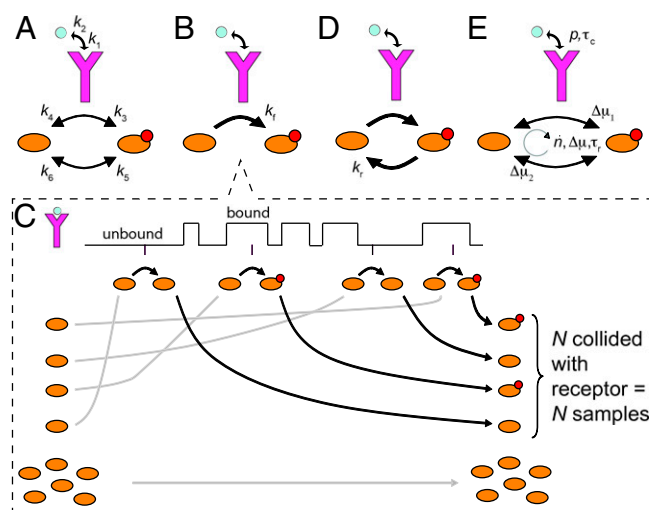


Fig. 2. Sensing at the molecular level. The sensing precision in terms of the rate constants $\{k_i\}$ (A) does not reveal the resource requirements (Eq. S8). To reveal these, the signaling network is viewed as a device that discretely samples the ligand-binding state of the receptor. The accuracy of sensing depends on how the samples are taken (B and C), erased (D), and on how reliable they are (E). (B) The ligand-bound receptor drives the modification of a downstream readout (i.e., the push-pull network $RL + x \rightarrow RL + x^*$). (C) The signaling network in B discretely samples the receptor state, illustrated for one receptor. The states of the receptor over time are encoded in the states of the N molecules that collided with it: the readout is modified if the receptor is bound; it is unmodified. Molecules that collide with the unbound receptor are indistinguishable from those that have never collided, leading to an additional error. (D) Active molecules can be degraded, erasing samples. (E) All reactions are in principle reversible, compromising the encoding of the receptor state into the readout. The sensing error is determined by collective variables that reveal the resource requirements for sensing: the probability p that the receptor is bound to ligand, the receptor–ligand correlation time τ_c , the flux \bar{n} , the relaxation time τ_r , and the free-energy drops $\Delta\mu_1$ and $\Delta\mu_2$ across the activation and deactivation reactions of the readout, respectively.

(19). In these systems, the receptor itself or the enzyme associated with it, such as CheA in *E. coli* (Fig. 1), catalyzes the chemical modification of a readout protein x , such as CheY. Active readout molecules x^* can decay spontaneously or be deactivated by an enzyme, like the phosphatase CheZ in *E. coli*.

The cell infers the ligand concentration c from the instantaneous concentration of the output x^* , by inverting the mean input–output relation $\bar{x}^*(c)$. Linearizing $\bar{x}^*(c)$ and using error propagation, the expected fractional error in the concentration estimate is then as follows (1, 5, 11):

$$\left(\frac{\delta c}{c}\right)^2 = \frac{1}{c^2} \frac{\sigma_{x^*}^2}{\left(\frac{dx^*}{dc}\right)^2} \quad [1]$$

The error is low if the readout responds sensitively to changes in ligand concentration, as measured by the gain dx^*/dc , but is not noisy, as quantified by the variance $\sigma_{x^*}^2$.

We can compute $\sigma_{x^*}^2$ from the linear-noise approximation (SI Text), and using Eq. 1, this yields Eq. S8 for the sensing error. It is a complicated expression in terms of the eight fundamental variables in the system: the six rate constants describing the forward and reverse rates of the three reactions (including ligand–receptor binding), and the total copy numbers X_T and R_T (Fig. 2A).

Inspired by the analysis of a simpler system, we can arrive at a much more illuminating expression for the sensing error, by viewing the signaling network as a device that samples the receptor state

(SI Text). The general principle is that the activation reaction, $x + RL \xrightarrow{k_f} x^* + RL$ generates samples of the ligand-binding state of the receptor by storing the receptor state in the stable modification states of the readout molecules (Fig. 2 B and C). Readout molecules that collide with a ligand-bound receptor are activated, whereas those that collide with an unbound receptor remain inactive. In this way, each readout molecule that has interacted with the receptor provides a memory or sample of the ligand-occupation state of that receptor molecule; collectively, the readout molecules encode the history of the receptor states. Intuitively, we expect that if there are N receptor–readout interactions, then the cell has N samples of the receptor state and the error in the concentration estimate, $\delta c/c$, is reduced by a factor of \sqrt{N} , or less if the samples are not independent. To derive the effective number of independent samples, we need to consider not only the creation of samples, but also the erasure of samples and the quality of the samples (Fig. 2 D and E). The decay of the readout, $x^* \xrightarrow{k_r} x$ (Fig. 2D), is equivalent to discarding or erasing samples. Additionally, reactions are microscopically reversible, which means that readout activation can occur independently of the receptor, $x \xrightarrow{k_{-f}} x^*$, and receptor-mediated modifications can occur in the wrong direction, $x^* + RL \xrightarrow{k_{-f}} x + RL$ (Fig. 2E). These reverse reactions compromise the encoding of the receptor state into the readout: an active x^* molecule no longer encodes the ligand-bound state of the receptor at a previous time with 100% fidelity, because it could have been activated independently of the receptor; similarly, x , rather than x^* , may reflect a modification by a ligand-bound receptor. These reverse reactions thus reduce the reliability of a receptor sample. Energy is needed to break time reversibility and to protect the coding.

How receptor samples are taken (Fig. 2 B and C), erased (Fig. 2D), and how they are stored in the readout x (Fig. 2E), determine the number of receptor samples, their independence, and their accuracy, which together set the sensing precision (SI Text):

$$\left(\frac{\delta c}{c}\right)^2 = \frac{1}{p(1-p)} \frac{1}{\bar{N}_I} + \frac{1}{(1-p)^2} \frac{1}{\bar{N}} \quad [2]$$

Here, p is the probability that a receptor is bound to ligand, and $1/(p(1-p))$ is the “instantaneous error,” i.e., the sensing error based on a single concentration estimate via a single receptor. The quantity \bar{N}_I , discussed below, is the average number of samples that are independent. The second term, with \bar{N} the total number of samples, accounts for the fact that the cell cannot distinguish between those molecules x that have collided with an unbound receptor (and hence provide information on the receptor occupancy), and those that have not collided with the receptor at all (Fig. 2C; SI Text). However, when p is small and/or \bar{N} is large, the second term is small compared with the first. Eq. 2 then shows that the sensing error has a form that one would expect for a sampling protocol: the sensing error is that of an estimate based on a single concentration measurement, $1/(p(1-p))$, divided by the average number of independent measurements, \bar{N}_I .

The number of independent measurements \bar{N}_I can be expressed in terms of collective variables that, as we will show, describe the resource limitations of the cell (Fig. 2E):

$$\bar{N}_I = \frac{1}{\underbrace{(1+2\tau_c/\Delta)}_{f_I}} \underbrace{\frac{(e^{\Delta\mu_1}-1)(e^{\Delta\mu_2}-1)}{e^{\Delta\mu}-1}}_{\bar{N}_{\text{eff}}} \underbrace{\frac{\bar{N}}{p}}_{\bar{n}\tau_r} \quad [3]$$

This expression has a clear interpretation. Cells count only those samples created less than a relaxation time τ_r in the past; nothing that happened earlier can influence the current state, including its ability to sense. Hence, τ_r is the effective integration time. The

quantity \bar{n} is the flux of x across the cycle of activation by the receptor and deactivation; it is given by $\bar{n} = k_f \bar{x} R_T p - k_{-f} \bar{x}^* R_T p$, where \bar{x} and \bar{x}^* are the average number of x and x^* in steady state. The product $\bar{n}\tau_r$ is thus the number of cycles of readout molecules involving collisions with ligand-bound receptor molecules during the system’s relaxation time τ_r . The quantity $\bar{n}\tau_r/p$ is the total number of readout cycles involving collisions with receptor molecules, be they ligand bound or not. It is thus the total number of receptor samples taken during τ_r , \bar{N} .

Not all of these samples are reliable. The effective number of samples taken during τ_r is $\bar{N}_{\text{eff}} = q\bar{N}$, where $0 \leq q \leq 1$ measures the quality of each sample. Here, $\Delta\mu_1$ and $\Delta\mu_2$ are the average free-energy drops across the activation and deactivation pathway respectively, in units of $k_B T$ (Fig. 2E); $\Delta\mu = \Delta\mu_1 + \Delta\mu_2$ is the total free-energy drop across the cycle. When $\Delta\mu = \Delta\mu_1 = \Delta\mu_2 = 0$, an active readout molecule is as likely to be created by the ligand-bound receptor as it is created spontaneously and there is no coding and no sensing; indeed, in this limit, $q = 0$ and $\bar{N}_{\text{eff}} = 0$. In contrast, when $\Delta\mu_1, \Delta\mu_2 \rightarrow \infty$, $q \rightarrow 1$ and $\bar{N}_{\text{eff}} \rightarrow \bar{N}$.

The factor f_I denotes the fraction of samples that are independent. It depends on the correlation time τ_c of receptor–ligand binding and on the time interval $\Delta = 2\tau_r/(\bar{N}_{\text{eff}}/R_T)$ between samples of the same receptor. Samples farther apart are more independent.

Fundamental Resources and Trade-Offs. We can use Eqs. 2 and 3 to understand how cellular resources limit the precision of sensing. A resource or combination of resources that fundamentally limits sensing is a (collective) variable Q_i that, when fixed, puts a non-zero lower bound on the sensing error, no matter how the other variables are varied. A fundamental resource class Q_i is thus mathematically defined by the following: $\text{MIN}_{Q_i=\text{const}}(\delta c/c)^2 = f(\text{const}) > 0$. To find these classes, we numerically or analytically minimized the sensing error, constraining (combinations of) variables yet optimizing over the other variables. As we show in SI Text, when only $R_T\tau_r/\tau_c$ is constrained, $(\delta c/c)^2 \geq 4/(R_T\tau_r/\tau_c)$; when only X_T is fixed, $(\delta c/c)^2 \geq 4/X_T$; when only $\dot{w}\tau_r$ is limiting, $(\delta c/c)^2 \geq 4/(\dot{w}\tau_r)$. When all resources are present in finite amounts, the minimum sensing error is set by the highest lower bound:

$$\left(\frac{\delta c}{c}\right)^2 \geq \text{MAX}\left(\frac{4}{R_T\tau_r/\tau_c}, \frac{4}{X_T}, \frac{4}{\dot{w}\tau_r}\right). \quad [4]$$

Fig. 3 A–C shows that the resource classes $R_T\tau_r/\tau_c$, X_T , and $\dot{w}\tau_r$ are indeed fundamental: the minimum sensing error is bounded by the limiting class and cannot be reduced by increasing another resource. Clearly, increasing a single resource, e.g., X_T , cannot reduce the sensing error indefinitely. The sensing error will eventually plateau, namely when it becomes limited by another resource, e.g., $R_T\tau_r/\tau_c$. These fundamental resource classes thus cannot compensate for each other in achieving a required sensing precision and hence do not trade-off against each other. However, within these classes, trade-offs are possible. We now elucidate why the fundamental classes cannot compensate each other, whereas resources within a given class can, leading to trade-offs between them.

Time/Receptor Copy Numbers, $R_T\tau_r/\tau_c$. An independent sample of the same receptor can be taken roughly every $2\tau_c$. Naturally, samples can be taken more frequently. In fact, cells can time-integrate as in the theory of Berg and Purcell (1): if $X_T \rightarrow \infty$, the receptors are sampled infinitely fast and $\Delta \rightarrow 0$ and $\bar{N}_{\text{eff}} \rightarrow \infty$. However, increasing X_T cannot reduce the sensing error ad infinitum, because the number of receptor samples that are independent will saturate at the Berg–Purcell factor, $R_T\tau_r/\tau_c$, and the sensing error $(\delta c/c)^2$ will plateau at $4/(R_T\tau_r/\tau_c)$. Indeed, $R_T\tau_r/\tau_c$ is the maximum number of independent concentration

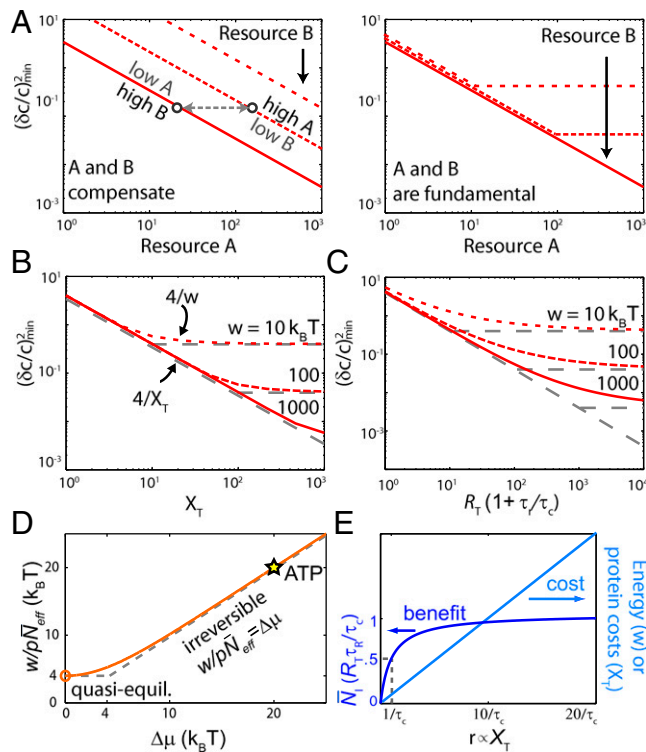


Fig. 3. Trade-offs in nonequilibrium sensing. (A) When two resources A and B compensate each other, one resource can always be decreased without affecting the sensing error, by increasing the other resource; concomitantly, increasing a resource will always reduce the sensing error. When both resources are instead fundamental, the sensing error is bounded by the limiting resource and cannot be reduced by increasing the other. (B and C) The three classes time/receptor copies, copies of downstream molecules, and energy are all required for sensing, with no trade-offs among them (Fig. 4). The minimum sensing error obtained by minimizing Eq. 2 is plotted for different combinations of (B) X_T and w , and (C) $R_T(1 + \tau_r/\tau_c)$ and w (SI Text). The curves track the bound for the limiting resource indicated by the gray lines, showing that the resources do not compensate each other. The plot for the minimum sensing error as a function of $R_T(1 + \tau_r/\tau_c)$ and X_T is identical to that of (C) with w replaced by X_T . (D) The energy requirements for sensing. In the irreversible regime ($\Delta\mu \rightarrow \infty$), the work to take one sample of a ligand-bound receptor, $w/(p\bar{N}_{\text{eff}})$, equals $\Delta\mu$, because each sample requires the turnover of one fuel molecule, consuming $\Delta\mu$ of energy. In the quasiequilibrium regime ($\Delta\mu \rightarrow 0$), each effective sample of the bound receptor requires $4k_B T$, which defines the fundamental lower bound on the energy requirement for taking a sample. When $\Delta\mu = 0$, the network is in equilibrium and both w and \bar{N}_{eff} are 0. ATP hydrolysis provides $20k_B T$, showing that phosphorylation of readout molecules makes it possible to store the receptor state reliably. The results are obtained from Eq. 3 with $\Delta\mu_1 = \Delta\mu_2 = \Delta\mu/2$. (E) Sampling more than once per correlation time requires more resources, although the benefit is marginal. As the sampling rate is increased by increasing the readout copy number X_T , the number of independent measurements \bar{N}_i saturates at the Berg–Purcell limit $R_T\tau_r/\tau_c$, but the energy consumption and protein cost ($\propto X_T$) continue to rise.

measurements—the total number of receptors R_T times the maximum number of independent measurements per receptor τ_r/τ_c . This shows that there is no fundamental relationship between sensing and receptor copy number: the latter can be traded against time to reach a desired sensing precision. Essentially, the error is determined by the total number of samples and it does not matter, as long as the samples are independent, whether these samples are from the same receptor over time or from many receptors at the same time.

Downstream Readout Molecules, X_T . The concentration measurements need to be stored in the readout molecules. Each readout

molecule provides at most one sample, because at any given time it exists in only one modification state, regardless of how many times it has collided with the receptor or how long the integration time τ_r is. There is no mechanistic sense in which a single molecule “integrates” the receptor state. As a consequence, no matter how the network is designed, how much time or energy it uses, or how many receptors it has, cells are fundamentally limited by the pool of readout molecules: the sensing error $(\delta c/c)^2 \geq 4/X_T$.

Energy, $\dot{w}\tau_r$. The free-energy drop across a cycle, $\Delta\mu$, must be provided by a fuel molecule such as ATP. The power, the rate at which the fuel molecules do work, is $\dot{w} = \dot{n}\Delta\mu$, and the total work performed during τ_r is $w \equiv \dot{w}\tau_r$. This work is spent on taking samples of receptor molecules that are bound to ligand, because only they can modify downstream readout molecules. Hence, the work needed to take one effective sample of a ligand-bound receptor is $w/(p\bar{N}_{\text{eff}})$, with \bar{N}_{eff} given by Eq. 3. Fig. 3D shows this quantity as a function of $\Delta\mu$. While $w/(p\bar{N}_{\text{eff}}) = \Delta\mu/q$ increases continuously with $\Delta\mu$, two limiting regimes can be observed.

When $\Delta\mu > 4k_B T$, the work to take one effective sample of a ligand-bound receptor becomes simply $w/(p\bar{N}_{\text{eff}}) = \Delta\mu$. In this regime, the readout reactions are essentially irreversible, $q \rightarrow 1$, and each sample requires the turnover of one fuel molecule, using $\Delta\mu$ of energy. Energy limits the accuracy of sensing, not because it limits the reliability q of each sample, but because it limits the total number of samples $\bar{N}_{\text{eff}} = \dot{n}\tau_r/p$ by limiting the receptor sampling frequency \dot{n} : $(\delta c/c)^2 \geq 1/(\dot{n}\tau_r) = \Delta\mu/(\dot{w}\tau_r)$. Intriguingly, this bound suggests that for a fixed amount of energy, $w = \dot{w}\tau_r$, spent during the relaxation time τ_r , the sensing error can be reduced to zero by reducing $\Delta\mu$ to zero. However, this lower bound only applies when $q \rightarrow 1$, i.e., when $\Delta\mu > 4k_B T$.

When $\Delta\mu < 4k_B T$, the system transitions to a quasiequilibrium regime in which each fuel molecule provides a small but nonzero amount of energy. In this regime, the system can still consume significant amounts of energy when the fuel molecules are consumed at a rapid rate \dot{n} by many distinct readout molecules. In the limit that $\dot{n} \rightarrow \infty$ and $\Delta\mu \rightarrow 0$ at fixed $\dot{w} = \dot{n}\Delta\mu$, the effective number of samples given by Eq. 3 reduces to $\bar{N}_{\text{eff}} \rightarrow \dot{w}\tau_r/(4p)$. Each readout–receptor interaction corresponds to an increasingly noisy measurement of the receptor state ($q \rightarrow 0$), but many noisy measurements ($\bar{N} = \dot{n}\tau_r/p \rightarrow \infty$) contain the same information as 1 perfect measurement—provided that collectively at least $4k_B T$ was spent on them. Indeed, as Fig. 3D shows, $4k_B T$ is the fundamental lower bound on the work needed to take one accurate sample of a ligand-bound receptor. It puts another bound on the sensing error: $(\delta c/c)^2 \geq 4/(\dot{w}\tau_r)$. The bound can be reached when $R_T\tau_r/\tau_c$ and X_T are not limiting, and $\Delta\mu \rightarrow 0$.

Eq. 4 shows that the sensing precision depends on the work done in the past relaxation time, $w = \dot{w}\tau_r$, setting up a trade-off among speed, power, and accuracy, as found in adaptation (13). When the response needs to be rapid, τ_r needs to be small and the power demand is high: the samples, which require energy, must be taken close together in time. However, when the cell can wait a long time τ_r before responding, the power \dot{w} required to make w large can be infinitesimal: the samples can be created far apart in time. There is no minimum power requirement for sensing.

Optimal Resource Allocation. Because the fundamental resource classes cannot compensate each other in achieving a desired sensing precision, any class that is in excess of the minimum amount necessary to achieve that precision is wasted. For example, the benefit of sampling the receptor faster by increasing X_T in reducing the sensing error saturates, whereas the total protein and energetic costs continue to rise with X_T (Fig. 3E). To the extent that all resources affect growth, evolutionary pressure should tend to drive systems so that no resource is wasted, which

occurs when all are equally limiting. Resource-optimal systems sample the receptor about once per correlation time and use just enough fuel and downstream molecules to do so. Quantitatively, from Eq. 4, all resources are equally limiting when

$$R_T \tau_r / \tau_c \approx X_T \approx w. \quad [5]$$

In an optimal sensing system, the number of independent concentration measurements $R_T \tau_r / \tau_c$ equals the number of readout molecules X_T that store these measurements and equals the work (in units of $k_B T$) to create the samples.

Comparison with Experiment. Eq. 5 makes a strong prediction for the optimal design of the large class of sensing systems that are based on the push-pull motif. We can test this prediction for the chemotaxis system of *E. coli* (Fig. 1), which has been well characterized experimentally. In this system, the receptor forms a complex with the kinase CheA. This complex, which is coarse-grained into R , can bind the ligand L and activate the intracellular messenger protein CheY (x) by phosphorylating it. Deactivation of CheY is catalyzed by CheZ, the effect of which is coarse-grained into the deactivation rate (*SI Text*).

The number of chemotaxis proteins depends on the growth rate: the number of receptors and CheY proteins varies as much as 10-fold as a function of strain and growth medium (23). Interestingly, however, these variations occur in concert for all components and thus hardly change their relative amounts (23). This is the scaling behavior predicted by Eq. 5, assuming that τ_r is robust to variations in the growth rate.

Not only the scaling of the number of CheY proteins, X_T , with the number of receptor-CheA complexes, R_T , can be tested, but also the magnitude of their ratio. A fit of the data of Li and Hazelbauer, shown in Fig. 1B, shows that $X_T / R_T \approx 3$ for different strains and growth media (23). Eq. 5 thus predicts that $\tau_r / \tau_c \approx 3$. The relaxation rate τ_r^{-1} is $\approx 2 \text{ s}^{-1}$ for the attractant response and $\approx 20 \text{ s}^{-1}$ for the repellent response (24), yielding $\tau_r \approx 100 \text{ ms}$. Hence, Eq. 5 predicts that $\tau_c \approx 30 \text{ ms}$. This prediction can be tested, assuming that the correlation time τ_c of the receptor-CheA complex is that of receptor-ligand binding. Specifically, we can estimate τ_c from the receptor-ligand dissociation rate k_{off} as $\tau_c \approx 1 / (2k_{\text{off}})$, ($p \approx 0.5$). The dissociation constant of Tar-aspartate (receptor-ligand) binding $K_D \approx 0.1 \mu\text{M}$ (25), and with an association rate $k_{\text{on}} \approx 10^9 \text{ M}^{-1} \cdot \text{s}^{-1}$ (26), this yields $k_{\text{off}} \approx 100 \text{ s}^{-1}$ and an estimated correlation time $\tau_c \approx 10 \text{ ms}$, in line with the prediction of Eq. 5.

Eq. 5 also predicts that the total number of CheY molecules X_T equals the chemical work $w = \dot{n} \Delta \mu \tau_r$ (in units of $k_B T$) to phosphorylate CheY during τ_r . In steady state, the flux \dot{n} of CheY phosphorylation balances the flux of CheYp dephosphorylation. The latter equals the inverse lifetime $\tau_r^{-1} \approx \tau_r^{-1}$ times the number of CheYp molecules in steady state, αX_T , where α is the fraction of CheY that is phosphorylated. This yields $w \approx \alpha \Delta \mu X_T$. The fraction $\alpha \approx 0.16$ (24). CheY phosphorylation is driven by ATP hydrolysis, which means that $\Delta \mu = 20 k_B T$. Hence $w \approx 3 X_T$, thus on the order of X_T , as Eq. 5 predicts. We thus argue that in the chemotaxis system of *E. coli*, the resources are optimally allocated, i.e., according to Eq. 5.

Discussion

Fig. 4 summarizes our analysis. The sensing precision is limited by three classes of resources. The resource class time/receptors $R_T \tau_r / \tau_c$ determines the maximum total number of independent concentration measurements that can be taken. However, these measurements need to be stored in downstream molecules. Moreover, energy is needed to store the samples reliably and to protect the coding. These three classes of resources are indeed fundamental. The sensing error is bounded by the limiting class like the weakest link in a chain, and other classes cannot compensate

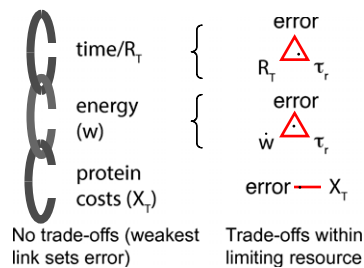


Fig. 4. The relationship between resources and the precision of biochemical sensing. The sensing precision is fundamentally limited by time and receptor copies, energy, and copies of downstream readouts. These three classes of resources cannot compensate each other, and it is the limiting resource that sets the fundamental limit to the precision of sensing. Within each class, however, trade-offs are possible: Power can be traded against speed to meet the energy requirement for reaching a desired sensing accuracy, whereas time can be traded against the number of receptors.

for it. For example, adding receptors and readout molecules does not improve sensing if not enough energy is used to take the samples (Fig. 3B); similarly, waiting more time to take another sample is not beneficial if the cell has no more readout molecules left to write the sample to, or cannot expend energy fast enough to accomplish the writing (Fig. 3C). However, within the fundamental resource classes, trade-offs are possible: time can be traded against the number of receptors to reach a required number of measurements, whereas power can be traded against speed to meet the energy requirement for a desired sensing accuracy. These design principles are in marked contrast to those of equilibrium sensing systems, which are not driven out of equilibrium via fuel turnover: the sensing precision of these systems is limited by the number of receptors; downstream networks can never improve the accuracy of sensing (27).

We find that at least $4k_B T$ is needed for reliably encoding a measurement. One of the most widely used coding strategies is phosphorylation, which requires ATP. In vivo, ATP hydrolysis provides about $20k_B T$. This is sufficient to take one receptor sample essentially irreversibly (Fig. 3D), which means that q reaches unity. Readout phosphorylation thus makes it possible to store the receptor state reliably.

Nonequilibrium networks can exhibit more complicated features than those of the simple push-pull motif, as in the MAPK cascade. The molecular picture for time integration suggests that our results hold generally, even in these more complicated systems. Indeed, we find the same or more severe resource limitations in signaling cascades and in networks with simple negative or positive feedback (*SI Text*). Although cascades can increase the response time (10), which increases information transfer, they do not make sensing more efficient in terms of energy or readout molecules.

In an optimally designed system, each fundamental resource is equally limiting. This leads to a specific prediction for the design of an optimal sensing system: the integration time, energy, and copy numbers of receptor and readout should satisfy Eq. 5. Importantly, this design principle of optimal resource allocation is independent of the resource costs. This is because the sensing precision is bounded by the limiting resource class; resources that are in excess cannot improve sensing and are thus wasted, no matter how cheap they are. Fig. 3C and D show that, even close to the optimum where all resources are equally limiting, the minimum error closely follows the lower bound of Eq. 4 (*SI Text*), supporting the idea that Eq. 5 is indeed fairly insensitive to the resource costs. It explains perhaps why Eq. 5 so successfully predicts the design of the *E. coli* chemotaxis system.

The optimal trade-off between nonfundamental resources within the fundamental resource classes will depend on their

fitness costs and benefits. For example, how receptors are traded against time in reaching a desired sensing precision will depend on the benefit of a fast response time and the cost of making the receptor proteins.

To understand how *E. coli* moves in a concentration gradient, we have to understand not only how the sensing system filters high-frequency ligand-binding noise by time averaging the receptor state—the topic of this study—but also how, on much longer timescales, the adaptation system computes the change in the concentration and filters low-frequency noise induced by the cell's random motion in the concentration gradient (28). Recently, Lan et al. (13) found a trade-off between the speed, power, and accuracy of adaptation, which mirrors the trade-off between power, speed, and precision of sensing observed here. Interestingly, the adaptation and sensing system share the receptor. In fact, the adaptation system continually performs work to keep the receptor activity close to 0.5. There is thus an energetic adaptation cost associated with the receptors, in addition to the energetic cost of synthesizing them. This adaptation cost will affect the trade-off between the nonfundamental resources receptors and time. It will not, however, affect the design principle of optimal resource allocation, which is based on the fundamental resource classes and hence insensitive to resource costs.

Whether other sensing systems satisfy the design principle of Eq. 5 remains an interesting question. Two-component systems are ideal for testing this once kinetic data and protein expression levels become available (29). Eq. 5 not only makes predictions for individual systems but also predicts that the fundamental

resources should vary proportionally to each other across different systems. For example, the relation predicts that the lifetime τ_r of the modified readout should increase, *ceteris paribus*, with its expression level X_T . The design principle of Eq. 5 can also be used to construct optimal synthetic networks that minimize resource consumption.

Finally, the process of sampling a time series, like the receptor state over time, defines a specific, familiar computation that could be conducted by any machine; it is instantiated in the biochemical system by the readout–receptor pair. We find that the free-energy drops across the “measurement” and “erasure” steps, $\Delta\mu_1$ and $\Delta\mu_2$, should be identical to minimize the energetic cost, even though the fuel molecule need only be involved in one of the reactions, preparing a nonequilibrium state that relaxes via the other. This allocation of energy differs from that typically considered in the computational literature, in which only the erasure step requires energy (30). In the cellular system, both steps are computational erasures: although only the erasure step erases memory of the receptor state, both steps erase the state of the molecule involved in the collision. Interestingly, when $p = 0.5$, the average work to measure the state of the receptor is $2k_B T$, which is perhaps surprisingly close to the Landauer bound, $k_B T \ln(2)$ (30).

ACKNOWLEDGMENTS. We thank Andrew Mugler, Thomas Ouldrige, and Tom Shimizu for critical readings of sections of the manuscript. This work is part of the research programme of the Foundation for Fundamental Research on Matter (FOM), which is part of the Netherlands Organisation for Scientific Research (NWO).

- Berg HC, Purcell EM (1977) Physics of chemoreception. *Biophys J* 20(2):193–219.
- Sourjik V, Berg HC (2002) Receptor sensitivity in bacterial chemotaxis. *Proc Natl Acad Sci USA* 99(1):123–127.
- Ueda M, Shibata T (2007) Stochastic signal processing and transduction in chemotactic response of eukaryotic cells. *Biophys J* 93(1):11–20.
- Wang K, Rappel WJ, Kerr R, Levine H (2007) Quantifying noise levels of intercellular signals. *Phys Rev E Stat Nonlin Soft Matter Phys* 75(6 Pt 1):061905.
- Bialek W, Setayeshgar S (2005) Physical limits to biochemical signaling. *Proc Natl Acad Sci USA* 102(29):10040–10045.
- Rappel WJ, Levine H (2008) Receptor noise and directional sensing in eukaryotic chemotaxis. *Phys Rev Lett* 100(22):228101.
- Endres RG, Wingreen NS (2009) Maximum likelihood and the single receptor. *Phys Rev Lett* 103(15):158101.
- Hu B, Chen W, Rappel WJ, Levine H (2010) Physical limits on cellular sensing of spatial gradients. *Phys Rev Lett* 105(4):048104.
- Mora T, Wingreen NS (2010) Limits of sensing temporal concentration changes by single cells. *Phys Rev Lett* 104(24):248101.
- Govern CC, ten Wolde PR (2012) Fundamental limits on sensing chemical concentrations with linear biochemical networks. *Phys Rev Lett* 109(21):218103.
- Kaizu K, et al. (2014) The Berg-Purcell limit revisited. *Biophys J* 106(4):976–985.
- Dekel E, Alon U (2005) Optimality and evolutionary tuning of the expression level of a protein. *Nature* 436(7050):588–592.
- Lan G, Sartori P, Neumann S, Sourjik V, Tu Y (2012) The energy-speed-accuracy tradeoff in sensory adaptation. *Nat Phys* 8(5):422–428.
- De Palo G, Endres RG (2013) Unraveling adaptation in eukaryotic pathways: Lessons from protocells. *PLoS Comput Biol* 9(10):e1003300.
- Mehta P, Schwab DJ (2012) Energetic costs of cellular computation. *Proc Natl Acad Sci USA* 109(44):17978–17982.
- Barato AC, Hartich D, Seifert U (2013) Information-theoretic versus thermodynamic entropy production in autonomous sensory networks. *Phys Rev E Stat Nonlin Soft Matter Phys* 87(4):042104.
- Skoge M, Naqvi S, Meir Y, Wingreen NS (2013) Chemical sensing by nonequilibrium cooperative receptors. *Phys Rev Lett* 110(24):248102.
- Goldbeter A, Koshland DE, Jr (1981) An amplified sensitivity arising from covalent modification in biological systems. *Proc Natl Acad Sci USA* 78(11):6840–6844.
- Alon U (2007) *Introduction to Systems Biology: Design Principles of Biological Networks* (CRC Press, Boca Raton, FL).
- Barkai N, Leibler S (1997) Robustness in simple biochemical networks. *Nature* 387(6636):913–917.
- Alon U, Surette MG, Barkai N, Leibler S (1999) Robustness in bacterial chemotaxis. *Nature* 397(6715):168–171.
- Kollmann M, Lovdok L, Bartholomé K, Timmer J, Sourjik V (2005) Design principles of a bacterial signalling network. *Nature* 438(7067):504–507.
- Li M, Hazelbauer GL (2004) Cellular stoichiometry of the components of the chemotaxis signaling complex. *J Bacteriol* 186(12):3687–3694.
- Sourjik V, Berg HC (2002) Binding of the *Escherichia coli* response regulator CheY to its target measured in vivo by fluorescence resonance energy transfer. *Proc Natl Acad Sci USA* 99(20):12669–12674.
- Vaknin A, Berg HC (2007) Physical responses of bacterial chemoreceptors. *J Mol Biol* 366(5):1416–1423.
- Danielson MA, Biemann HP, Koshland DE, Jr, Falke JJ (1994) Attractant- and disulfide-induced conformational changes in the ligand binding domain of the chemotaxis aspartate receptor: A ^{19}F NMR study. *Biochemistry* 33(20):6100–6109.
- Govern CC, ten Wolde PR (2014) Energy dissipation and noise correlations in biochemical sensing. *Phys Rev Lett*, in press.
- Sartori P, Tu Y (2011) Noise filtering strategies in adaptive biochemical signaling networks: Application to *E. coli* chemotaxis. *J Stat Phys* 142(6):1206–1217.
- Gao R, Stock AM (2013) Probing kinase and phosphatase activities of two-component systems in vivo with concentration-dependent phosphorylation profiling. *Proc Natl Acad Sci USA* 110(2):672–677.
- Landauer R (1961) Irreversibility and heat generation in the computing process. *IBM J Res Develop* 5(3):183–191.