# Kinetic versus Energetic Discrimination in Biological Copying

Pablo Sartori[1] and Simone Pigolotti[2]

[1]Max Planck Institute for the Physics of Complex Systems, Noethnitzer Strasse 38, 01187 Dresden, Germany
[2]Departament de Física i Enginyeria Nuclear, Universitat Politecnica de Catalunya Edificio GAIA, Rambla Sant Nebridi s/n, 08222 Terrassa, Barcelona, Spain

(Received 2 October 2012; published 1 May 2013)

We study stochastic copying schemes in which discrimination between a right and a wrong match is achieved via different kinetic barriers or different binding energies of the two matches. We demonstrate that, in single-step reactions, the two discrimination mechanisms are strictly alternative and cannot be mixed to further reduce the error fraction. Close to the lowest error limit, kinetic discrimination results in a diverging copying velocity and dissipation per copied bit. On the other hand, energetic discrimination reaches its lowest error limit in an adiabatic regime where dissipation and velocity vanish. By analyzing experimentally measured kinetic rates of two DNA polymerases, T7 and Pol$\gamma$, we argue that one of them operates in the kinetic and the other in the energetic regime. Finally, we show how the two mechanisms can be combined in copying schemes implementing error correction through a proofreading pathway.

DOI: 10.1103/PhysRevLett.110.188101 PACS numbers: 87.10.Vg, 05.70.Ln, 87.18.Tt

Living organisms need to process signals in a fast and reliable way. Copying information is a task of particular relevance, as it is required for the replication of the genetic code, the transcription of DNA into mRNA, and its translation into a protein. Reliability is fundamental, since errors can result in the costly (or harmful) production of a nonfunctional protein. Indeed, cells have developed mechanisms to reduce the copying error rate $\eta$ to values as low as $\eta \sim 10^{-4}$ for protein transcription-translation [1] and $\eta \sim 10^{-10}$ for DNA replication [2]. Such mechanisms include multiple discrimination steps [1,2] and pathways to undo wrong copies as in proofreading [3–6] or backtracking [7].

Biological information is copied by thermodynamic machines that operate at a finite temperature. There is agreement that this fact alone implies a lower limit on the error rate. However, contrasting results have been obtained regarding the nature of this limit. In particular, it is not clear when it is reached in a slow and quasiadiabiatic regime, or in a fast and dissipative one. As clarified by Bennett [8], information can be copied adiabatically. Indeed, the copying scheme proposed in Hopfield's seminal proofreading paper [3] reaches its minimum error at zero velocity and zero dissipation [9]. In contrast, a copolymerization model proposed a few years later by Bennett [10–12] achieves its minimum error in a highly dissipative regime, where velocity and dissipation diverge. Some of the biological literature has favored that the minimum error is achieved in near-equilibrium conditions [9]. This view is however not unanimous [13]. Recent biophysical literature supports a dissipative minimum error limit [11,12,14,15]. Similar disagreements are also present in models including proofreading. The proofreading model in Ref. [10] dissipates systematically less than the corresponding copying, while in other models [3,4], at low errors, dissipation comes mainly from the proofreading step.

In this Letter, we show how these contrasting results can be rationalized, noting that a copy can be performed either discriminating through binding energies adiabatically, energetic discrimination, or discriminating through binding barriers dissipatively, kinetic discrimination. We begin by presenting a model for copying a single bit of information in the spirit of those proposed in Refs. [8,16–18]; see Fig. 1(a). A biomachine such as a polymerase binds and unbinds monomers of different species to a template, trying to match it. We then move to the case of copolymerization, Fig. 1(b), where a polymerase assembles a polymer chain to match a template strand. Finally, we discuss two proofreading schemes, Figs. 1(c) and 1(d), where the polymerase is assisted by an exonuclease that tends to remove wrong matches.

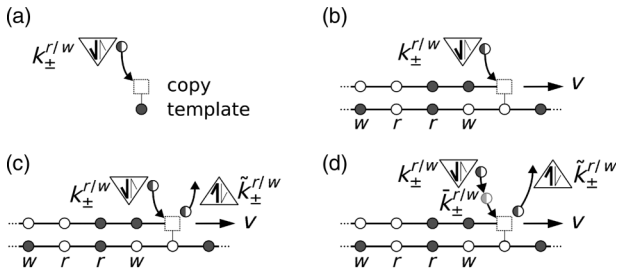


FIG. 1. (a) Copying of one bit. The lower-vertex triangle represents a biomachine, such as a polymerase, binding or unbinding right and wrong matches with rates $k_{\pm}^{r/w}$. (b) Copolymerization. A template strand (bottom) is copied into a new strand (top). Right ($r$) and wrong ($w$) matching monomers are added and removed ($\pm$) with rates $k_{\pm}^{r/w}$ resulting in a growth velocity $v$. (c) and (d) Proofreading schemes. The polymerase is assisted by an exonuclease which removes wrong matches, represented by an upper-vertex triangle and characterized by $\tilde{k}_{\pm}^{r/w}$. In (d), copies are made via an intermediate state characterized by $\bar{k}_{\pm}^{r/w}$.

 

*Stochastic copying strategies of a single bit.*—The copying machine is described as a three-states system. Two are bound states in which the right ($r$) or wrong ($w$) molecule is attached to the machine. The third is a "blank" state ($\varnothing$), representing the unbound state of the machine before a match is done. To help physical intuition and following Ref. [10], we define the rates from the free energy landscape in Fig. 2(a). Right and wrong matching are characterized by a difference in barrier height $\delta$ and in the energy of the final states $\gamma$. The energy $\epsilon$ is a chemical driving. All energies are in units of $k_BT$, where $k_B$ is the Boltzmann constant and $T$ the temperature.

The four rates $k_{\pm}^{r/w}$ connecting the unbound state with the right and wrong states can be written as Kramers rates from energy barriers of Fig. 2(a) as

$$k_+^r = \omega e^{\epsilon+\delta}; \quad k_-^r = \omega e^{\delta}; \quad k_+^w = \omega e^{\epsilon}; \quad k_-^w = \omega e^{\gamma}, \tag{1}$$

where $\omega$ is an overall rate scale. The master equation for the probabilities $p_r$ and $p_w$ of finding the system in the right or wrong state reads

$$\begin{aligned} \dot{p}_r &= (1 - p_r - p_w)k_+^r - k_-^r p_r, \\ \dot{p}_w &= (1 - p_r - p_w)k_+^w - k_-^w p_w, \end{aligned} \tag{2}$$

where $p_\varnothing$ has been eliminated by normalization. We study the time-dependent error rate $\eta(t) = p_w(t)/[p_r(t) + p_w(t)]$ for the system prepared in the unbound state, $p_r(t=0) = p_w(t=0) = 0$. At short times, $t \ll \omega^{-1}$, one has $p_r \approx tk_+^r$ and $p_w \approx tk_+^w$. To shorten notation, we define the function $f(x) = e^{-x}/(1 + e^{-x})$ mapping energies into errors. The short-time error is then $\eta(t \to 0) = f(\delta)$. In the opposite limit of $t \gg \omega^{-1}$, the system reaches equilibrium so that $\eta(t \to \infty) = f(\gamma)$ by detailed balance. At intermediate times, one can demonstrate from the analytical solution of Eqs. (2) that $\eta(t)$ is a monotonic function for any choice of rates (see Supplemental Material [19]): increasing with time when $\delta > \gamma$ [i.e., $f(\delta) < f(\gamma)$] and decreasing when $\gamma > \delta$ [i.e., $f(\delta) > f(\gamma)$]. For $\delta = \gamma$, the error is time independent. The three cases are shown in Fig. 2(b).

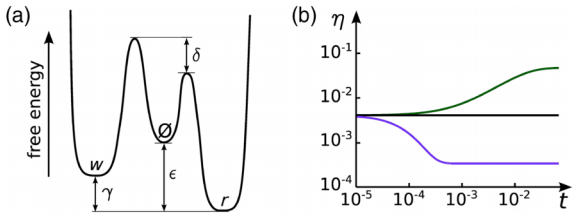


FIG. 2 (color online). (a) Energy diagram for copying rates. The barrier height difference $\delta$ biases right additions, while the energy difference $\gamma$ favors wrong removals. The right and wrong chemical drivings are $\epsilon$ and $\epsilon - \gamma$. (b) Time evolution of the error in the single bit copy for the following three parameter choices: $\gamma < \delta$ (green curve, $\gamma = 3$ and $\delta = 5.5$), $\gamma > \delta$ (blue curve, $\gamma = 8$ and $\delta = 5.5$), and $\gamma = \delta = 5.5$ (black curve). The other parameters are $\epsilon = 5$ and $\omega = 4$.

To maximize accuracy, the copying reaction must be arrested when $\eta(t)$ is at its minimum value, quenching the system into either a right or wrong copy outcome. In an enzymatic reaction, this corresponds to the irreversible transformation of bound states into products [3]. In Ref. [8], where bits are encoded in ferromagnets, it corresponds to decoupling from an external transverse field. We define the kinetic discrimination regime $\delta > \gamma$, where optimal accuracy requires stopping the process as fast as possible. If $\gamma > \delta$, the energetic discrimination regime, optimal accuracy is approached at very long times, when the reaction equilibrates. In all cases, accuracy cannot be improved by combining the two mechanisms, as the lower limit on the error is determined by either $\gamma$ or $\delta$. Notice that in an energetic discrimination scheme, the quench can be performed slowly, at no dissipation [8]. In a kinetic scheme, the quench has to be fast and dissipative.

*Kinetic and energetic discrimination in copolymerization.*—In copolymerization, a polymerase stochastically adds and removes monomers to a tip of the growing copy strand, trying to match them with those on the template strand (see Fig. 1(b) and Refs. [8,11,12,19]). The model is defined by the incorporation and removal rates of right $k_{\pm}^r$ and wrong $k_{\pm}^w$ matching monomers, defined by Eqs. (1) and Fig. 2(a). The chemical drivings of the polymerase for right and wrong bases are $\epsilon$ and $\epsilon - \gamma$. These bias monomer addition over removal and ensure growth of the copied strand at an average velocity $v \geq 0$. Monomer addition or removal and polymerase forward or back stepping are thus tightly coupled (relaxing this has no effect on our results [19]). Previous studies on copolymerization assumed isoenergetic strands, i.e., $\gamma = 0$ [10–12,14]. We relax this assumption and study how the copying velocity $v$, and the rate of entropy production or dissipated chemical work $\dot{S}$ [20], depend on the error rate $\eta$ for a general choice of $\delta$ and $\gamma$. It is straightforward to show that $v = k_+^r - (1 - \eta)k_-^r + k_+^w - \eta k_-^w$ [10,12] and also that $\dot{S}$ is given by

$$\dot{S} = v\Delta S = v(1 - \eta)\epsilon + v\eta(\epsilon - \gamma) + vH(\eta), \tag{3}$$

where $\Delta S = \dot{S}/v$ is the dissipation per added monomer, and $H(\eta) = -\eta\log(\eta) - (1-\eta)\log(1-\eta)$ is the Shannon entropy of the error rate $\eta$. The first two terms in Eq. (3) represent the distinct chemical driving forces of right and wrong bases, multiplied by the flux of right and wrong incorporated bases. The last term of Eq. (3) corresponds to the information entropy increase due to incorporation of errors, hence information, into the chain [10–12].

By imposing steady state flux conservation, we express $\epsilon$ in terms of $(\eta, \delta, \gamma)$. Substituting, we obtain $\Delta S(\eta, \delta, \gamma)$ and $v(\eta, \delta, \gamma)$ [19], presented in Figs. 3(a) and 3(c) for a fixed value of $\gamma$ and different values of $\delta$. The physical range of admissible errors depends on $\gamma$ and $\delta$ (see Ref. [19]) as

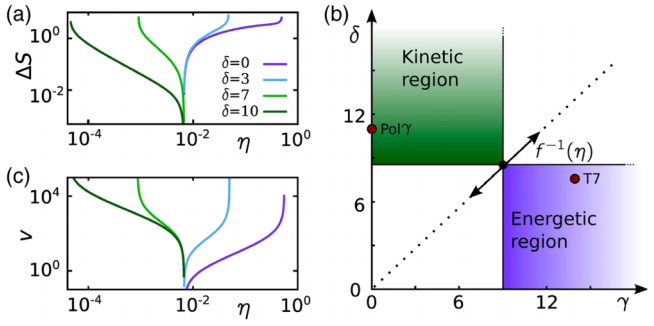


FIG. 3 (color online). (a) Dissipation per step in copolymerization. In all curves $\gamma = 5$, while $\delta$ varies. All curves tend to zero at $\eta = f(\gamma) \approx 6.7 \times 10^{-3}$. Blue curves are in the energetic region, $\gamma > \delta$, while green curves are in the kinetic region $\delta > \gamma$. (b) $\gamma$-$\delta$ phase diagram, showing the kinetic and energetic discrimination regions compatible with an error $\eta \sim f(7) \approx 9 \times 10^{-4}$, and estimated values of $(\gamma, \delta)$ for Pol$\gamma$ and T7. Tuning $\eta$ shifts the limit $f^{-1}(\eta)$ of phase regions along the line $\gamma = \delta$. (c) Behavior of the velocity $v$ for the parameter choices in (a).

$$\min[f(\delta), f(\gamma)] < \eta < \max[f(\delta), f(\gamma)], \tag{4}$$

with $f(x)$ as previously defined. We now study the dissipative limit, $\eta \to f(\delta)$, and the adiabatic limit, $\eta \to f(\gamma)$.

When $\eta \to f(\delta)$, the chemical driving diverges as $\epsilon \sim -\log|\eta - f(\delta)|$ [19]. Substituting into Eq. (3) shows that also $\Delta S$ diverges [see Fig. 3(a)] as

$$\Delta S \sim \eta\epsilon \sim \eta \log|\eta - f(\delta)| \quad \text{for } \eta \to f(\delta). \tag{5}$$

Since $\epsilon \gg 1$, the information entropy $H(\eta)$ in Eq. (3) is negligible, and dissipation is dominated by the chemical terms. As an effect of the strong driving, the velocity diverges as $|\eta - f(\delta)|^{-1}$; see Fig. 3(c).

When $\eta \to f(\gamma)$, both $v$ and $\Delta S$ tend to zero [see Figs. 3(a) and 3(c)]: all the chemical energy is invested in copying the information without any being wasted. The chemical driving is then

$$\epsilon = \log(1-\eta) = \log[1 - f(\gamma)] < 0 \quad \text{for } \eta = f(\gamma). \tag{6}$$

Note that $\epsilon$ is small and negative to compensate for the small positive entropic driving caused by $H(\eta)$ in Eq. (3).

By inverting Eq. (4), the values of $\gamma$ and $\delta$ compatible with a given error $\eta$ must satisfy either $\gamma < f^{-1}(\eta) < \delta$ or $\delta < f^{-1}(\eta) < \gamma$, with $f^{-1}(x) = \log(1 + 1/x)$ the inverse of $f(x)$. This defines the two disconnected kinetic discrimination $(\delta > \gamma)$ and energetic discrimination $(\gamma > \delta)$ regions of the $(\gamma, \delta)$ plane in Fig. 3(b).

In the kinetic region, both $\Delta S$ and $v$ diverge in the minimum error limit, so that accuracy comes at the cost of high dissipation. In the energetic region, accurate copying comes at the cost of the copying velocity, which goes to zero in the adiabatic minimum error limit. This fundamental difference is at the core of the discrepancies between enzymatic copying models [3] that assumed lack of forward discrimination, $\delta = 0$ in our language (see Ref. [19] for mapping), and copolymerization studies [10–12] that assumed isoenergetic strands, $\gamma = 0$. Our results show that it is impossible to interpolate between the two, as they belong to two separate regions of parameter space.

*Operating regimes of T7 and* Pol$\gamma$ *polymerases.*—We now analyze two specific biological copying systems: DNA replication of the phage T7 [2,21] and replication of human DNA by Pol$\gamma$ [22]. A recent experimental study [22] points at the strong and asymmetric backward rates as the leading discriminatory mechanism in T7. We derived from [22] the copolymerization rates by assuming equilibrium nucleotide binding with dissociation constants $K_r = 28\ \mu$M and $K_w = 200\ \mu$M for right and wrong base matching. Considering nucleotide concentrations in a range of $[dNTP] \sim 0.5$–$50\ \mu$M, we obtain the binding states $1/(1 + K_{r/w}/[dNTP])$. Multiplying them by the forward rates (360 and 0.2 Hz for right and wrong bases, respectively), we obtain $k_+^{r/w}$. The backward rates are $k_-^r \approx 2$ Hz and $k_-^w \approx 0.04$ Hz [22]. These values give an error range $\eta \sim 10^{-6}$–$10^{-4}$, in agreement with Ref. [2]. Usual estimates of the error assume linear binding, approximation valid for low $[dNTP]$ and yielding the lowest end of the error range. The velocities are $v \sim 5$–$250$ bases per second, in agreement with the saturation rate measured in Ref. [22]. By inverting Eqs. (1), we can infer $\gamma \approx 14$ and $\delta \approx 8$. Since $\gamma > \delta$, we conclude that T7 operates in the energetic regime [see Fig. 3(b)].

DNA duplication by Pol$\gamma$ was analyzed in Ref. [11] with a variant of the copolymerization model, where different monomer species are characterised by different rates. Agreement with experimental data in Ref. [21] was obtained assuming that the copy be isoenergetic ($\gamma = 0$). We simplify the analysis in Ref. [11] by averaging over the different monomer species. Using the same driving $\epsilon \approx 5$ determined for T7, we obtain $\delta \approx 11$ and a range of error rates $\eta \sim 10^{-5}$–$10^{-3}$. In the limit of low $[dNTP]$, it agrees with the estimates in Refs. [11,21]. As $\gamma = 0$, Pol$\gamma$ lies in the kinetic discrimination region [Fig. 3(b)]. While here as in Ref. [11], $\gamma = 0$ was assumed for simplicity, a nonzero value of $\gamma$ but smaller than $\delta$ would not alter our main conclusion.

The estimates of $\delta$ and $\gamma$ above indicate that while the two polymerases achieve a similar error rate $\eta$, they operate in different regimes, implying different tradeoffs. In T7, lowering $[dNTP]$ (effectively, the chemical driving) can reduce the error $\eta$. This also reduces the dissipation $\Delta S$, at the cost of a smaller speed $v$. This situation is similar to that of the blue curves in Figs. 3(a) and 3(c). In Pol$\gamma$, a smaller error requires a stronger driving, hence dissipation [11]. This gives a higher polymerization rate, as in the green curves of Figs. 3(a) and 3(c).

*Combining copying strategies in proofreading schemes.*—We now explore the possibility of combining the two mechanisms in multistep copying schemes, involving a proofreading pathway. In proofreading, an initially

copied base can be removed via an alternative pathway; see Figs. 1(c) and 1(d). Such an erasing pathway is characterized by a discrimination which, *a priori*, can be energetic $\gamma_p$ or kinetic $\delta_p$, a distinct time scale $1/\omega_p$, and a (backward) driving $\epsilon_p$. In an effective proofreading scheme, the minimal copying error of Eq. (4) is reduced by an additional proofreading factor, in principle energetic $f(\gamma_p)$ or kinetic $f(\delta_p)$. We discuss two proofreading schemes. In both of them, the proofreading rates $\tilde{k}_{\pm}^{r/w}$ have the same structure as the copying ones, apart from a backward driving [19]. In the first, Fig. 1(c), the copying step is identical to that in the copolymerization model, as in Bennett's proofreading model [10]. In the second, Fig. 1(d), the copying step leads to an intermediate state, taken to its final form via rates $\bar{k}_{\pm}^{r/w}$ without further discrimination, as in Hopfield's model [3]. By imposing flux balance at the steady state, we solved both models analytically [19]. We fixed the discrimination factors, and for each error $\eta$ minimized $\Delta S$ over the remaining free parameters [19], obtaining the curves of minimum dissipation versus error in Fig. 4.

As shown in Ref. [19], there are no regimes in any of the two proofreading schemes where the error is lowered by the energetic factor $f(\gamma_p)$, while error reduction by a kinetic proofreading factor $f(\delta_p)$ is feasible; see Figs. 4(a) and 4(b). Proofreading is thus only effective when it operates in the kinetic regime. This result is consistent with Landauer's principle [23], as erasure of information (errors) constitutes an intrinsically dissipative process. Further, by looking at the minimum errors in Fig. 4, one can conclude that, while kinetic proofreading is always effective when combined with kinetic copying (green curves), it is only compatible with adiabatic copying when an intermediate state is present (blue curves). This is a key difference between the proofreading schemes in Refs. [3,10]: without an intermediate state it is impossible to find a regime where copies are produced adiabatically and undone very quickly. The combination of kinetic proofreading with an adiabatic copying step has the advantage of a lower dissipation [see Fig. 4(b), green versus blue lines].

In this Letter, we have shown how each copying step in stochastic copying can be unambiguously classified into one of two radically different classes—kinetic and energetic discrimination. These regimes are reminiscent of kinetic and thermodynamic control in chemistry, where, however, the two discrimination factors appear in parallel competing pathways [24]. The existence of an energetic regime in the copolymerization model complements the view in the literature [10–12] that low copy errors are achieved only in a highly dissipative regime. It also demonstrates how entropy-driven growth, a phenomenon studied in the large error regime [10,11,14,15], can be exploited to reliably copy information. Copolymerization is thus compatible with the principle of reversible computing stating that a copy can be performed adiabatically [8]. The analysis of two DNA polymerases, T7 and Pol$\gamma$, shows that the first operates in the energetic regime, while the second in the kinetic one. Both mechanisms are thus used by biological systems. Finally, our study of proofreading proves that the two regimes discussed here can be combined in more complex copying schemes.

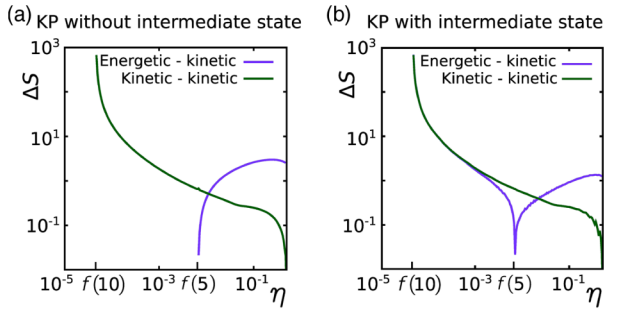


FIG. 4 (color online). (a) Minimum dissipation in proofreading without an intermediate step, Fig. 1(c). For both curves, $\gamma_p = 0$ and $\delta_p = 5$. In the case of energetic discrimination in copying and kinetic discrimination in proofreading (energetic-kinetic), the other parameters are $\gamma = 5$, $\delta = 0$. In the kinetic-kinetic case, we used $\gamma = 0$ and $\delta = 5$. (b) Minimum dissipation in proofreading with an intermediate step, Fig. 1(d). Parameters are as in (a). In both panels, the expected minimum errors $f(10)$ and $f(5)$, depending on whether proofreading is effective or not, are marked.

Our conceptual framework can be applied to a wider range of problems related to stochastic discrimination. Examples are detection of antigens by $T$-cell receptors [25] and discrimination of a binary input in neural dynamics [26]. At the subcellular level, thermal fluctuations dominate and impose constraints on biological tasks. While the thermodynamics of biomechanical systems such as molecular motors is well understood [27], the role of fluctuations in biological information processing, such as bacterial chemotaxis, presents still many open questions [28]. Our work shows that the emerging tradeoffs may be complex and may depend on the region in parameter space where the system operates.

This work was partially supported by a Max Planck Society Scholarship (to P. S.) and a Ramon y Cajal Grant (to S. P.). We are grateful to A. Bernacchia, J. Garcia-Ojalvo, N. Mitarai, L. Granger, M. A. Muñoz, and Y. Tu for a critical reading of the manuscript. S P. acknowledges partial support from Spanish research ministry through Grant No. FIS2012-37655-C02-01.

[1] H. S. Zaher and R. Green, Cell **136**, 746 (2009).
[2] K. A. Johnson, Annu. Rev. Biochem. **62**, 685 (1993).
[3] J. J. Hopfield, Proc. Natl. Acad. Sci. U.S.A. **71**, 4135 (1974).
[4] J. Ninio, Biochimie **57**, 587 (1975).
[5] R. R. Freter and M. A. Savageau, J. Theor. Biol. **85**, 99 (1980).
[6] A. Murugan, D. A. Huse, and S. Leibler, Proc. Natl. Acad. Sci. U.S.A. **109**, 12 034 (2012).

[7] J. W. Shaevitz, E. A. Abbondanzieri, R. Landick, and S. M. Block, Nature (London) **426**, 684 (2003).
[8] C. H. Bennett, Int. J. Theor. Phys. **21**, 905 (1982).
[9] M. Johansson, M. Lovmar, and M. Ehrenberg, Curr. Opin. Microbiol. **11**, 141 (2008).
[10] C. H. Bennett, BioSystems **11**, 85 (1979).
[11] D. Andrieux and P. Gaspard, Proc. Natl. Acad. Sci. U.S.A. **105**, 9516 (2008).
[12] M. Esposito, K. Lindenberg, and C. Van den Broeck, J. Stat. Mech. (2010) P01008.
[13] R. C. Thompson and A. M. Karim, Proc. Natl. Acad. Sci. U.S.A. **79**, 4922 (1982).
[14] C. H. Bennett and M. Donkor, *Information Theory Workshop* (IEEE, Porto, 2008).
[15] C. Jarzynski, Proc. Natl. Acad. Sci. U.S.A. **105**, 9451 (2008).
[16] R. Kawai, J. M. R. Parrondo and C. Van den Broeck, Phys. Rev. Lett. **98**, 080602 (2007).
[17] T. Sagawa and M. Ueda, Phys. Rev. Lett. **102**, 250602 (2009).
[18] L. Granger and H. Kantz, Phys. Rev. E **84**, 061110 (2011).
[19] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.110.188101 for details on mathematical demonstrations and numerical simulations.
[20] H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics* (Wiley, New York, 1985).
[21] H. R. Leei and K. A. Johnson, J. Biol. Chem. **281**, 36 236 (2006)
[22] Y. C. Tsai and K. A. Johnson, Biochemistry **45**, 9675 (2006)
[23] R. Landauer, IBM J. Res. Dev. **5**, 183 (1961).
[24] M. A. Fox and J. K. Whitesell *Organic Chemistry* (Jones and Bartlett, Burlington, MA, 2004).
[25] T. W. McKeithan, Proc. Natl. Acad. Sci. U.S.A. **92**, 5042 (1995).
[26] X. J. Wang, Neuron **36**, 955 (2002).
[27] A. Parmeggiani, F. Jülicher, A. Ajdari, and J. Prost, Phys. Rev. E **60**, 2127 (1999).
[28] G. Lan, P. Sartori, S. Neumann, V. Sourjik, and Y. Tu, Nat. Phys. **8**, 422 (2012).