Template-Directed Copolymerization, Random Walks along Disordered Tracks, and Fractals

Pierre Gaspard
Center for Nonlinear Phenomena and Complex Systems, Université libre de Bruxelles (ULB), Code Postal 231, Campus Plaine, B-1050 Brussels, Belgium
(Received 10 April 2016; published 30 November 2016)

In biology, template-directed copolymerization is the fundamental mechanism responsible for the synthesis of DNA, RNA, and proteins. More than 50 years have passed since the discovery of DNA structure and its role in coding genetic information. Yet, the kinetics and thermodynamics of information processing in DNA replication, transcription, and translation remain poorly understood. Challenging issues are the facts that DNA or RNA sequences constitute disordered media for the motion of polymerases or ribosomes while errors occur in copying the template. Here, it is shown that these issues can be addressed and sequence heterogeneity effects can be quantitatively understood within a framework revealing universal aspects of information processing at the molecular scale. In steady growth regimes, the local velocities of polymerases or ribosomes along the template are distributed as the continuous or fractal invariant set of a so-called *iterated function system*, which determines the copying error probabilities. The growth may become sublinear in time with a scaling exponent that can also be deduced from the iterated function system.

DOI: 10.1103/PhysRevLett.117.238101

The growth of copolymers such as DNA, RNA, and proteins is catalyzed by molecular machines running on templates where the required information is coded in nucleotide sequences. In general, template sequences constitute disordered media for the stochastic motion of molecular machines [1–6]. Since the motion is powered by chemical energy, it is biased and the machine drifts along the template. Accordingly, the kinetic processes of template-directed copolymerization are examples of random walks along disordered one-dimensional chains [7–12]. Yet, the disorder is not fixed during copolymerization. Indeed, the molecular machine synthesizing the copolymer modifies its own motion upon the occurrence of replication errors. In this regard, the growth of the copy differs from random walks on quenched disorder that have previously been investigated. It is only in the limit of perfect error-free replication that copy and template sequences would be identical, and the disorder of the copy would be quenched. In general, copy sequences and their statistical properties are self-generated by the copolymerization process, which is a challenging aspect for kinetic and transport theory.

DNA replication and other biological copolymerization processes such as transcription and translation are often modeled by supposing that rates depend only on whether nucleotide pairing is correct or incorrect, which reduces the kinetics to the one of free copolymerization as if template sequences were homogeneous [13–19]. However, this assumption overlooks sequence heterogeneity effects associated with the fact that the template forms a disordered medium for the motion of molecular machines such as polymerases or ribosomes [1–6]. Moreover, experimental data are available for the rates of all possible pairings in the case of DNA replication [20]. A major challenge is thus to understand the sequence-specific effects of template-directed copolymerization processes and their implications for thermodynamics.

The purpose of the present Letter is to show that the aforementioned difficulties can be overcome in steady growth regimes where the molecular machine is running at a constant mean velocity. In these regimes, copy sequences turn out to form Bernoulli or Markov chains, depending on whether previously incorporated monomeric units influence the molecular machine motion. Accordingly, the kinetic equations ruling the time evolution of the process can be solved to obtain the nucleotide incorporation probabilities by a nonlinear recurrence defined along the template sequence. This recurrence is here demonstrated to be controlled by an *iterated function system* (IFS) [21–23], which remarkably generates fractals under specific conditions. In this way, the probabilities of nucleotide pairings as well as the growth velocity and thermodynamic quantities such as entropy production can be computed more accurately and much faster than with Monte Carlo simulations. Furthermore, it is shown that, as a consequence of sequence heterogeneity, the catalyst random drift may become anomalous with elongation interrupted by frequent stalling and depolymerization steps. The present theory can determine with precision anomalous transport regimes where the mean length of the copy sequence grows sublinearly in time with an exponent that can be calculated, providing a framework for understanding sequence heterogeneity effects.

In the following, the theory will be first developed for kinetics generating Bernoulli chains and, thereafter, for DNA replication, transcription, and translation.

In template-directed copolymerization by a catalyst such as DNA polymerase, monomers $m \in \{1, 2, \ldots, M\}$ (i.e., the nucleotides dNTP) are attached to or detached from the growing copy, which is the nascent DNA strand. Replication proceeds by pairing between the monomeric units $m_l{:}n_l$ of the copy $\omega = m_1 m_2 \ldots m_l$ and template $\alpha = n_1 n_2 \ldots n_l \ldots$. The simplest kinetics supposes that the attachment and detachment rates $W_{\pm m_l, n_l}$ depend only on pairings between the reacting monomer m_l and the corresponding template unit n_l. The kinetic equations can be established for this reaction network. At the single-molecule level, they rule the time evolution of the probabilities to find the copy with a given sequence. With this kinetics, the probability of a given sequence factorizes as for Bernoulli chains:

$$P_t\begin{pmatrix} m_1 \ldots m_l \\ n_1 \ldots n_l n_{l+1} \cdots \end{pmatrix} \simeq p_t(l) \prod_{j=1}^{l} \mu(m_j, j), \qquad (1)$$

in terms of the probability $p_t(l)$ that the copy has the length l at time t and the probabilities $\mu(m_j, j)$ that the copy monomeric unit m_j is paired with the unit n_j at the location j of the template. The length probability $p_t(l)$ obeys the equations for a random walk along a one-dimensional disordered chain [7–12]. The novel aspect is that the hopping rates of this random walk here depend on the probabilities μ of pair formations, which are themselves generated by the process. In steady growth regimes, we may introduce the local velocity x_l, i.e., the rate of elongation by one monomeric unit at the location l of the template. Its inverse x_l^{-1} is the local dwell time of the catalyst. Now, the probability $\mu(m_l, l)$ to find the unit m_l at this location is resulting from the balance between its attachment probability given by the ratio of the attachment rate W_{+m_l, n_l} to the local velocity x_{l-1} and its detachment probability given by the ratio of the detachment rate W_{-m_l, n_l} to the local velocity x_l, multiplied by the probability $\mu(m_l, l)$ itself: $\mu(m_l, l) = (W_{+m_l, n_l}/x_{l-1}) - (W_{-m_l, n_l}/x_l) \times \mu(m_l, l)$. Inverting this relation, we get

$$\mu(m_l, l) = \frac{x_l}{x_{l-1}} \frac{W_{+m_l, n_l}}{W_{-m_l, n_l} + x_l}. \qquad (2)$$

Since these probabilities are normalized to unity, the following IFS is found:

$$x_{l-1} = f_{n_l}(x_l) \quad \text{with} \quad f_n(x) \equiv x \sum_{m=1}^{M} \frac{W_{+m,n}}{W_{-m,n} + x}. \qquad (3)$$

We notice that this nonlinear recurrence is iterated backwards along the template $\alpha = n_1 n_2 \ldots n_l \ldots$. The mean growth velocity v is obtained as

$$\frac{1}{v} = \left\langle \frac{1}{x_l} \right\rangle \equiv \lim_{L \to \infty} \frac{1}{L} \sum_{l=1}^{L} \frac{1}{x_l} \qquad (4)$$

in terms of the local velocities x_l [24].

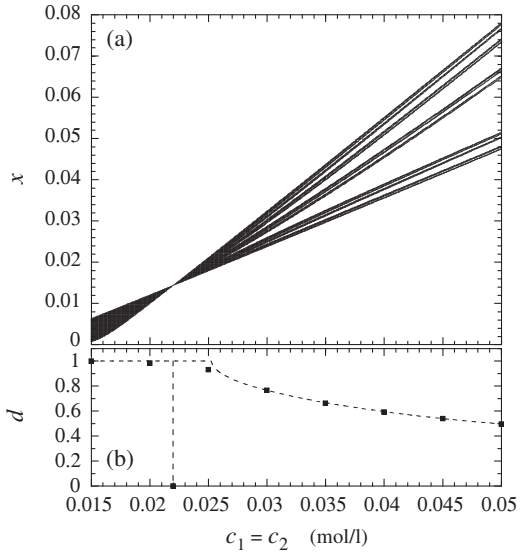

FIG. 1. Kinetics generating Bernoulli chains with the rate constants $k_{+1,1} = 1$, $k_{+1,2} = 0.3$, $k_{+2,1} = 0.2$, $k_{+2,2} = 2$, $k_{-1,1} = 0.01$, $k_{-1,2} = 0.02$, $k_{-2,1} = 0.03$, and $k_{-2,2} = 0.04$. (a) Invariant set of the IFS (3) for the local velocity $x = x_l$ versus the concentration $c_1 = c_2$. (b) Box counting dimension d. The fractal is generated by 10^6 IFS iterations and the dimension computed with 5×10^5 points.

In order to illustrate the IFS method, we consider a kinetics with a mass-action law for the copolymerization of two species. The attachment and detachment rates are, respectively, given by $W_{+m,n} = k_{+m,n} c_m$ and $W_{-m,n} = k_{-m,n}$, where c_m is the constant concentration of monomer m (in moles per liter). Figure 1 shows the IFS invariant set (giving the distribution of local velocities) versus the common concentration $c_1 = c_2$, as well as the corresponding box counting dimension, in the case of a random template. The set is fractal beyond the critical concentration $c_1 = c_2 \simeq 0.02526$, where gaps open. Below, the invariant set has a continuous support that shrinks to a single point at the concentration $c_1 = c_2 \simeq 0.02199$.

The velocity and thermodynamic quantities are plotted in Fig. 2. The results of Monte Carlo simulations [41] are in excellent agreement with the predictions of the IFS method. The internal thermodynamic entropy production (denoted with the subscript i) has the rate

$$\frac{1}{k_B} \frac{d_i S}{dt} = vA = v[\epsilon + D(\omega|\alpha)] \geq 0, \qquad (5)$$

expressed in terms of Boltzmann's constant k_B, the velocity (4), the affinity A (i.e., the entropy production per monomeric unit), the free-energy driving force ϵ, and the conditional Shannon disorder $D(\omega|\alpha)$ of the copy with respect to the template [24]. This disorder is calculated with the probability (2) and can be expressed as the difference $D(\omega|\alpha) = D(\omega) - I(\omega, \alpha)$ between the overall copy disorder and the mutual information between the copy and

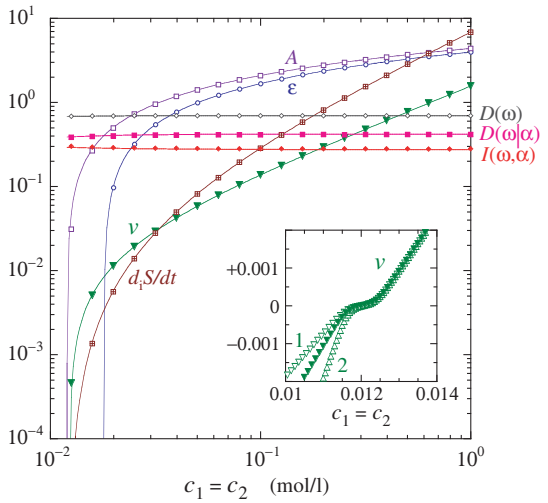

FIG. 2. Kinetics generating Bernoulli chains under the same conditions as in Fig. 1: Velocity v (solid triangles), affinity A (open squares), free-energy driving force ϵ (open circles), conditional disorder $D(\omega|\alpha)$ (solid squares), overall copy disorder $D(\omega)$ (open diamonds), mutual information $I(\omega,\alpha)$ (solid diamonds), and entropy production rate d_iS/dt (crossed squares) versus the concentration $c_1 = c_2$. The dots are obtained by statistics over 10^2 copy sequences of maximum length 10^5 generated by Monte Carlo simulations. The lines are the results of the IFS (3) running over sequences of length 10^5. The inset shows the velocity v at the crossover between copolymerization and depolymerization for an initial copy given by a random sequence (solid triangles) and the periodic sequences 111… and 222… (open triangles). For the inset, the statistics is performed over 10^3 copies of maximum length 10^6. Note that the plateau velocity is not strictly zero because of the finite-time effect.

template, which characterizes the replication fidelity [42]. We see in Fig. 2 that the velocity v vanishes at the concentration $c_1 = c_2 \simeq 0.01238$, together with the entropy production rate (5). In contrast with homogeneous copolymerization processes, this critical concentration does not coincide with thermodynamic equilibrium. Indeed, the copolymer is still growing at lower concentrations but sublinearly in time, so that there is a plateau at $v = 0$, as shown in the inset in Fig. 2. In this anomalous transport regime, the mean length increases as $\langle l \rangle \sim t^{\gamma}$ with an exponent $0 < \gamma < 1$, which is determined by the distribution of the local velocities $x = x_l$. If its probability density is singular as $p(x) \sim x^{\gamma-1}$, the sum of L inverse local velocities has a Lévy distribution of parameter γ, so that the length scales in time with the same exponent. For random templates, the scaling exponent is obtained by solving the equation $\sum_{n=1}^{M} |f_n'(0)|^{-\gamma} = M$, where $f_n'(0)$ is the derivative of the function (3) at $x = 0$ [24]. The exponent γ is zero at the concentration $c_1 = c_2 \simeq 0.01201$, which corresponds to thermodynamic equilibrium, because it is at this critical concentration that the affinity A vanishes.

At still lower concentrations, the copy undergoes depolymerization if its initial length is large enough. Interestingly, the depolymerization speed depends on the initial copy sequence besides the template sequence. This is also illustrated by the inset in Fig. 2, showing the depolymerization velocity of different initial copy sequences. The depolymerization of the periodic sequence 222… is faster than for a random one, while the sequence 111… has the slowest depolymerization. Within numerical errors, the depolymerization velocity vanishes at the concentration $c_1 = c_2 \simeq 0.01165$ for the three sequences. Between this critical concentration and equilibrium, depolymerization is anomalous. The plateau of vanishing velocity in the inset in Fig. 2 is characteristic of random walks in disordered media and shows that anomalous transport also manifests itself in template-directed copolymerization.

The IFS method applies to DNA replication, transcription, and translation. For DNA replication, the attachment and detachment rates depend not only on the last nucleotide that is incorporated in the copy but also on the previous one. Besides, the kinetics is of the Michaelis-Menten type, because base pairing is faster than elongation. As a consequence, the rates depend on three consecutive template nucleotides, and they are denoted for simplicity as $W_{\pm m_l m_{l-1},l}$ with $l \equiv n_{l+1}n_l n_{l-1}$. Because of the dependence on previously incorporated nucleotides, the probability (1) here factorizes as for Markov chains. In this case, the IFS for the local and partial velocities $v_{m_l,l}$ associated with the four nucleotide species $m_l \in \{\mathrm{A},\mathrm{C},\mathrm{G},\mathrm{T}\}$ reads

$$v_{m_{l-1},l-1} = \sum_{m_l} \frac{W_{+m_l m_{l-1},l}}{W_{-m_l m_{l-1},l} + v_{m_l,l}} v_{m_l,l}, \tag{6}$$

which is running backward along the template sequence. Besides, the bulk probabilities of the copy nucleotides at the location l of the template are given by $\bar{\mu}(m_l, l) = u_{m_l,l} v_{m_l,l}$ in terms of the variables $u_{m_l,l}$ that are computed by the forward recurrence:

$$u_{m_l,l} = \sum_{m_{l-1}} \frac{W_{+m_l m_{l-1},l}}{W_{-m_l m_{l-1},l} + v_{m_l,l}} u_{m_{l-1},l-1}, \tag{7}$$

once the partial velocities $v_{m_l,l}$ are determined by the backward recurrence (6). The bulk probabilities are normalized to unity, so that $\sum_{m_l} u_{m_l,l} v_{m_l,l} = 1$. The velocity is given by Eq. (4) with $x_l = (\sum_{m_l} u_{m_l,l})^{-1}$ [24]. In steady growth regimes, the properties of replication kinetics with sequence heterogeneity are surprisingly given with full accuracy by Eqs. (6) and (7), which provide an algorithm much more efficient than Monte Carlo simulations. In the limit where the rates differ only between correct and incorrect pairings, the IFS (6) has a fixed point for the invariant set and the sequence heterogeneity effects disappear.

The proposed algorithm is applied to a realistic model of exonuclease-deficient human mitochondrial DNA polymerase, allowing us to go beyond the mean-field

analysis of Refs. [18,19] that neglected sequence heterogeneity effects. The kinetic parameters of this polymerase have been measured experimentally [20,24]. Sequence heterogeneity effects are more pronounced if the nucleotide concentrations are imbalanced [43], e.g., if $[\mathrm{dATP}] = 10[\mathrm{dCTP}] = 100[\mathrm{dGTP}] = 1000[\mathrm{dTTP}]$. For this example, the velocity and thermodynamic quantities are depicted in Fig. 3 as a function of the dATP concentration, showing the accuracy of the results obtained with Eqs. (6) and (7) plotted as lines with respect to the dots given by Monte Carlo simulations. For a comparable accuracy, the computational time is 10^5 times shorter with the recurrences (6) and (7) than with Monte Carlo simulations. As the dATP concentration decreases, the velocity and entropy production rate also decrease to vanish at the critical concentration $[\mathrm{dATP}] \simeq 1.2 \times 10^{-7}$. Below, the copy growth becomes anomalous, the length sublinearly increasing in time as $\langle l \rangle \sim t^{\gamma}$ with $0 < \gamma < 1$. The reason is the occurrence of numerous stalling and depolymerization events in this regime. The entropy irreversibly produced during the time interval t also increases as $\langle \Delta_i S \rangle \sim t^{\gamma}$. However, the affinity A, the free-energy driving force ϵ, and the conditional disorder $D(\omega|\alpha)$ continue to take nonvanishing values, because they are defined per incorporated nucleotide, as seen in Fig. 3. Thermodynamic equilibrium happens at $[\mathrm{dATP}] \simeq 1.2 \times 10^{-8}$, where the affinity vanishes, $A = 0$. The inset in Fig. 3 shows that the velocity is zero in some interval of dATP concentrations, which is due to anomalous transport caused by sequence heterogeneity. Below equilibrium, the copy undergoes depolymerization, which is itself sublinear in time down to $[\mathrm{dATP}] \simeq 2.5 \times 10^{-9}$ and linear in time for still lower concentrations.

If the velocity is positive enough, the IFS (6) generates a fractal invariant set in the space of the four partial velocities, as illustrated in Fig. 4 at the low concentration $[\mathrm{dATP}] = 5 \times 10^{-7}$. Its box counting dimension has the value $d \simeq 0.3$ in the four-dimensional space, so that this fractal is thin. This dimension increases as the velocity decreases towards the anomalous transport regime.

For the transcription of DNA into RNA, the rates depend on $k = 8$–9 consecutive base pairs of the DNA-RNA hybrid duplex in the polymerase: $W_{\pm m_l m_{l-1} \ldots m_{l-k}, l}$ with $l \equiv n_{l+1} n_l n_{l-1} \ldots n_{l-k}$. Besides the main reaction pathway, backtracking pathways into translocation modes are also observed [5,44,45]. Most remarkably, the IFS method still applies with $m_l m_{l-1}$ replaced by $m_l m_{l-1} \ldots m_{l-k}$ in Eqs. (6) and (7). Instead of Eq. (4), the velocity is here given by $v = \langle [x_l(1 - f_l)]^{-1} \rangle^{-1}$, where $0 \le f_l \le 1$ is the fraction of time spent into backtracking at transcript length l [24]. As a consequence, elongation is slowed down and sequence dependence enhanced. In anomalous transport regimes around the equilibrium concentration or near the stall force, the local velocities $x_l(1 - f_l)$ have a power-law distribution, as in the previous cases. An important implication is that the slowing down promotes error correction by proofreading mechanisms [19,46,47].

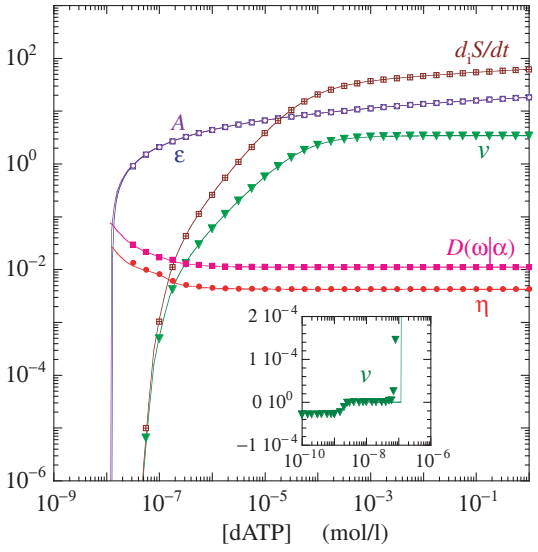

FIG. 3. Exo$^-$ human mitochondrial DNA polymerase: velocity v, affinity A, free-energy driving force ϵ, conditional Shannon disorder $D(\omega|\alpha)$, error probability η, and entropy production rate d_iS/dt versus dATP concentration. The dots are obtained by Monte Carlo simulations generating 10^5 copies of maximum length 10^5. The lines depict the results of the IFS run over sequences of length 10^6. The inset shows the velocity v at the crossover between copolymerization and depolymerization versus the dATP concentration.

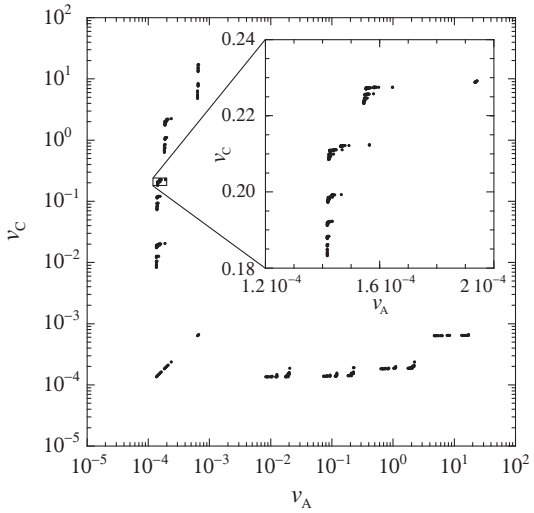

FIG. 4. Exo$^-$ human mitochondrial DNA polymerase: projection of the fractal invariant set of the IFS (6) from the space of the partial velocities (v_A, v_C, v_G, v_T) into the plane (v_A, v_C) at the concentration $[\mathrm{dATP}] = 5 \times 10^{-7}$; otherwise, under the same conditions as in Fig. 3. Inset: Enlargement of the fractal. The fractal is generated by 10^7 iterations of the IFS. Here, the velocity has the value $v = 2.53 \times 10^{-2}$ nt/s.

The IFS method can also be applied to the translation of RNA into proteins by ribosomes, where similar results are expected.

In conclusion, this Letter presents a complete theory for the sequence-dependent motion of molecular machines catalyzing biological template-directed copolymerization, overcoming the difficulty that, if template disorder is quenched, this is not the case for copy disorder, which is self-generated during motion. In steady growth regimes, the theory shows that an IFS rules the statistical properties of copy sequences, which is computationally much faster than Monte Carlo simulations. By varying the external conditions, the IFS invariant set determining the local velocities undergoes metamorphoses from fractal to continuous. The motion becomes anomalous if the invariant set accumulates at zero local velocities. In this way, the IFS method can explain the sequence heterogeneity effects on the statistical and thermodynamic properties of DNA replication, transcription, and translation.

This research is financially supported by the Université libre de Bruxelles (ULB), the FNRS, and the Belgian Federal Government under the Interuniversity Attraction Poles Project No. P7/18.

[1] T. Harms and R. Lipowsky, Phys. Rev. Lett. **79**, 2895 (1997).
[2] F. Jülicher and R. Bruinsma, Biophys. J. **74**, 1169 (1998).
[3] H.-Y. Wang, T. Elston, A. Mogilner, and G. Oster, Biophys. J. **74**, 1186 (1998).
[4] Y. Kafri, D. K. Lubensky, and D. R. Nelson, Biophys. J. **86**, 3373 (2004).
[5] L. Bai, A. Shundrovsky, and M. D. Wang, J. Mol. Biol. **344**, 335 (2004).
[6] L. Bai, R. M. Fulbright, and M. D. Wang, Phys. Rev. Lett. **98**, 068103 (2007).
[7] J. Bernasconi, S. Alexander, and R. Orbach, Phys. Rev. Lett. **41**, 185 (1978).
[8] J. Bernasconi and W. R. Schneider, J. Phys. A **15**, L729 (1982).
[9] B. Derrida and Y. Pomeau, Phys. Rev. Lett. **48**, 627 (1982).
[10] B. Derrida, J. Stat. Phys. **31**, 433 (1983).
[11] C. Aslangul, M. Barthelemy, N. Pottier, and D. Saint-James, J. Stat. Phys. **59**, 11 (1990).
[12] J. P. Bouchaud, A. Comtet, A. Georges, and P. Le Doussal, Ann. Phys. (N.Y.) **201**, 285 (1990).
[13] C. H. Bennett, BioSystems **11**, 85 (1979).
[14] F. Cady and H. Qian, Phys. Biol. **6**, 036011 (2009).
[15] A. K. Sharma and D. Chowdhury, Phys. Rev. E **86**, 011913 (2012).
[16] P. Sartori and S. Pigolotti, Phys. Rev. Lett. **110**, 188101 (2013).
[17] T. Saito, Phys. Rev. E **89**, 062716 (2014).
[18] P. Gaspard, Phys. Rev. E **93**, 042419 (2016).
[19] P. Gaspard, Phys. Rev. E **93**, 042420 (2016).
[20] H. R. Lee and K. A. Johnson, J. Biol. Chem. **281**, 36236 (2006).
[21] M. F. Barnsley and S. Demko, Proc. R. Soc. A **399**, 243 (1985).
[22] C. Van den Broeck and T. Tél, in *From Phase Transitions to Chaos*, edited by G. Györgyi, I. Kondor, L. Sasvári, and T. Tél (World Scientific, Singapore, 1992), pp. 227–236.
[23] T. Wichmann, A. Giacometti, and K. P. N. Murthy, Phys. Rev. E **52**, 481 (1995).
[24] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.117.238101 for the detailed derivation and further implications of the IFS method, which includes Refs. [25–40].
[25] D. Andrieux and P. Gaspard, J. Chem. Phys. **130**, 014901 (2009).
[26] P. Gaspard and D. Andrieux, J. Chem. Phys. **141**, 044908 (2014).
[27] I. Prigogine, *Introduction to Thermodynamics of Irreversible Processes* (Charles C. Thomas, Springfield, IL, 1955).
[28] J. Schnakenberg, Rev. Mod. Phys. **48**, 571 (1976).
[29] G. Nicolis, Rep. Prog. Phys. **42**, 225 (1979).
[30] Luo Jiu-Li, C. Van den Broeck, and G. Nicolis, Z. Phys. B **56**, 165 (1984).
[31] D.-Q. Jiang, M. Qian, and M.-P. Qian, *Mathematical Theory of Nonequilibrium Steady States* (Springer, Berlin, 2004).
[32] P. Gaspard, J. Chem. Phys. **120**, 8898 (2004).
[33] A. A. Johnson and K. A. Johnson, J. Biol. Chem. **276**, 38090 (2001).
[34] S. S. Patel, I. Wong, and K. A. Johnson, Biochemistry **30**, 511 (1991).
[35] T. W. Traut, Mol. Cell. Biochem. **140**, 1 (1994).
[36] J. K. Heinonen, *Biological Role of Inorganic Pyrophosphate* (Springer, New York, 2001).
[37] M. Dangkulwanich, T. Ishibashi, S. Liu, M. L. Kireeva, L. Lubkowska, M. Kashlev, and C. J. Bustamante, eLife **2**, e00971 (2013).
[38] T. Ishibashi, M. Dangkulwanich, Y. Coello, T. A. Lionberger, L. Lubkowska, A. S. Ponticelli, M. Kashlev, and C. Bustamante, Proc. Natl. Acad. Sci. U.S.A. **111**, 3419 (2014).
[39] Y.-G. Shu, Y.-S. Song, Z.-C. Ou-Yang, and M. Li, J. Phys. Condens. Matter **27**, 235105 (2015).
[40] P. Gaspard, J. Stat. Phys. **164**, 17 (2016).
[41] D. T. Gillespie, J. Comput. Phys. **22**, 403 (1976).
[42] D. Andrieux and P. Gaspard, Proc. Natl. Acad. Sci. U.S.A. **105**, 9516 (2008).
[43] D. Kumar, A. L. Abdulovic, J. Viberg, A. K. Nilsson, T. A. Kunkel, and A. Chabes, Nucleic Acids Res. **39**, 1360 (2011).
[44] J. W. Shaevitz, E. A. Abbondanzieri, R. Landick, and S. M. Block, Nature (London) **426**, 684 (2003).
[45] E. A. Galburt, S. W. Grill, A. Wiedmann, L. Lubkowska, J. Choy, E. Nogales, M. Kashlev, and C. Bustamante, Nature (London) **446**, 820 (2007).
[46] M. Voliotis, N. Cohen, C. Molina-París, and T. B. Liverpool, Phys. Rev. Lett. **102**, 258101 (2009).
[47] M. Depken, J. M. R. Parrondo, and S. W. Grill, Cell biology international reports **5**, 521 (2013).