Energetic costs of cellular computation

Pankaj Mehta[a,1] **and David J. Schwab**[b]

[a]Department of Physics, Boston University, Boston, MA 02215; and [b]Department of Molecular Biology and Lewis-Sigler Institute, Princeton University, Princeton, NJ 08854

Edited by Herbert Levine, University of California, San Diego, La Jolla, CA, and approved September 10, 2012 (received for review May 9, 2012)

Cells often perform computations in order to respond to environmental cues. A simple example is the classic problem, first considered by Berg and Purcell, of determining the concentration of a chemical ligand in the surrounding media. On general theoretical grounds, it is expected that such computations require cells to consume energy. In particular, Landauer's principle states that energy must be consumed in order to erase the memory of past observations. Here, we explicitly calculate the energetic cost of steady-state computation of ligand concentration for a simple two-component cellular network that implements a noisy version of the Berg–Purcell strategy. We show that learning about external concentrations necessitates the breaking of detailed balance and consumption of energy, with greater learning requiring more energy. Our calculations suggest that the energetic costs of cellular computation may be an important constraint on networks designed to function in resource poor environments, such as the spore germination networks of bacteria.

biophysics | signaling | inference | nonequilibrium

The relationship between information and thermodynamics remains an active area of research despite decades of study (1–4). An important implication of the recent experimental confirmation of Landauer's principle, relating the erasure of information to thermodynamic irreversibility, is that any irreversible computing device must necessarily consume energy (2, 3). The generality of Landauer's argument suggests that it is true regardless of how the computation is implemented. A particularly interesting class of examples relevant to systems biology and biophysics is that of intracellular biochemical networks that compute information about the external environment. These biochemical networks are ubiquitous in biology, ranging from the quorum-sensing and chemotaxis networks in single-cell organisms to networks that detect hormones and other signaling factors in higher organisms.

A fundamental issue is the relationship between the information processing capabilities of these biochemical networks and their energetic costs (5–8). It is known that energetic costs place important constraints on the design of physical computing devices as well as on neural computing architectures in the brain and retina (9–11), suggesting that these constraints may also influence the design of cellular computing networks.

The best studied example of a cellular computation is the estimation of the steady-state concentration of a chemical ligand in the surrounding environment (12–14). This problem was first considered in the seminal paper by Berg and Purcell who showed that the information a cell can acquire about its environment is fundamentally limited by stochastic fluctuations in the occupancy of the membrane-bound receptor proteins that detect the ligand (12). In particular, they considered the case of a cellular receptor that binds ligands with a concentration-dependent rate k_4^{off} and unbinds particles at a uniform rate k_4^{off} (see Fig. 1). They argued that cells could estimate the ambient chemical concentration by calculating the average time a receptor is bound during a chosen measurement time $T \gg 1$. Recently, however, it was shown that the optimal strategy for a cell is instead to calculate the average duration of the unbound intervals during T or, equivalently, the total time that the receptor was unbound during T. This later computation implements maximum likelihood estimation (MLE) (14). In these previous studies, the biochemical network downstream of the receptors that implements the desired computations was largely ignored because the authors were primarily interested in calculating fundamental limits on how precisely cells can compute external concentrations. However, calculating energetic costs requires us to explicitly model the downstream biochemical networks that implement these computations (15).

Here, we consider a simple two-component biochemical network that encodes information about ligand concentration in the steady-state concentration of the activated form of a downstream protein (as shown in Fig. 1). Such two-component networks are a common signal transduction motif found in bacteria and are often used to sense external signals through receptor-catalyzed phosphorylation of a downstream response regulator (16). The membrane-bound receptors can be in an either an active "on" state or an inactive "off" state. For simplicity, as in previous works (12–14), we assume that the binding affinity of the on state is extremely high such that all ligand-bound receptors are always in the on state and all unbound receptors are in the off state. Receptors can switch between the off state and on state at a concentration-dependent rate k_4^{off} and from the on state to the off state at a concentration-independent rate k_4^{on}. Receptors additionally convert a downstream signaling protein from an inactive form X to an active form X^*, by, for example, phosphorylation, at a state-dependent rate k_2^s, where $s = \text{on}, \text{off}$. The proteins are deactivated at a state-independent rate k_1. The dependence of k_2^s on the receptor state is what propagates information about ligand concentration downstream. The biochemical network described above contains all the basic elements of a computing device (see Table 1).

Importantly, the deactivation rate of the off state is small yet must be nonzero for thermodynamic consistency (17). We also note that for the case where proteins are activated through phosphorylaltion, k_2^{off} includes nonspecific phosphorylation arising from other kinases as well as contributions from the reverse reactions of the phosphotases. The inactivation rate sets the scale for the effective measurement time $T \propto k_1^{-1}$ because it is the rate at which information encoded in downstream proteins is lost due to inactivation. In order to compute external concentrations accurately, the measurement time must be much longer than the typical switching times between receptor states, $k_1 \ll k_4^{\text{on}}, k_4^{\text{off}}$. We show below that this simple network in fact implements a noisy version of the original Berg–Purcell calculation. Our explicit construction allows us to study the relationship between information and power consumption in this network.

The paper is organized as follows. In the first section, we compute the steady-state behavior of the system, first within the linear-noise approximation and then through an exact solution of the corresponding master equation. In the next two sections, we

Author contributions: P.M. and D.J.S. designed research, performed research, contributed new reagents/analytic tools, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: pankajm@bu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1207814109/-/DCSupplemental.

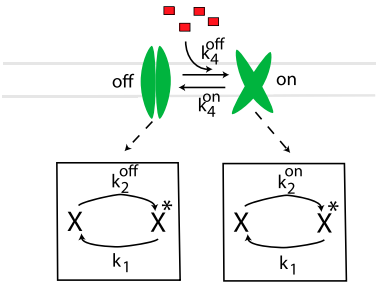

Fig. 1. A cellular network for the computation of an external ligand concentration. External ligands are detected by a receptor that can exist in two conformations: A high-activity on state and a low-activity off state. Receptors switch between states at rate k_{off} and k_{on}. Receptors in state $s = \{on, off\}$ can post-translationally activate (i.e., phosphorylate) a downstream protein at a rate k_2^s. The protein is deactivated (i.e., dephosphorylated) at a constant rate k_1.

quantify the efficacy of chemosensation through the variance in estimated concentration (using the linear-noise approximation) and calculate the power consumption required to maintain the network in steady state (using the exact solution). Finally, we show there exists a tradeoff between these two quantities and discuss implications for the design of cellular sensing systems. We begin each section with a general discussion that summarizes the major ideas, results, and interpretations followed by a more technical and mathematical discussion.

Characterization of the Network's Steady-State Properties

In this section, we derive the steady-state properties of the cellular network considered above. The network translates the external ligand concentration into an internal concentration of activated proteins. This mapping between external ligand concentration and downstream protein number is probabilistic due to the stochasticity inherent in biochemical networks. For this reason, we can characterize the output of the network by a probability distribution of activated proteins, $p(n)$. The shape of the distribution $p(n)$ depends explicitly on the kinetic parameters (k_1, k_2^s, $k_4^{on/off}$) (see Fig. 2). In the following sections, we focus on the fast-switching regime where receptors switch between the on and off states quickly as compared to the deactivation rate, and as we will see, $p(n)$ is unimodal (see Fig 2*B*).

In this regime, two important characteristics of the distribution $p(n)$ are the mean, protein number, $\bar{n}$, and variance $\langle \delta n^2 \rangle$. The mean protein number is the cell's best estimate of the external ligand concentration. On the other hand, the variance characterizes the cell's uncertainty about external ligand concentrations due to stochasticity in the underlying biochemical network. Though it is possible to decrease the variance by increasing the mean protein number, it will be always be nonzero due to fluctuations in the state of the receptors (12).

We begin by first deriving the steady-state mean and variance of the number of activated proteins, n, within the linear-noise approximation. Afterwards, we will study the full probability distribution. The deterministic dynamics of the biochemical network in Fig. 1 is captured by simple rate-equations for the time dependence of the receptor state probabilities and the mean number of activated proteins. We can augment these equations to account for stochastic fluctuations within the linear-noise approximation by adding appropriate Langevin noise terms. In the remainder of this paper,

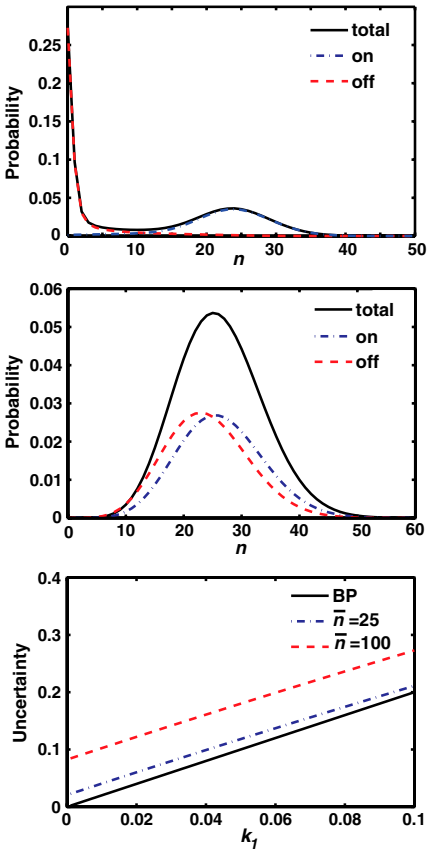

Fig. 2. *Top* Slow switching regime, $k_1 \gg k_4^{on}, k_4^{off}$, with a bimodal distribution of activated proteins. Probability of having n activated proteins at steady-state (black solid line), probability of having n activated proteins when receptor is in the on state (blue dash-dot line), probability of having n activated proteins when receptor is in the off state (red dashed line). *Middle* Fast switching regime, $k_1 \ll k_4^{on}, k_4^{off}$, where the distribution of activated proteins is unimodal. Total probability (black solid line), probability when receptor is in the on state (blue dash-dot line), probability when receptor is in the off state (red dashed line). *Bottom* The uncertainty in ligand concentration, $(\delta c_{rms}/\bar{c})^2$ as a function of k_1 with mean number of active proteins $\bar{n} = 25$ (dashed red line) and $\bar{n} = 100$. This can be compared to the Berg–Purcell result (solid black line). Parameters: $k_2^{off} = 0.01$, k_1^{on}, $k_4^{on} = k_4^{off} = 1$.

we assume for simplicity that proteins are abundant and ignore saturation effects. The dynamics of the circuit is therefore described by a pair of Langevin equations for the probabilities p_{on} and p_{off} (i.e. $1 - p_{on}$) for the receptor to be in the on or off states, respectively, together with the number of activated proteins n,

$$\frac{dp_{on}}{dt} = k_4^{off}(1 - p_{on}) - k_4^{on} p_{on} + \eta_r(t), \qquad [1]$$

$$\frac{dn}{dt} = k_2^{on} p_{on} + k_2^{off}(1 - p_{on}) - k_1 n + \eta_n(t). \qquad [2]$$

The variance of the Langevin terms is given by the Poisson noise in each of the reactions:

Table 1. Summary of the biological realization of basic computational elements

Computational element	Biological realization
Memory	Number of activated (phosphorylated) proteins
Writing to memory	Concentration dependent activation (phosphorylation) of proteins
Erasure of memory	Deactivation (dephosphorylation) of proteins
Energy dissipation	Entropy production due to a lack of detailed balance in chemical kinetic of activation/deactivation

$$\langle \eta_n(t)\eta_n(t')\rangle = [k_2^{\text{on}}\bar{p}_{\text{on}} + k_2^{\text{off}}(1-\bar{p}_{\text{on}}) + k_1\bar{n}]\delta(t-t')$$
$$\langle \eta_r(t)\eta_r(t')\rangle = [k_4^{\text{off}}(1-\bar{p}_{\text{on}}) + k_4^{\text{on}}\bar{p}_{\text{on}}]\delta(t-t'), \quad [3]$$

where $\delta(t-t')$ denotes the Dirac-delta function and the overbar denotes the mean steady-state value of the respective quantity (18, 19).

At steady-state, we can calculate the mean probability and number of proteins by setting the time derivative in Eq. **2** equal to zero while ignoring noise terms, yielding

$$\bar{p}_{\text{on}} = 1 - \bar{p}_{\text{off}} = \frac{K_4^{\text{off}}}{K_4^{\text{off}} + K_4^{\text{on}}}, \quad [4]$$

and

$$\bar{n} = (K_2^{\text{on}} - K_2^{\text{off}})\bar{p}_{\text{on}} + K_2^{\text{off}}, \quad [5]$$

where we have defined the dimensionless parameters $K_j^s = k_j^s/k_1$ with $j = \{2, 4\}$ and $s = \{\text{on}, \text{off}\}$. For the biologically realistic case $K_2^{\text{off}} \ll K_2^{\text{on}} p_{\text{on}}$, the mean number of proteins is simply proportional to the kinase activity in the on state times the probability of being in the on state, $\bar{n} \approx K_2^{\text{on}} p_{\text{on}}$, as expected. One can further calculate the variance in protein numbers (*SI Text*)

$$\langle (\delta n)^2 \rangle = \bar{n} + (\Delta K_2^{\text{on}})^2 \frac{\bar{p}_{\text{on}}\bar{p}_{\text{off}}}{1 + K_4^{\text{on}} + K_4^{\text{off}}}. \quad [6]$$

The first term on the right-hand side of the equation results from Poisson noise in the synthesis and degradation of activated protein, whereas the second term is due to stochastic fluctuations in the state of the receptors.

In addition to the mean and variance, we will need the full steady-state probability distribution for n to calculate the power consumption of the network. The steady-state distribution can be calculated from the master equation for the probability, $p_s(n)$, of there being n active proteins with the receptor in a state s:

$$\frac{dp_s(n)}{dt} = k_1(n+1)p_s(n+1) + k_2^s p_s(n-1) + k_4^s p_s(n) - (k_1 n + k_2^s + k_4^s)p_s(n), \quad [7]$$

where $\bar{s} =$ off (on) when $s =$ on (off). This equation is similar to those found in (20, 21) and the steady-state distribution can be solved via a generating function approach (see *SI Text*).

Depending on the parameters, the steady-state distributions can have two qualitatively distinct behaviors as shown in Fig. 2. In the "slow switching" regime with $k_2^{\text{off}} \ll k_2^{\text{on}}$ and k_4^{on}, $k_4^{\text{off}} \ll k_1$, receptors switch at rates much slower than the protein deactivation rate k_1. The result is a bimodal distribution of activated proteins that can be intuitively understood to arise from the superposition of the probability distributions of activated proteins when the receptor is the on and off states. Similar behavior has been found in the steady-state expression of a self-regulation of a gene (22). As k_2^{on} approaches k_2^{off}, the distributions in the two states merge and the overall probability distribution becomes unimodal. On the other hand, in the "fast switching" regime, characterized by $k_4^{\text{on}}, k_4^{\text{off}} \ll k_1$, the distribution of activated proteins is always unimodal. In this limit, the measurement time, $T \propto k_1^{-1}$, is much longer than the average time a receptor remains in the on or off state, and the biochemical network "time-averages" out the stochastic fluctuations in receptor states. In what follows, we restrict our considerations to this latter regime.

Quantification of Learning

The biochemical circuit in Fig. 1 "computes" the external concentration of a chemical ligand. As emphasized by Berg and Purcell in their seminal paper (12), the chief obstacle in determining external concentration is the stochastic fluctuations in the state of the ligand-binding receptors. Berg and Purcell argued that a reasonable measure of how much cells learn is the uncertainty cells have about external concentration as measured by the variance of the estimated concentration, $(\delta c)^2$. $(\delta c)^2$ measures the uncertainty about the external ligand concentration based on the probability distribution of downstream activated proteins. Using standard arguments, we show below that this uncertainty is directly related to the variance, $(\delta n)^2$, of the corresponding protein probability distribution. This framework allows us to use the results of the last section to quantify how much cells learn about external ligand concentrations as a function of kinetic parameters. The results are plotted in Fig. 2*C*.

To compute uncertainty, Berg and Purcell assumed that the cell computes the average receptor occupancy by time-averaging over a measurement time T. They showed (12) that

$$\frac{(\delta c_{\text{BP}})^2}{c^2} = \frac{2k_4^{\text{on}}}{T\bar{p}_{\text{on}}} = 2/N_b, \quad [8]$$

where $k_4^{\text{off}} = k_+ c$, $k_4^{\text{on}} = k_-$ is independent of c, and N_b is the number of binding events during the time T. It was later shown that cells could compute concentration more accurately by implementing MLE with (14, 23)

$$\frac{(\delta c_{\text{ML}})^2}{c^2} = \frac{1}{2} \times \frac{(\delta c_{\text{BP}})^2}{c^2}. \quad [9]$$

The factor of two decrease in uncertainty derives from the fact that MLE ignores noise due to unbinding of ligands from the cell. We note that this "forgetting" of ligand unbinding, in turn, requires the receptors themselves to be out of equilibrium, further contributing to the system's energy consumption. We do not, however, consider the additional energetic costs of nonequilibrium receptors here.

To quantify learning in our biochemical circuit, we follow Berg and Purcell and estimate the fluctuations in $(\delta c)^2$ as

$$\frac{(\delta c)^2}{c^2} = \left(c\frac{\partial \bar{n}}{\partial c}\right)^{-2} (\delta n)^2, \quad [10]$$

with $(\delta n)^2 = \langle n^2 \rangle - \bar{n}^2$. Substituting $k_4^{\text{off}} = k_+ c$ and $k_4^{\text{on}} = k_-$ and computing the derivative using Eq. **5** gives

$$\left(c\frac{\partial \bar{n}}{\partial c}\right)^2 = (\bar{p}_{\text{on}}\bar{p}_{\text{off}}\Delta K_2)^2. \quad [11]$$

Substituting Eqs. **5** and **6** into Eq. **10** yields

$$\frac{(\delta c)^2}{c^2} = \frac{\bar{n}}{(\bar{p}_{\text{on}}\bar{p}_{\text{off}}\Delta K_2)^2} + \frac{1}{(\bar{p}_{\text{on}}\bar{p}_{\text{off}})(1 + K_4^{\text{on}} + K_4^{\text{off}})}. \quad [12]$$

Similar to the linear-noise calculation, the first term on the right-hand side arises from the Poisson fluctuations in activated protein number, while the second term results from the stochastic fluctuations in the state of receptors. Fig. 2 shows the uncertainty, $(\delta c)^2/c^2$, as a function of the degradation rate of activated protein, k_1, when $\bar{n} = 25$ and $\bar{n} = 100$ and $k_2^{\text{on}} \gg k_2^{\text{off}}$.

By identifying the degradation rate with the inverse measurement time, $k_1 = 2T^{-1}$, we can also compare the results with Berg–Purcell. The factor of two is due to the slight difference in how the variance of the average receptor occupancy is calculated for a biochemical network when compared to the original

Berg–Purcell calculation (23). As shown in Fig. 2, when $\bar{n}$ is increased, the Poisson noise in protein production is suppressed and the performance of the cellular network approaches that due to Berg-Purcell. To make the connection with Berg–Purcell more explicit, it is helpful to rewrite Eq. **12** in terms of the average number of binding events, N_b, during the averaging time, T:

$$\frac{(\delta c)^2}{c^2} = \frac{\bar{n}}{(\bar{n} - K_2^{\text{off}})^2 p_{\text{off}}^2} + \frac{2}{N_b}\left(1 - \frac{k_1}{k_4^{\text{on}} + k_4^{\text{off}} + k_1}\right). \quad [13]$$

This form allows us to see the approach to the Berg–Purcell limit. When the measurement time is much longer than the timescale of fluctuations in receptor activity, i.e., $k_4^{\text{on,off}} \gg k_1$ (or equivalently $K_2^{\text{on}} \gg K_2^{\text{off}} \gg 1$) and the average number of activated proteins is large, $\bar{n} \gg K_2^{\text{off}} \gg 1$, the expression above reduces to $(\delta c)^2/c^2 \approx 2/N_b$ in agreement with Eq. **8**.

Power Consumption and Entropy Production in Steady State

We now compute the energy consumed by the circuit in Fig. 1 as a function of the kinetic parameters. To do so, we exploit the fact that dynamics of the circuit can be formulated as a nonequilbrium Markov process (see Fig. 3). A nonequilibrium steady-state (NESS) necessarily implies the breaking of detailed balance in the underlying Markovian dynamics and, therefore, possesses a nonzero entropy production rate. The entropy production rate is precisely the amount of power consumed by the biochemical circuit in maintaining the nonequilibrium steady state. Thus, by calculating the entropy production rate as function of kinetic parameters, we can calculate the power consumed by the biochemical network implementing the computation. To calculate entropy production, we utilize the full steady-state probability distribution and standard formulas from the theory on nonequilibrium Markov processes.

Consider a general Markov process with states labeled by σ and transition probability from σ to σ′ given by $k(\sigma, \sigma')$. Defining the steady-state probability of being in state σ by P_σ, the entropy production rate, EP, for a NESS is given by (24)

$$EP = \sum_{\sigma,\sigma'} P(\sigma) k(\sigma, \sigma') \log \frac{k(\sigma, \sigma')}{k(\sigma', \sigma)}. \quad [14]$$

For our problem this general formula reduces to

$$EP = k_1 \sum_{s=\{\text{on,off}\}} \sum_n p_s(n) \left(K_2^s \log \frac{K_2^s}{n+1} - n \log \frac{K_2^s}{n} \right), \quad [15]$$

where we again use $K_2^s = k_2^s/k_1$ (see *SI Text*). In deriving this formula, we assumed that the receptors were in thermal equilibrium and obeyed detailed balance. Using explicit expressions for the steady-state distributions, $p_s(n)$, we can calculate the energy consumption of the network as a function of kinetic parameters (see *SI Text*). The physical content of this expression is summarized in Fig. 3. The expression states that any nonzero cyclic flux must necessarily produce entropy. Otherwise, one would have a chemical version of a perpetual motion machine. Figs. 3 and 4 show the power consumption as a function of $\Delta K_2 = K_2^{\text{on}} - K_2^{\text{off}}$ and k_1. Notice that the power consumption tends to zero as both these parameters go to zero. Note that we cannot, however, set $k_1 = 0$ identically because there then no longer exists a steady-state distribution.

Energetics, Information, and Landauer's Principle

We now highlight the fundamental connection between the energy consumed by the network and the information the network acquires about the environment and briefly discuss its relation to Landauer's principle. First, note that learning information about the environment requires energy consumption by the network. This relationship can be seen in Fig. 3, which shows that as $\Delta k_2 \to 0$, the uncertainty about the concentration tends to infinity. Moreover, we show in the *SI Text* that the entropy production, Eq. **15**, is zero if and only if $\Delta k_2 = 0$. In conjunction with Eq. **12**, which diverges as $\Delta k_2 \to 0$, these observations imply that learning requires consuming energy. Mathematically, in the limit where $\Delta k_2 = 0$, the dynamics of the Markov process in Fig. 3 become "one-dimensional" instead of a two-legged ladder, and the dynamics obeys detailed balance. Physically, in this limit the number of downstream proteins becomes insensitive to external ligand concentrations because all information about concentration is contained in the relative probabilities of being in the on or off state.

Second, as shown in Fig. 4, the power consumption of the circuit tends to zero as $k_1 \to 0$. This result is consistent with, and a manifestation of, Landauer's principle: Entropy production stems from erasing memory in a computing device. The number of activated proteins serves the function of a memory of ligand concentration, which is erased at the dephosphorylation rate k_1. Thus, as the erasure rate of the memory tends to zero, the device consumes less energy per unit time, as expected. Yet despite the fact that the power consumption tends to zero as k_1 decreases, the total energy consumed per measurement, namely the power times the measurement time, $T \simeq 2k_1^{-1}$, still increases (see Fig. 4). Thus, learning more requires consuming more total energy despite the fact that power consumption is decreasing. In effect, one is approaching the reversible computing limit where memory is erased adiabatically. Again note, however, that when erasure is

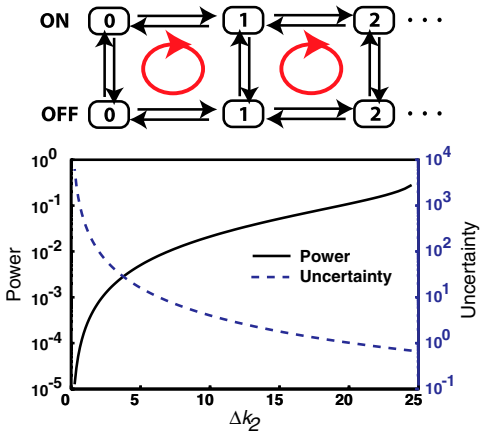

Fig. 3. *Upper* The probabilistic Markov process underlying the circuit in Fig. 1 takes the form of a two-legged ladder. Any nonzero cyclic flux (depicted in red) results in entropy production and power consumption. *Lower* Power consumption (solid black line) and uncertainty (dashed purple line) as a function of $\Delta k_2 = k_2^{\text{on}} - k_2^{\text{off}}$ when $\bar{n} = 25$, and $k_4^{\text{on}} = k_4^{\text{off}} = 1$.

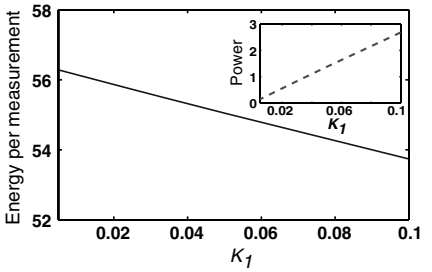

Fig. 4. Total energy per independent measurement (Power × k_1^{-1}) as a function of k_1 when $\bar{n} = 25$. (*Inset*) Power consumption as a function of k_1 over the same parameter range. Note that although the system's power consumption decreases with decreasing k_1, the energy per measurement increases.

performed infinitely slowly, $k_1 = 0$, the system no longer has an NESS and our formalism does not apply.

Finally, we note that one of the important open problems in our understanding of chemosensation is a full characterization of the constraints placed on the measurement time T. In principle, cells can always learn more by measuring the environment for longer periods of time. However, in most biological systems, these measurement times are observed to be quite short. There are a number of constraints that can limit this measurement time, including rotational diffusion (12) (in the case of swimming bacteria) and the restrictions placed on motility. Here, we highlight another restriction that may be important in resource-starved environments: Sensing external concentration necessarily requires cells to consume energy.

Discussion and Conclusion

Cells often perform computations using elaborate biochemical networks that respond to environmental cues. One of the most common simple networks found in bacteria are two-component networks where a receptor phosphorylates a downstream response regulator (16). In this work, we have shown that these simple two-component networks can implement a noisy version of the Berg–Purcell strategy to compute the concentration of external ligands. Furthermore, by mapping the dynamics of the biochemical network to nonequilibrium steady-states in Markov processes, we explicitly derived expressions for the power consumed by the network and showed that learning requires energy consumption. Taken together, these calculations suggest that, much like man-made and neural computing (3, 9–11), energetic considerations may place important constraints on the design of biochemical networks that implement cellular computations. They also suggest a fundamental tradeoff between the efficiency of cellular computing and the requisite energy consumption.

Bacterial cells such as *Bacillus subtilis* can sporulate during times of environmental stress and remain metabolically dormant for many years. Although sporulation is relatively well understood, the reverse process of germination is much more difficult to study. One current model for how a spore knows when to germinate in response to external cues involves integrating the signal and triggering commitment when an accumulation threshold is reached (25, 26). Such a scheme corresponds to the limit of vanishingly small k_1 in our model, so that power consumption is minimized at the expense of retaining the entire integrated signal. Our results indicate that this behavior may be due to the extreme energetic constraints imposed on a metabolically dormant spore, rather than an evolutionarily optimized strategy.

An important insight of this work is that even a simple Berg–Purcell strategy for sensing external concentrations requires the consumption of energy. It is likely that more complicated strategies that increase how much cells learn, such as maximum likelihood, require additional energetic inputs. For example, it was argued in (23) that MLE can be implemented by a network similar to the perfect adaptation network where bursts are produced in response to binding events. These bursts break detailed balance and therefore require energy consumption. It will be interesting to investigate further how the tradeoff between learning and energy consumption manifests itself in the design of computational strategies employed by cells.

In this work, we restricted ourselves to the simple case where cells calculate the steady-state concentration of an external signal. In the future, it will be useful to generalize the analysis performed here to other computations such as responding to temporal ramps (23) and spatial gradients (27, 28). It will also be interesting to understand how to generalize the considerations here to arbitrary biochemical networks. An important restriction on our work is that we reduced our considerations to nonequilibrium steady states. It will be interesting to ask how to generalize the work here to biochemical networks with a strong temporal component.

ACKNOWLEDGMENTS. P.M. and D.J.S. would like to thank the Aspen Center for Physics, where this work was initiated. We are especially grateful to Thierry Mora for clarifying the relationship between the rate k_1 and the average integration time T. This work was partially supported by National Institute of Health Grants K25GM086909 (to P.M.). DS was partially supported by Defense Advanced Research Projects Agency grant HR0011-05-1-0057 and National Science Foundation grant PHY-0957573.

1. Landauer R (1961) Irreversibility and heat generation in the computing process. *IBM J Res Dev* 5:183–191.
2. Berut A, et al. (2012) Experimental verification of Landauers principle linking information and thermodynamics. *Nature* 483:187–189.
3. Bennett C (1982) The thermodynamics of computation: A review. *Int J Theor Phys* 21:905–940.
4. Del Rio L, Berg J, Renner R, Dahlsten O, Vedral V (2011) The thermodynamic meaning of negative entropy. *Nature* 474:61–63.
5. Qian H, Reluga T (2005) Nonequilibrium thermodynamics and nonlinear kinetics in a cellular signaling switch. *Phys Rev Lett* 94:28101.
6. Qian H (2003) Thermodynamic and kinetic analysis of sensitivity amplification in biological signal transduction. *Biophys Chem* 105:585–593.
7. Lan G, Sartori P, Neumann S, Sourjik V, Tu Y (2012) The energy-speed-accuracy tradeoff in sensory adaptation. *Nat Phys* 8:422–428.
8. Tu Y (2008) The nonequilibrium mechanism for ultrasensitivity in a biological switch: Sensing by Maxwell's demons. *Proc Natl Acad Sci USA* 105:11737–11741.
9. Laughlin S (2001) Energy as a constraint on the coding and processing of sensory information. *Curr Opin Neurobiol* 11:475–480.
10. Laughlin S, van Steveninck R, Anderson J (1998) The metabolic cost of neural information. *Nat Neurosci* 1:36–41.
11. Balasubramanian V, Kimber D, II M (2001) Metabolically efficient information processing. *Neural Comput* 13:799–815.
12. Berg H, Purcell E (1977) Physics of chemoreception. *Biophys J* 20:193–219.
13. Bialek W, Setayeshgar S (2005) Physical limits to biochemical signaling. *Proc Natl Acad Sci USA* 102:10040–10045.
14. Endres R, Wingreen N (2009) Maximum likelihood and the single receptor. *Phys Rev Lett* 103:158101.
15. Magnasco M (1997) Chemical kinetics is turing universal. *Phys Rev Lett* 78:1190–1193.
16. Laub M, Goulian M (2007) Specificity in two-component signal transduction pathways. *Annu Rev Genet* 41:121–145.
17. Beard D, Qian H (2008) *Chemical biophysics: Quantitative analysis of cellular systems* (Cambridge Univ Press, Cambridge, MA).
18. Detwiler P, Ramanathan S, Sengupta A, Shraiman B (2000) Engineering aspects of enzymatic signal transduction: Photoreceptors in the retina. *Biophys J* 79:2801–2817.
19. Mehta P, Goyal S, Wingreen N (2008) A quantitative comparison of sRNA-based and protein-based gene regulation. *Mol Syst Biol* 4:221.
20. Iyer-Biswas S, Hayot F, Jayaprakash C (2009) Stochasticity of gene products from transcriptional pulsing. *Phys Rev E* 79:031911.
21. Visco P, Allen R, Evans M (2009) Statistical physics of a model binary genetic switch with linear feedback. *Phys Rev E* 79:031923.
22. Hornos JE, et al. (2005) Self-regulating gene: An exact solution. *Phys Rev E* 72–051907.
23. Mora T, Wingreen N (2010) Limits of sensing temporal concentration changes by single cells. *Phys Rev Lett* 104:248101.
24. Lebowitz J, Spohn H (1999) A gallavotti–cohen-type symmetry in the large deviation functional for stochastic dynamics. *J Stat Phys* 95:333–365.
25. Yi X, Liu J, Faeder J, Setlow P (2011) Synergism between different germinant receptors in the germination of Bacillus subtilis spores. *J Bacteriol* 193:4664–4671.
26. Indest K, Buchholz W, Faeder J, Setlow P (2009) Workshop report: Modeling the molecular mechanism of bacterial spore germination and elucidating reasons for germination heterogeneity. *J Food Sci* 74:R73–R78.
27. Endres R, Wingreen N (2008) Accuracy of direct gradient sensing by single cells. *Proc Natl Acad Sci USA* 105:15749–15754.
28. Hu B, Chen W, Rappel W, Levine H (2010) Physical limits on cellular sensing of spatial gradients. *Phys Rev Lett* 105:48104.

Supporting Information

Mehta and Schwab 10.1073/pnas.1207814109

SI Text

Variance from the Linear-Noise Approximation. Consider the Langevin equations for the ordinary differential equations found in the main text,

$$\frac{dp_{\rm on}}{dt} = k_4^{\rm off}(1-p_{\rm on}) - k_4^{\rm on}p_{\rm on} + \eta_{\rm r}(t)$$
$$\frac{dn}{dt} = k_2^{\rm on}p_{\rm on} + k_2^{\rm off}(1-p_{\rm on}) - k_1 n + \eta_n(t). \quad [S1]$$

We can linearize (with bar denoting average) to get

$$\frac{d\delta p_{\rm on}}{dt} = -(k_4^{\rm off}+k_4^{\rm on})\delta p_{\rm on} - \tau_P^{-1}\delta p_{\rm on} + \eta_{\rm r}(t)$$
$$\frac{d\delta n}{dt} = (k_2^{\rm on}-k_2^{\rm off})\delta p_{\rm on} - k_1\delta n + \eta_n, \quad [S2]$$

where

$$\langle \eta_r(t)\eta_r(t)\rangle = k_4^{\rm off}(1-\bar{p}_{\rm on}) + k_4^{\rm on}\bar{p}_o n = 2(k_4^{\rm off}+k_4^{\rm on})\bar{p}_{\rm on}(1-\bar{p}_{\rm on})$$
$$= 2\tau_P^{-1}\bar{p}_{\rm on}(1-\bar{p}_{\rm on}) \quad [S3]$$

and $\langle \eta_n(t)\eta_n(t)\rangle = 2k_1\bar{n}$. Now we Fourier transform the above equations and use the fact that

$$\langle(\delta n)^2\rangle = \int \frac{d\omega}{2\pi}\langle \delta\hat{n}(\omega)\delta\hat{n}^*(\omega)\rangle \quad [S4]$$

to get

$$\langle(\delta n)^2\rangle = \bar{n} + \frac{(\Delta k_2^{\rm on})^2}{k_1^2}\frac{k_1}{k_1+\tau_P^{-1}}$$
$$= \bar{n} + (\Delta K_2^{\rm on})^2\frac{\bar{p}_{\rm on}\bar{p}_{\rm off}}{1+K_4^{\rm on}+K_4^{\rm off}}, \quad [S5]$$

where we have used $K_i^s = k_i^s/k_1$.

Generating Function for Probability Distribution. The steady-state distribution can be calculated from the master equation for the probability, $p_s(n)$, of there being n active proteins with the receptor in a state s:

$$\frac{dp_s(n)}{dt} = k_1(n+1)p_s(n+1) + k_2^s p_s(n-1) + k_4^{\bar{s}}p_{\bar{s}}(n)$$
$$-(k_1 n + k_2^s + k_4^s)p_s(n), \quad [S6]$$

where $\bar{s}$ = off (on) when s = on (off). At steady state, the left-hand side of Eq. **S6** is equal to zero and

$$K_4^{\bar{s}}p_{\bar{s}}(n) = -(n+1)p_s(n+1) - K_2^s p_s(n-1)$$
$$+(n+K_2^s+K_4^s)p_s(n). \quad [S7]$$

Eq. **S7** is similar to those equation is similar to those found previously 1, 2 and can be solved via a generating function approach. First, we define a pair of generating functions:

$$G_s(n) = \sum_{n=0}^{\infty} p_s(n)z^n, \quad [S8]$$

with s = on, off. We can then rewrite Eq. **S7** in terms of the generating functions as

$$[(z-1)\partial_z - K_2^s(z-1) + K_4^s]G_s(z) = K_4^{\bar{s}}G_{\bar{s}}(z). \quad [S9]$$

This equation must be supplemented by initial conditions for the functions, $G_s(z)$, which follow from the observation that $G_{\rm on}(1) = \bar{p}_{\rm on}$ and $G_{\rm off}(1) = \bar{p}_{\rm off} = 1-\bar{p}_{\rm on}$, where $\bar{p}_{\rm on}$ is given by Eq. **S4** of the main text.

Adding the equations for s = on, off and dividing through by $(z-1)$ gives

$$(\partial_z - K_2^{\rm on})G_{\rm on}(z) = -(\partial_z - K_2^{\rm off})G_{\rm off}(z). \quad [S10]$$

To proceed further, it is useful to define the quantity $H_s(z)$ related to the generating functions $G_s(z)$ by

$$G_s(z) = e^{K_2^s z}H_s(z). \quad [S11]$$

It is clear that

$$(\partial_z - K_2^s)G_s(z) = e^{K_2^s z}\partial_z H_s(z). \quad [S12]$$

Thus, we can rewrite Eq. **S10** as

$$e^{K_2^{\rm on}z}\partial_z H_{\rm on}(z) = -e^{K_2^{\rm off}z}\partial_z H_{\rm off}(z). \quad [S13]$$

Similarly, we can rewrite Eq. **S9** as

$$(z-1)e^{K_2^s z}\partial_z H_s(z) + K_4^s e^{K_2^s z}H_s(z) = K_4^{\bar{s}}e^{K_2^{\bar{s}}z}H_{\bar{s}}(z). \quad [S14]$$

Multiplying the equation by $e^{-K_2^{\bar{s}}z}$, taking the derivative with respect to z, substituting Eq. **S13**, and defining $\Delta K_2^s = K_2^{\bar{s}} - K_2^s s$ one has

$$\partial_z H_s(z) - \Delta K_2^s(z-1)\partial_z H_s(z) + (z-1)\partial_z^2 H_s(z)$$
$$-K_4^s\Delta K_2^s H_s(z) + K_4^s\partial_z H_s(z) = -K_4^{\bar{s}}\partial_z H_s(z). \quad [S15]$$

Regrouping terms one has

$$(z-1)\partial_z^2 H_s(z) + (1-\Delta K_2^s(z-1) + K_4^s$$
$$+K_4^{\bar{s}})\partial_z H_s(z) - K_4^s\Delta K_2^s H_s(z) = 0. \quad [S16]$$

Defining

$$u = \Delta K_2^s(z-1), \quad [S17]$$

one can rewrite Eq. **S16** in terms of u as

$$u\partial_u^2 H_s(u) + (1+K_4^s+K_4^{\bar{s}}-u)\partial_u H_s(z) - K_4^s H_s(u) = 0. \quad [S18]$$

Eq. **S18** is just the confluent hypergeometric equation. We can immediately write the solutions in terms of confluent hypergeometric functions of the first and second kind. In particular, the

general solution to this equation is given by (assuming that we need a power series in u with integer cofficients) the confluent geometric function:

$$H_s(u) = c_{s1}F_1(K_4^s; 1 + K_4^s + K_4^{\bar{s}}; u), \quad \text{[S19]}$$

with c_s a constant of integration. Thus, we have

$$G_s(z) = c_s e^{K_2^s z}{}_1F_1[K_4^s; 1 + K_4^s + K_4^{\bar{s}}; \Delta K_2^s(z-1)]. \quad \text{[S20]}$$

To determine the constants, notice that

$$G_s(1) = c_s e^{K_2^s} = \bar{p}_s = \frac{K_2^{\bar{s}}}{K_2^s + K_2^{\bar{s}}}, \quad \text{[S21]}$$

so that

$$c_s = \frac{e^{-K_2^s}K_2^{\bar{s}}}{K_2^s + K_2^{\bar{s}}}. \quad \text{[S22]}$$

Combining Eqs. **S20** and **S22** gives the final expression

$$G_s(z) = \frac{K_4^{\bar{s}}e^{K_2^s(z-1)}}{K_4^s + K_4^{\bar{s}}}{}_1F_1[K_4^s; 1 + K_4^s + K_4^{\bar{s}}; \Delta K_2^s(z-1)]. \quad \text{[S23]}$$

Variance from Generating Functions. We now calculate the variance $(\delta n)^2$ directly from the generating function and see that it matches the linear noise calculation. To do so, write

$$(\delta n)^2 = \sum_s (\partial_z z \partial_z G_s(z))|_{z=1} - \bar{n}^2 = \sum_s (z\partial_z^2 G_s(z))|_{z=1} + \bar{n} - \bar{n}^2. \quad \text{[S24]}$$

First, note that

$$\partial_z G_s(z) = \frac{K_4^s}{K_4^s + K_4^{\bar{s}}}\left[K_2^s e_1^{K_2^s(z-1)}F_1(\cdot,\cdot;\Delta K_2^s(z-1)) + e^{K_2^s(z-1)}\frac{(K_2^{\bar{s}} - K_2^s)K_4^s}{1 + K_4^s + K_4^{\bar{s}}}{}_1F_1(+,+;\Delta K_2^s(z-1))\right], \quad \text{[S25]}$$

where we have used

$$z\partial_{z\,1}F_1(a,b;z) = z\frac{b}{a}{}_1F_1(a+,b+;z). \quad \text{[S26]}$$

Taking the second derivative and setting $(z = 1)$ yields

$$(z\partial_z^2 G_s(z))|_{z=1} = \frac{K_4^s}{K_4^s + K_4^{\bar{s}}}\left[(K_2^s)^2 + \frac{\Delta K_2^s K_4^s}{1 + K_4^s + K_4^{\bar{s}}}\left(2K_2^s + \frac{\Delta K_2^s}{2 + K_4^s + K_4^{\bar{s}}}\right)\right]. \quad \text{[S27]}$$

After some algebra, one has

$$\sum_s (z\partial_z^2 G_s(z))|_{z=1} = p_{\text{on}}(K_2^{\text{on}})^2 + p_{\text{off}}(K_2^{\text{off}})^2 - p_{\text{on}}p_{\text{off}}(\Delta K_2^s)^2\frac{K_4^{\text{on}} + K_4^{\text{off}}}{1 + K_4^{\text{on}} + K_4^{\text{off}}}. \quad \text{[S28]}$$

Eq. **S24** then becomes

$$(\delta n)^2 = \bar{n} + p_{\text{on}}(K_2^{\text{on}})^2 + p_{\text{off}}(K_2^{\text{off}})^2 - p_{\text{on}}p_{\text{off}}(\Delta K_2^s)^2 \times \frac{K_4^{\text{on}} + K_4^{\text{off}}}{1 + K_4^{\text{on}} + K_4^{\text{off}}} - \bar{n}^2, \quad \text{[S29]}$$

$$= \bar{n} + p_{\text{on}}(K_2^{\text{on}})^2 + p_{\text{off}}(K_2^{\text{off}})^2 - p_{\text{on}}p_{\text{off}}(\Delta K_2^s)^2 \times \frac{K_4^{\text{on}} + K_4^{\text{off}}}{1 + K_4^{\text{on}} + K_4^{\text{off}}} - (p_{\text{on}}K_2^{\text{on}} + p_{\text{off}}K_2^{\text{off}})^2, \quad \text{[S30]}$$

$$= \bar{n} + p_{\text{on}}p_{\text{off}}\frac{(K_2^{\text{on}} - K_2^{\text{off}})^2}{1 + K_4^{\text{on}} + K_4^{\text{off}}}, \quad \text{[S31]}$$

$$= \bar{n} + p_{\text{on}}p_{\text{off}}\frac{(\Delta K_2^{\text{on}})^2}{1 + K_4^{\text{on}} + K_4^{\text{off}}}. \quad \text{[S32]}$$

This result is identical to [**S5**] calculated using the linear noise approximation.

Entropy Production. Consider a general Markov process with states labeled by σ and transition probability from σ to σ' given by $k(\sigma, \sigma')$. Defining the steady-state probability of being in state σ by P_σ, the entropy production rate for a Markov process is given by 3

$$\frac{dS}{dt} = \sum_{\sigma,\sigma'} P(\sigma)k(\sigma,\sigma')\log\frac{k(\sigma,\sigma')}{k(\sigma',\sigma)} - \sum_{\sigma,\sigma'} P(\sigma')k(\sigma,\sigma')\log\frac{k(\sigma,\sigma')}{k(\sigma',\sigma)}. \quad \text{[S33]}$$

The first term of the right-hand side is the entropy production whereas the second term is the entropy expelled from the system. As steady state, $\frac{dS}{dt} = 0$ and the entropy production rate, EP, is equal to the energy expelled and is given by

$$EP = \sum_{\sigma,\sigma'} P(\sigma)k(\sigma,\sigma')\log\frac{k(\sigma,\sigma')}{k(\sigma',\sigma)}. \quad \text{[S34]}$$

For the biochemical network described under consideration, the entropy production becomes

$$EP = \sum_{s=\text{on,off},n} p_s(n)\left[k_2^s\log\frac{k_2^s}{k_1(n+1)} + k_1 n\log\frac{k_1 n}{k_2^s} + k_4^s\log\frac{k_4^s}{k_4^{\bar{s}}}\right]. \quad \text{[S35]}$$

Because the receptors are in thermodynamic equilibrium, from detailed balance we know that

$$\sum_{s,n} p_s(n)k_4^s\log\frac{k_4^s}{k_4^{\bar{s}}} = 0, \quad \text{[S36]}$$

so that

$$EP = k_1\sum_{s=\text{on,off}}\sum_n p_s(n)\left(K_2^s\log\frac{K_2^s}{n+1} - n\log\frac{K_2^s}{n}\right), \quad \text{[S37]}$$

where $K_2^s = k_2^s/k_1$. The calculation is completed by substitution of the steady-state distributions, $p_s(n)$, which can be found from Eq. **S23**.

Energy Consumption is Required for Signaling. We show is that $EP = 0$ if and only if $\Delta K_2^s = 0$. To do so we will use the expressions for the probability distribution, Eq. **S23**, and average entropy production, Eq. **S37**. Let us start by writing Eq. **S37** as

$$k_1^{-1} EP = -(\langle n\rangle_{\text{on}} \log K_2^{\text{on}} + \langle n\rangle_{\text{off}} \log K_2^{\text{off}}) + \sum_n (K_2^{\text{on}} p_{\text{on}}(n) \log K_2^{\text{on}} + K_2^{\text{off}} p_{\text{off}}(n) \log K_2^{\text{off}}) + \sum_n ([p_{\text{on}}(n+1) - K_2^{\text{on}} p_{\text{on}}(n)] \log(n+1)) + \sum_n ([p_{\text{off}}(n+1) - K_2^{\text{off}} p_{\text{off}}(n)] \log(n+1)), \quad \textbf{[S38]}$$

where $\langle n\rangle_s = \sum_n n p_s(n)$ for $s = \text{on, off}$. We will now examine each of these terms one by one. Notice that

$$\langle n\rangle_s = \partial_z G_s(z=1) = \frac{K_2^{\bar{s}} K_4^{\bar{s}}}{K_4^s + K_4^{\bar{s}}}, \quad \textbf{[S39]}$$

so we have that

$$(\langle n\rangle_{\text{on}} \log K_2^{\text{on}} + \langle n\rangle_{\text{off}} \log K_2^{\text{off}}) = \sum_s \frac{K_2^{\bar{s}} K_4^{\bar{s}}}{K_4^s + K_4^{\bar{s}}} \log K_2^s. \quad \textbf{[S40]}$$

Notice that because $\sum p_s(n) = G_s(1)$, the second term in the sum above is just

$$\sum_n (K_2^{\text{on}} p_{\text{on}}(n) \log K_2^{\text{on}} + K_2^{\text{off}} p_{\text{off}}(n) \log K_2^{\text{off}}) = \sum_s \frac{K_2^s K_4^{\bar{s}}}{K_4^s + K_4^{\bar{s}}} \log K_2^s. \quad \textbf{[S41]}$$

Substituting these expressions into Eq. **S38** yields

$$\frac{A_p}{k_1} = -\sum_s \frac{K_4^s \Delta K_2^s}{K_4^s + K_4^{\bar{s}}} \log K_2^s + \sum_n ([p_{\text{on}}(n+1) - K_2^{\text{on}} p_{\text{on}}(n)] \log(n+1)) + \sum_n ([p_{\text{off}}(n+1) - K_2^{\text{off}} p_{\text{off}}(n)] \log(n+1)). \quad \textbf{[S42]}$$

In order for Eq. **S42** to equal zero, we know that

$$p_s(n+1) = \frac{K_2^s}{n+1} p_s(n), \quad \textbf{[S43]}$$

because each of the terms in the sum above must equal zero individually. Therefore, the generating functional in detailed balance, $G_s^{DB}(z)$, must be given by

$$G_s^{DB} \propto e^{K_2^s z}. \quad \textbf{[S44]}$$

Comparing with **[S23]** implies that $\Delta K_2^s = 0$. In addition, the first term of [38] also disappears in this case. Thus, we have detailed balance if and only if $\Delta K_2^s = 0$.

1. Iyer-Biswas S, Hayot F, Jayaprakash C (2009) Stochasticity of gene products from transcriptional pulsing. *Phys Rev E* 79:031911.
2. Visco P, Allen R, Evans M (2009) Statistical physics of a model binary genetic switch with linear feedback. *Phys Rev E* 79:031923.
3. Lebowitz J, Spohn H (1999) A Gallavotti–Cohen-type symmetry in the large deviation functional for stochastic dynamics. *J Stat Phys* 95:333–365.