New Journal of Physics **17** (2015) 045016 (*17 pages*)

Force-Velocity Relation for Copolymerization Processes

Pierre Gaspard
Center for Nonlinear Phenomena and Complex Systems – Department of Physics, Université Libre de Bruxelles, Code Postal 231, Campus Plaine, B-1050 Brussels, Belgium

A study is reported of copolymerization processes subjected to an external force in the light of recent advances on the kinetics and thermodynamics of copolymer growth. Different processes are considered: the free copolymerization of Bernoulli and first-order Markov chains and copolymerization with a template. For every process, the dependence on the force is analyzed for the elongation rate, the growth velocity, the disorder in the copolymer sequence, the thermodynamic entropy production, as well as related quantities. It is shown that disorder in the copolymer sequence can generate forces of entropic origin. They are characterized by the value of their stall force in the force-velocity relation of the corresponding process.

I. INTRODUCTION

Molecular motors or growing polymers generate active forces, characterized by their force-velocity relation. Since Huxley's pioneering work on muscle contraction [1], different biophysical systems have been studied including actin-myosin and kinesin-microtubule linear motors, F_oF_1-ATPase and bacterial flagellum rotary motors, as well as growing microtubules [2–19]. In these systems, motion is powered by the free energy of chemical or electrochemical reactions. In particular, pushing forces are generated by the polymerization of microtubules, which provide free energy by binding together protein subunits [13–17].

These processes can be influenced by external forces exerted on such motors, for instance with optical tweezers. Motion can even be stopped by opposing an external force, called the stalling force. At this value of the force, the detachment rate of monomeric units compensates the attachment rate, so that elongation is stopped. These considerations have been much developed for processes taking place in periodic potentials or with identical monomeric units. However, many processes involve different species of monomers or proceed along a heterogeneous filament, as it is the case for DNA or RNA polymerases [20, 21]. For such copolymerization processes, theoretical work has shown that disorder in the sequence of monomeric units remarkably contributes to the thermodynamic entropy production and may even drive the growth of the copolymer [22–28]. In this regard, we may wonder if this sequence disorder could generate a pushing force of entropic origin during copolymerization.

The purpose of the present paper is to show that this can indeed be established on the basis of recent advances [23–25]. According to these advances, the thermodynamic entropy produced per monomeric unit during the growth of copolymers is equal to the sum of the contributions from both chemical free energy and disorder in the growing sequence. If an external force is applied, its mechanical work should also be included. Under the assumption of tight mechano-chemical coupling, the disorder per monomeric unit in the growing copolymer sequence would thus contribute to the value of the stalling force, besides chemical free energy. For certain concentrations of monomers, the copolymer could even generate a pushing force although the pure polymers grown in the absence of comonomers could not.

A stochastic approach is used to describe the random attachments and detachments of monomeric units to the growing copolymer. In this approach, the thermodynamics of these nonequilibrium processes can be established and the entropy production obtained. Fluctuation relations are used in order to characterize the growth processes at the level of the fluctuating numbers of monomeric units incorporated in the copolymer.

The general results of this paper are presented in Section II for copolymerization processes without or with a template. The cases of growing Bernoulli and first-order Markov chains are developed respectively in Sections III and IV. The case of copolymerization with a template is given in Section V. Conclusions are drawn in Section VI. Furthermore, a multivariate fluctuation relation is proved in the Appendix A for the numbers of monomeric units or doublets incorporated in a growing copolymer.

II. GENERAL RESULTS

A. Free copolymerization

We consider the growth of a copolymer. Its description is coarse grained at the level of the sequence $\omega = m_1 m_2 \cdots m_{l-1} m_l$ of its monomeric units $m_j \in \{1, 2, ..., M\}$. M denotes the number of different monomeric species. The chain has a length equal to $l = |\omega|$ units. The solution surrounding the copolymer constitutes a large enough reservoir of monomers so that their concentrations $\{C_m\}_{m=1}^{M}$ remain fixed during the growth process. Moreover, the solution is also a heat bath so that the growth proceeds at constant temperature T. One end of the copolymer is anchored on a surface while monomers are attached to the other end. At this end, a catalyst or enzyme may lower the activation energies of monomer attachments. The reversed reaction of detachment is also possible. Accordingly, the chain length l undergoes a biased random walk. An external force F is exerted at the growing end of the chain, for instance on the catalyst or enzyme, as shown in Fig. 1a. We assume a tight coupling between the mechanical movement of the growing end at the velocity v, and the chemical process of chain elongation at the rate r. Moreover, the dynamics is supposed to be quasi one-dimensional. In other circumstances, the growing copolymer can exert a pushing force on a wall if the copolymer is rigid enough to maintain the directionality of the generated force [13–16].

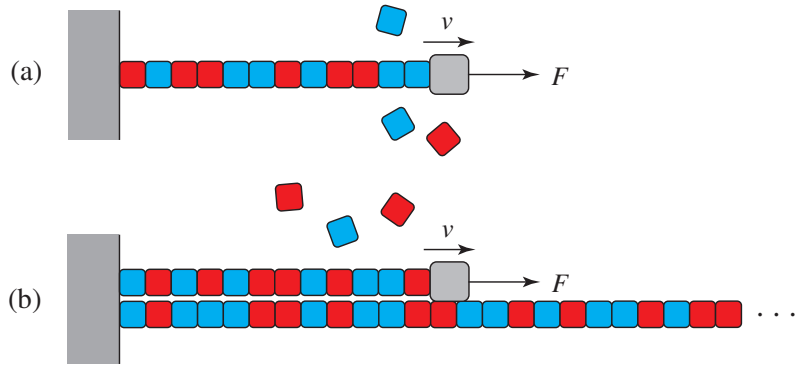

FIG. 1: Schematic representations of force generation by growing copolymers in the cases: (a) of free copolymerization, (b) of copolymerization with a template. F denotes the force exerted on the copolymer chain end, while v is its growth velocity.

Since the attachment and detachment events are random, the process is stochastic and described by the master equation

$$\frac{d}{dt} P_t(\omega) = \sum_{\omega'} \left[P_t(\omega')\, W(\omega' \to \omega) - P_t(\omega)\, W(\omega \to \omega') \right] \tag{1}$$

for the probability $P_t(\omega)$ to find the sequence ω at the time t. The rates of opposite transitions have ratios given by

$$\frac{W(\omega \to \omega')}{W(\omega' \to \omega)} = \mathrm{e}^{\beta \left[G_c(\omega) - G_c(\omega') - F(x_\omega - x_{\omega'}) \right]}, \tag{2}$$

where $\beta = (k_\mathrm{B} T)^{-1}$ is the inverse temperature, k_B denotes Boltzmann's constant, $G_c(\omega)$ is the chemical free enthalpy or Gibbs' free energy of the copolymer ω in the surrounding solution, and x_ω is the position of the copolymer growing end, on which the external force F is exerted.

At the time t, the total entropy of the system is given by

$$S_t = \sum_{\omega} P_t(\omega)\, S(\omega) - k_\mathrm{B} \sum_{\omega} P_t(\omega)\, \ln P_t(\omega)\,, \tag{3}$$

where the first term is the average contribution of the entropy $S(\omega)$ of the copolymer if it has the sequence ω, while the second term is the contribution due to the distribution of the probability $P_t(\omega)$ over the different possible sequences $\{\omega\}$ [29]. The contribution of vibrational, rotational, and translational atomic movements is taken into account in the entropy $S(\omega)$.

Now, the total entropy (3) has a time evolution ruled by the master equation (1). The time derivative of the total entropy can be decomposed as

$$\frac{dS_t}{dt} = \frac{d_\mathrm{e} S}{dt} + \frac{d_\mathrm{i} S}{dt} \tag{4}$$

into the entropy exchange

$$\frac{d_{\rm e}S}{dt} = \sum_{\omega,\omega'} P_t(\omega) W(\omega \to \omega') \left[S(\omega') - S(\omega) - k_{\rm B} \ln \frac{W(\omega \to \omega')}{W(\omega' \to \omega)} \right] \tag{5}$$

and the entropy production

$$\frac{d_{\rm i}S}{dt} = \frac{k_{\rm B}}{2} \sum_{\omega,\omega'} \left[P_t(\omega) W(\omega \to \omega') - P_t(\omega')\, W(\omega' \to \omega) \right] \ln \frac{P_t(\omega)\, W(\omega \to \omega')}{P_t(\omega')\, W(\omega' \to \omega)} \geq 0\,. \tag{6}$$

This latter is always non negative in accordance with the second law of thermodynamics.

Copolymerization is assumed to proceed in a regime of steady growth at a constant mean elongation rate

$$r = \frac{d\langle l \rangle}{dt} \qquad \text{with} \quad \langle l \rangle = \sum_{\omega} P_t(\omega)\, |\omega| \tag{7}$$

and mean growth velocity

$$v = \frac{d\langle x_\omega \rangle}{dt} \qquad \text{with} \quad \langle x_\omega \rangle = \sum_{\omega} P_t(\omega)\, x_\omega\,, \tag{8}$$

which are proportional to each other

$$\frac{d\langle x_\omega \rangle}{dt} = \delta\, \frac{d\langle l \rangle}{dt}\,, \tag{9}$$

δ being the average size of a monomeric unit. In this regime and after a long enough time t, the probability $P_t(\omega)$ is supposed to factorize as

$$P_t(\omega) \simeq p_t(l)\, \mu_l(\omega) \tag{10}$$

into the time-dependent probability $p_t(l)$ that the copolymer contains l monomeric units and the stationary probability distribution $\mu_l(\omega)$ that the copolymer of length $l = |\omega|$ has the sequence ω [23–25]. Using these assumptions and Eq. (2), the entropy production (6) can be expressed as

$$\frac{1}{k_{\rm B}} \frac{d_{\rm i}S}{dt} = r\, (\beta\, F\, \delta - \beta\, g_c + D) \geq 0\,, \tag{11}$$

where

$$g_c = \lim_{l \to \infty} \frac{1}{l} \sum_{\omega} \mu_l(\omega)\, G_c(\omega) \tag{12}$$

is the mean chemical free enthalpy per monomeric unit and

$$D = \lim_{l \to \infty} -\frac{1}{l} \sum_{\omega} \mu_l(\omega)\, \ln \mu_l(\omega) \geq 0 \tag{13}$$

is the Shannon disorder per monomeric unit in the sequence. We notice that the entropy production (11) can be written as $\frac{1}{k_{\rm B}} \frac{d_{\rm i}S}{dt} = rA$, as the product of the elongation rate r with the affinity $A \equiv \beta\, F\, \delta - \beta\, g_c + D$ [23–25]. The tight mechano-chemical coupling explains that a single rate and a single affinity describe this process.

The expression (11) shows that the elongation of the chain can be powered by three possible mechanisms: The first one is the mechanical force F, which can drive elongation if it is positive and large enough, but which is opposed to elongation if it is negative. The second one is the chemical free energy of monomeric attachment, which is driving elongation if the free-energy landscape goes down as the chain grows, i.e., if $g_c < 0$. The third one is the disorder D in the sequence, which is always non negative and can thus also drive elongation if it is large enough to dominate the two other terms [23].

At equilibrium, the elongation rate r and the affinity A are vanishing together with the thermodynamic entropy production (11). Accordingly, the stall force, which stops elongation $r = 0$, is given by

$$F_{\rm st} = \frac{1}{\delta}\, (g_c - k_{\rm B} T\, D)\,. \tag{14}$$

The stall force is negative if it is opposed to elongation, as it should in particular if the free-energy landscape is favorable to the growth when $g_c < 0$.

For the growth of a polymer composed of a single monomeric species, the disorder is vanishing $D = 0$, in which case the stall force is solely determined by the chemical free energy. However, for the growth of a copolymer, the sequence disorder also contributes to the stall force. We see that, even if the free-energy landscape is not favorable to the growth when $g_c > 0$, the sequence disorder can nevertheless generates a contribution such that the stall force is negative opposing the growth if $D > \beta g_c$. In this case, the sequence disorder generates a pushing force of entropic origin, although the copolymer evolves in an adverse free-energy landscape with $g_c > 0$.

B. Copolymerization with a template

In the fundamental biological processes of DNA replication or transcription, the copolymerization of DNA or RNA takes place with a template of sequence $\alpha = n_1 n_2 \cdots n_{l-1} n_l n_{l+1} \cdots$, which has its own statistical distribution $\nu_l(\alpha)$. Now, the transition rates $W(\omega \to \omega')$ also depend on the template sequence, for instance, if the monomers should first form Watson-Crick pairs with the corresponding monomeric units of the template, before their attachment to the growing DNA or RNA strand by some polymerase, as depicted in Fig. 1b.

The previous framework continues to hold. If elongation is again assumed to proceed in a regime of steady growth, the probability $P_t(\omega)$ that the copy has the sequence ω at the time t can also be factorized as

$$P_t(\omega) \simeq p_t(l)\, \mu_l(\omega|\alpha) \tag{15}$$

into the probability $p_t(l)$ that its length is equal to $l = |\omega|$ and the probability $\mu_l(\omega|\alpha)$ that the copy has the sequence ω given that the template has the sequence α [23]. By a reasoning similar as in the previous Subsection II A, the thermodynamic entropy production is obtained as

$$\frac{1}{k_{\rm B}} \frac{d_{\rm i} S}{dt} = r\, [\beta\, F\, \delta - \beta\, g_c + D(\omega|\alpha)] \geq 0\,, \tag{16}$$

in terms of the mean chemical free enthalpy per monomeric unit

$$g_c = \lim_{l\to\infty} \frac{1}{l} \sum_{\alpha,\omega} \nu_l(\alpha)\, \mu_l(\omega|\alpha)\, G_c(\omega) \tag{17}$$

and the conditional Shannon disorder per monomeric unit of the copy ω with respect to the template α:

$$D(\omega|\alpha) = \lim_{l\to\infty} -\frac{1}{l} \sum_{\alpha,\omega} \nu_l(\alpha)\, \mu_l(\omega|\alpha)\, \ln \mu_l(\omega|\alpha) \geq 0\,. \tag{18}$$

Besides this conditional Shannon disorder, we can introduce the overall Shannon disorder of the copy as

$$D(\omega) = \lim_{l\to\infty} -\frac{1}{l} \sum_{\omega} \mu_l(\omega)\, \ln \mu_l(\omega) \geq 0\,, \tag{19}$$

with the probability distribution of the copy whatever the template sequence can be:

$$\mu_l(\omega) = \sum_{\alpha} \nu_l(\alpha)\, \mu_l(\omega|\alpha)\,. \tag{20}$$

We can also introduce the mutual information between the copy and the template as

$$I(\omega,\alpha) = \lim_{l\to\infty} \frac{1}{l} \sum_{\alpha,\omega} \mu_l(\omega,\alpha)\, \ln \frac{\mu_l(\omega,\alpha)}{\nu_l(\alpha)\, \mu_l(\omega)} \geq 0\,, \tag{21}$$

in terms of the joint probability $\mu_l(\omega,\alpha) = \nu_l(\alpha)\, \mu_l(\omega|\alpha)$. This mutual information is always lower or equal to the overall disorders in the copy and the template: $I(\omega,\alpha) \leq {\rm Min}\{D(\omega), D(\alpha)\}$. By a standard formula of information theory [32], the mutual information can be written as

$$I(\omega,\alpha) = D(\omega) - D(\omega|\alpha) \geq 0\,, \tag{22}$$

which characterizes fidelity in the copying process [23].

The thermodynamic entropy production (16) becomes

$$\frac{1}{k_{\rm B}}\frac{d_{\rm i}S}{dt} = r\left[\beta\, F\, \delta - \beta\, g_c + D(\omega) - I(\omega,\alpha)\right] \geq 0\,. \tag{23}$$

Consequently, the stall force is here given by

$$F_{\rm st} = \frac{1}{\delta}\left[g_c - k_{\rm B}T\, D(\omega) + k_{\rm B}T\, I(\omega,\alpha)\right]\,. \tag{24}$$

The stall force is negative and opposed to elongation if the chemical free-energy landscape is favorable to the growth or if sequence disorder in the copy is large enough. Instead, the mutual information has a non-negative contribution to the stall force.

In the following sections, these general results are illustrated in particular cases.

III. THE GROWTH OF BERNOULLI CHAINS

A. The kinetic process

In the simplest growth processes, the rates only depend on the attaching or detaching monomeric unit:

$$m_1 m_2 \cdots m_{l-1} \;+\; m_l \underset{w_{-m_l}}{\overset{w_{+m_l}}{\rightleftharpoons}} m_1 m_2 \cdots m_{l-1} m_l\,. \tag{25}$$

The attachment and detachment rates are supposed to obey mass action law and to have an exponential dependence on the external force F:

$$w_{+m} \;=\; k^0_{+m}\, {\rm e}^{\beta F \delta_{+m}}\, C_m\,, \tag{26}$$

$$w_{-m} \;=\; k^0_{-m}\, {\rm e}^{\beta F \delta_{-m}}\,, \tag{27}$$

where C_m is the concentration of monomers m in the surrounding solution [33, 34]. The ratio of these transition rates is given by

$$\frac{w_{+m}}{w_{-m}} = \frac{k^0_{+m}}{k^0_{-m}}\, {\rm e}^{\beta F \delta_m}\, C_m \qquad \text{with} \quad \delta_m = \delta_{+m} - \delta_{-m}\,, \tag{28}$$

for $m = 1, 2, ..., M$. We notice that, in general, the rates may have non-exponential dependences on the force, as long as their ratio has the required exponential dependence [35].

The master equation of this process takes the following form [24]

$$\begin{aligned}\frac{d}{dt}P_t(m_1\cdots m_{l-1}m_l) = w_{+m_l}P_t(m_1\cdots m_{l-1}) + \sum_{m_{l+1}=1}^{M} w_{-m_{l+1}}P_t(m_1\cdots m_{l-1}m_l m_{l+1}) \\ - \left(w_{-m_l} + \sum_{m_{l+1}=1}^{M} w_{+m_{l+1}}\right) P_t(m_1\cdots m_{l-1}m_l)\,.\end{aligned} \tag{29}$$

Inserting the factorization (10), the stationary probability distribution of the sequences is found to factorize itself into the product

$$\mu_l(m_1\cdots m_{l-1}m_l) = \mu(m_1)\,\cdots\,\mu(m_{l-1})\,\mu(m_l) \tag{30}$$

of the probabilities of the monomeric units

$$\mu(m) = \frac{w_{+m}}{w_{-m}+r}\,, \tag{31}$$

where r is the mean elongation rate given by the non-negative root of

$$\sum_{m=1}^{M}\frac{w_{+m}}{w_{-m}+r} = 1\,, \tag{32}$$

if it exists. Therefore, the growing copolymer forms a Bernoulli chain [24].

B. The thermodynamic entropy production

The thermodynamic entropy production is here given by Eq. (11) with the mean monomeric size

$$\delta = \sum_{m=1}^{M} \mu(m)\, \delta_m \,, \tag{33}$$

the mean chemical free enthalpy per monomeric unit

$$g_c = -k_{\rm B}T \sum_{m=1}^{M} \mu(m)\, \ln \frac{k^0_{+m}\, C_m}{k^0_{-m}} \,, \tag{34}$$

and the Shannon disorder per monomeric unit

$$D = -\sum_{m=1}^{M} \mu(m)\, \ln \mu(m) \geq 0 \,, \tag{35}$$

as it should for a Bernoulli chain [24].

C. The stall force

The stall force has the expression (14), which can be compared with the stall forces of pure polymers grown in the absence of comonomers. If all the concentrations would be vanishing except the one of the species m, the disorder would also be vanishing $D = 0$ and the stall force would take the value

$$F_{{\rm st},m} = -\frac{k_{\rm B}T}{\delta_m}\, \ln \frac{k^0_{+m}\, C_m}{k^0_{-m}} \,. \tag{36}$$

Using Eqs. (33) and (34), the stall force (14) of the copolymer grown in a solution where all the monomers are present at the concentrations $\{C_m\}_{m=1}^{M}$ can be written as

$$F_{\rm st,copol} = \frac{1}{\delta} \left[\sum_{m=1}^{M} \mu(m)\, \delta_m\, F_{{\rm st},m} - k_{\rm B}T\, D \right] . \tag{37}$$

As a corollary, the Shannon disorder per monomeric unit can be obtained as

$$D = \frac{1}{k_{\rm B}T} \sum_{m=1}^{M} \mu(m)\, \delta_m\, (F_{{\rm st},m} - F_{\rm st,copol}) \geq 0 \,. \tag{38}$$

Since the disorder is always non negative, the stall force of the copolymer is always bounded from above by the weighted sum of the stall forces of the pure polymers grown with equal corresponding concentrations of monomers:

$$F_{\rm st,copol} \leq \frac{1}{\delta} \sum_{m=1}^{M} \mu(m)\, \delta_m\, F_{{\rm st},m} \,. \tag{39}$$

D. Illustrative example

In order to illustrate the predictions, the growth of Bernoulli chains composed of $M = 2$ monomeric species can be simulated using Gillespie's algorithm [36, 37] with the following rate constants and concentrations

$$k^0_{+1} = k^0_{+2} = 1 \,, \quad k^0_{-1} = 0.01 \,, \quad k^0_{-2} = 0.001 \,, \quad C_1 = 0.01 \,, \quad C_2 = 0.0005 \,, \tag{40}$$

and $\delta_{\pm 1} = \delta_{\pm 2} = \pm\delta/2$.

Figure 2 depicts the different quantities of interest as a function of the external force F rescaled by $(\beta\delta)^{-1} = k_{\rm B}T/\delta$. We observe that the elongation rate r, the affinity A, and the entropy production $d_{\rm i}S/dt$ are all vanishing at the stall

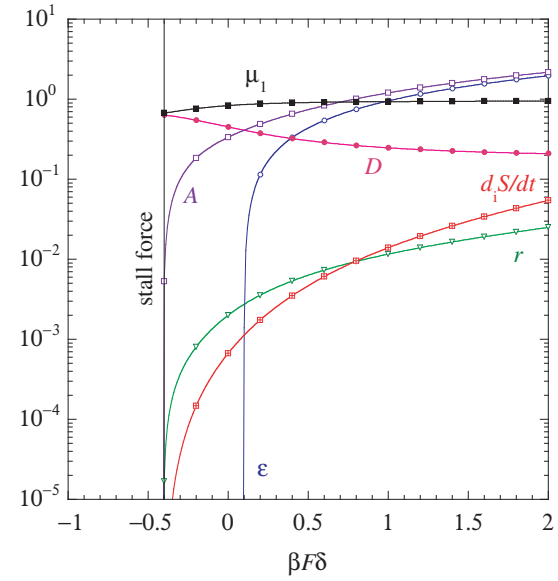

FIG. 2: Growth of Bernoulli chains in the conditions (40) versus the rescaled external force $\beta F\delta$. r is the elongation rate (giving the force-velocity relation), μ_1 the fraction of monomeric units $m = 1$ in the sequence, D the Shannon disorder (35), $\varepsilon = \beta(F\delta - g_c)$ the free-energy driving power, $A = \varepsilon + D$ the affinity, $d_iS/dt = rA$ the entropy production in units with $k_B = 1$. The solid lines are the theoretical predictions by Eqs. (11) and (31)-(35). The symbols are the results of numerical simulations with Gillespie's algorithm.

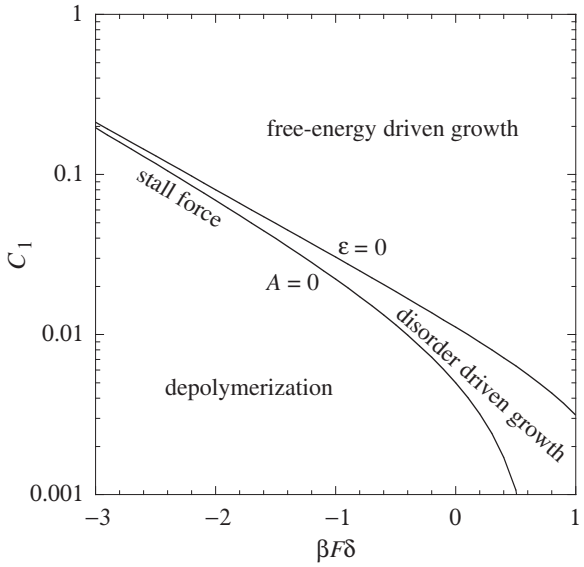

FIG. 3: Growth of Bernoulli chains in the conditions (40) but for a varying concentration C_1 of monomer 1 in solution: the different regimes of depolymerization and growth driven by free energy or sequence disorder are shown versus the concentration C_1 and the rescaled external force $\beta F\delta$.

force $\beta F_{\rm st,copol}\delta = \ln(2/3) \simeq -0.4055$. However, this stall force does not coincide with the value $\beta F\delta = \beta g_c \simeq 0.09878$ of the external force where the free-energy driving power $\varepsilon = \beta(F\delta - g_c)$ is vanishing. This difference finds its origin in the sequence disorder $D = A - \varepsilon$, which shifts the stall force towards a negative value according to Eq. (37). This is an example where the stall force is exerted to oppose a growth driven by sequence disorder. Since the velocity is given by the elongation rate r multiplied by the mean monomeric size (33), $v = r\delta$, the force-velocity relation is shown in Fig. 2 as the curve with the triangles.

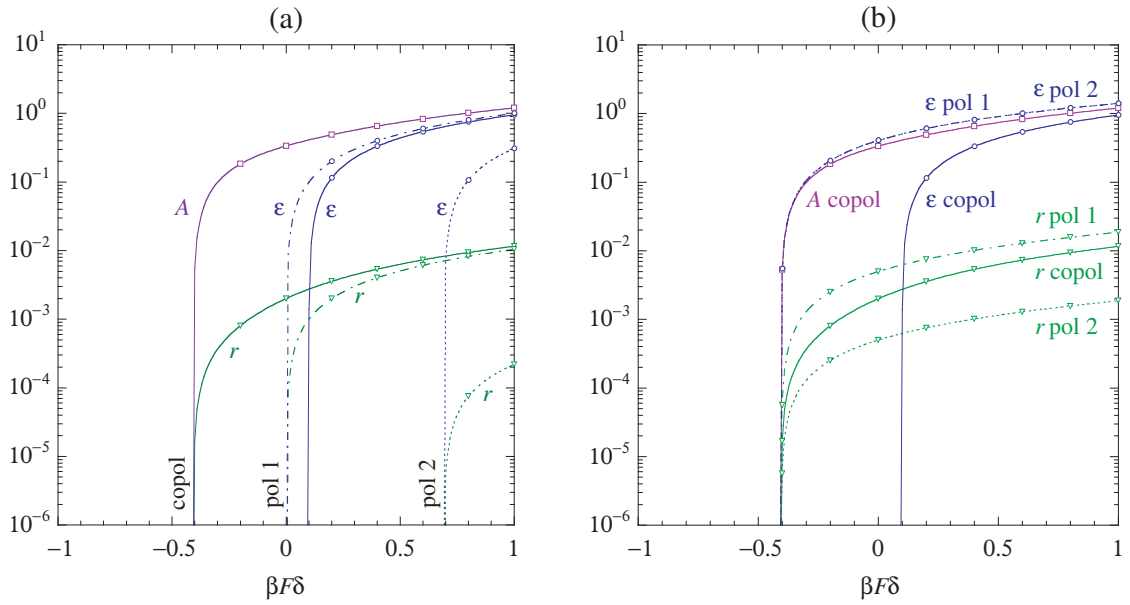

FIG. 4: Growths of Bernoulli chains in the conditions (40) (solid lines and symbols) and of pure polymers with different concentrations: (a) For the polymer 1, the concentrations are $C_1 = 0.01$ and $C_2 = 0$ (dashed-dotted lines and symbols). For the polymer 2, the concentrations are $C_1 = 0$ and $C_2 = 0.0005$ (dashed lines and symbols). (b) For the polymer 1, the concentrations are $C_1 = 0.015$ and $C_2 = 0$ (dashed-dotted lines and symbols). For the polymer 2, the concentrations are $C_1 = 0$ and $C_2 = 0.0015$ (dashed lines and symbols). The elongation rate r (giving the force-velocity relation), the free-energy driving power $\varepsilon = \beta(F\delta - g_c)$, and the affinity $A = \varepsilon + D$ are plotted versus the rescaled external force $\beta F\delta$. The lines are the theoretical predictions by Eqs. (11) and (31)-(35). The symbols are the results of numerical simulations with Gillespie's algorithm.

Figure 3 shows the different possible regimes in the plane of the control parameters given by the concentration C_1 of monomers 1 and the external force F, otherwise in the conditions (40). If the concentration of monomers 1 is too low or the external force too opposing, the copolymer undergoes depolymerization and shrinks. Above the stall force where $A = 0$, the copolymer is growing, first driven by the entropic effect of its sequence disorder and, next, by the free energy of binding the monomeric units. The two growth regimes are separated by the condition of vanishing free-energy driving power $\varepsilon = \beta(F\delta - g_c) = 0$.

Figure 4 compares the growths of the copolymer and pure polymers composed of monomeric units $m = 1$ or $m = 2$. For pure polymers, there is no disorder, $D = 0$, so that the affinity coincides with the free-energy driving power $A = \epsilon$. The affinity of pure polymers is thus given by their free-energy driving power in Fig. 4. Moreover, their stall force varies with the concentration of monomers as shown in Eq. (36). Accordingly, different comparisons can be made depending on the concentrations chosen for the pure polymers.

In Fig. 4a, the concentrations are chosen to illustrate the inequality (39) between the stall forces of the copolymer and the polymers. With this aim, the concentrations should take equal values, because Eqs. (37) and (38) are deduced for the same values of the concentrations in the growths of the copolymer and the corresponding pure polymers. For the concentrations $C_1 = 0.01$ and $C_2 = 0.0005$, the copolymer is composed of monomeric units 1 and 2 with the fractions $\mu(1) = 2/3$ and $\mu(2) = 1/3$ and its stall force is thus equal to $\beta F_{\mathrm{st,copol}}\delta = \ln(2/3) \simeq -0.4055$. According to Eq. (36), the stall force of polymer 1 is vanishing at the concentration $C_1 = 0.01$, $\beta F_{\mathrm{st},1}\delta = 0$, while the stall force of polymer 2 is positive at the concentration $C_2 = 0.0005$: $\beta F_{\mathrm{st},2}\delta = \ln 2 \simeq 0.6931$. Consequently, the inequality (39) is satisfied, as observed in Fig. 4a. Furthermore, the weighted difference (38) between these stall forces gives the value of the Shannon disorder per monomeric unit at the copolymer stall force: $D = \ln 3 - (2/3)\ln 2 \simeq 0.6365$.

In contrast, the concentrations can be chosen for the pure polymers to develop equal stall forces as the copolymer, as shown in Fig. 4b. Using Eq. (36), the conditions $\beta F_{\mathrm{st},1}\delta = \beta F_{\mathrm{st},2}\delta = \beta F_{\mathrm{st,copol}}\delta = \ln(2/3)$ are satisfied with, on the one hand, the concentrations $C_1 = 0.015$ and $C_2 = 0$ for the growth of polymer 1 and, on the other hand, the concentrations $C_1 = 0$ and $C_2 = 0.0015$ for the growth of polymer 2. As seen in Fig. 4b, all the elongation rates and the affinities are vanishing at the same value of the stall force for these concentrations. To get this effect, the concentrations needed for the growths of the pure polymers have to be increased with respect to the concentrations (40) used for the growth of the corresponding copolymer. Different comparisons are thus possible by tuning the concentrations.

E. Multivariate fluctuation relation

In order to follow the stochastic evolution of the copolymerization process, we may consider the random numbers $\{N_m\}_{m=1}^M$ of monomeric units of the different species that are incorporated in the growing chain. These numbers are ruled by a multivariate fluctuation relation, which is proved in the Appendix A. If $p_t(\mathbf{N}) = p_t(N_1, N_2, ..., N_M)$ denotes the probability that $\mathbf{N} = \{N_m\}_{m=1}^M$ monomeric units of the different species have been incorporated into the chain during the time interval $[0, t]$, the probabilities of opposite fluctuations obey the following multivariate fluctuation relation

$$\frac{p_t(\mathbf{N})}{p_t(-\mathbf{N})} \simeq_{t\to\infty} \mathrm{e}^{\mathbf{A}\cdot\mathbf{N}} \,, \tag{41}$$

where $\mathbf{A} = \{A_m\}_{m=1}^M$ are the affinities

$$A_m \equiv \ln \frac{w_{+m}}{w_{-m}\,\mu(m)} \,, \tag{42}$$

associated with the incorporation of each species $m = 1, 2, ..., M$. The aforementioned affinity is given by the average value of the monomer-specific affinities (42): $A = \beta\, F\, \delta - \beta\, g_c + D = \sum_{m=1}^M \mu(m)\, A_m$. The multivariate fluctuation relation can be used in order to determine the monomer-specific affinities by setting to zero all the numbers except one. In the case where $M = 2$, we get both relations:

$$\frac{p_t(N_1, 0)}{p_t(-N_1, 0)} \simeq_{t\to\infty} \mathrm{e}^{A_1 N_1} \,, \tag{43}$$

$$\frac{p_t(0, N_2)}{p_t(0, -N_2)} \simeq_{t\to\infty} \mathrm{e}^{A_2 N_2} \,. \tag{44}$$

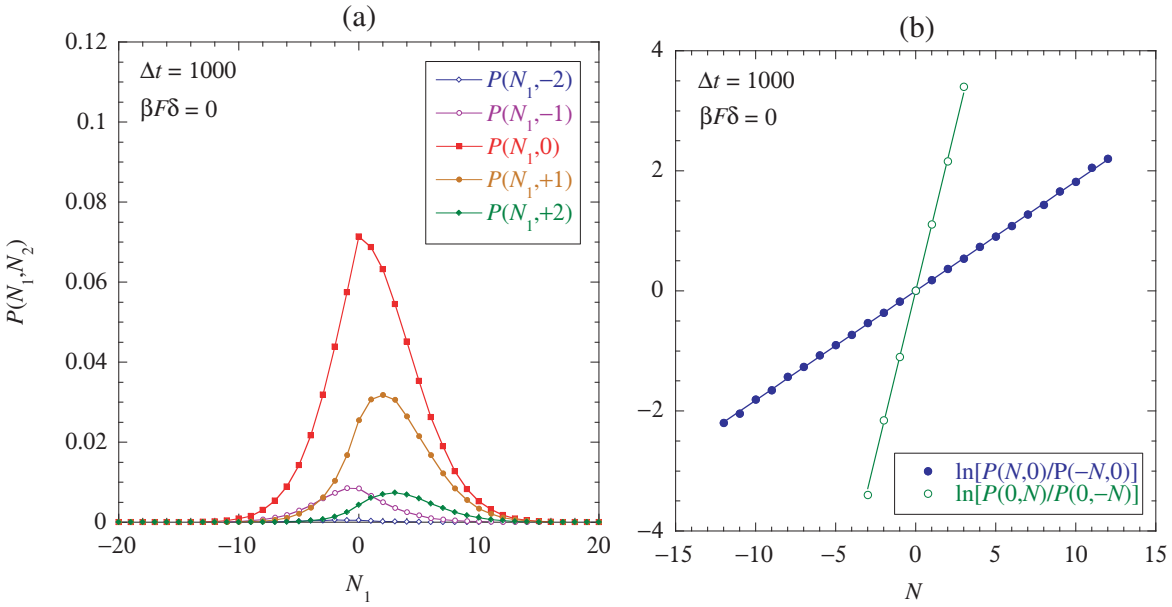

FIG. 5: Multivariate fluctuation relation (41) applied to the stochastic evolution of Bernoulli chains in the conditions (40) and the absence of external force $\beta F\delta = 0$: (a) The probability distribution $p_t(N_1, N_2)$ for the incorporation of N_1 and N_2 monomeric units $m = 1$ and $m = 2$ over a time interval $[0, t]$ with $t = \Delta t = 1000$ versus N_1, for $N_2 = 0, \pm 1, \pm 2$. (b) Logarithms of the ratios of the probabilities of opposite fluctuations $\pm N$ versus the number N with $N = N_1$ (filled circles) and $N = N_2$ (open circles). The symbols give the simulation results and the lines the linear dependences with the affinities $A_1 = 0.1823$ and $A_2 = 1.0986$, as expected from Eqs. (43) and (44). The statistics is carried out over 10^6 random paths.

The multivariate fluctuation relation is applied to copolymerization in the conditions (40) without external force in Fig. 5 and at the stall force in Fig. 6. In the absence of external force, the copolymer is growing so that the probability

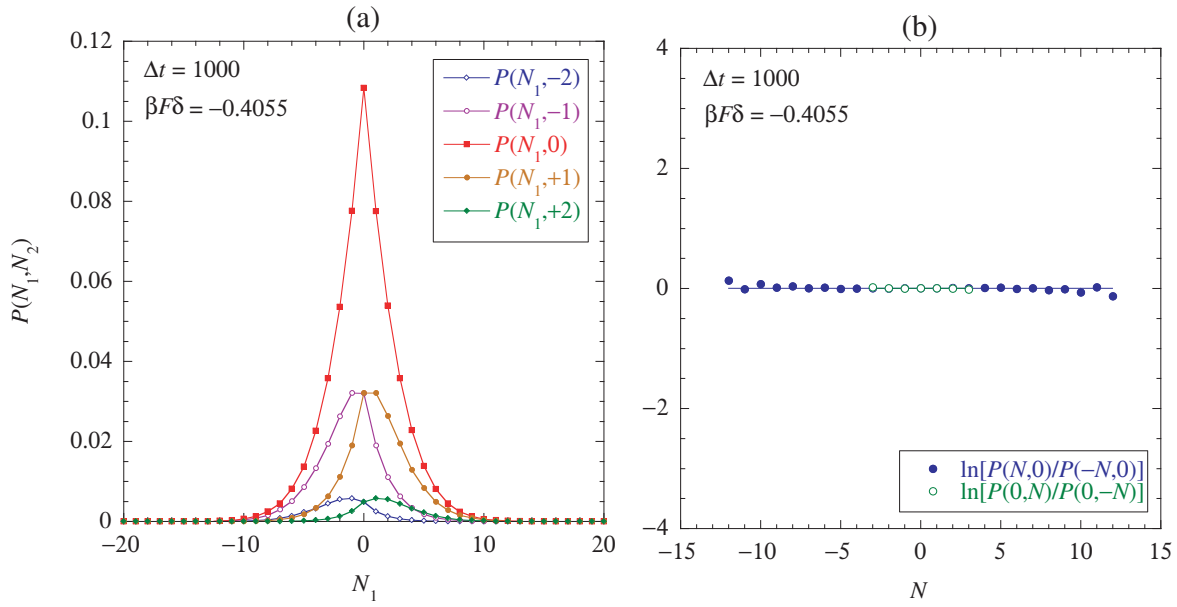

FIG. 6: Multivariate fluctuation relation (41) applied to the stochastic evolution of Bernoulli chains in the conditions (40) at the stall force $\beta F\delta = -0.4055$: (a) The probability distribution $p_t(N_1, N_2)$ for the incorporation of N_1 and N_2 monomeric units $m = 1$ and $m = 2$ over a time interval $[0, t]$ with $t = \Delta t = 1000$ versus N_1, for $N_2 = 0, \pm 1, \pm 2$. (b) Logarithms of the ratios of the probabilities of opposite fluctuations $\pm N$ versus the number N with $N = N_1$ (filled circles) and $N = N_2$ (open circles). The symbols give the simulation results and the lines the linear dependences (43) and (44) with the affinities $A_1 = A_2 = 0$, as expected from Eqs. (43) and (44). The statistics is carried out over 10^6 random paths.

distribution $p_t(N_1, N_2)$ is biased towards the direction of growth and the dependences (43) and (44) determine the affinities (42) driving the incorportation of monomers 1 and 2, as seen in Fig. 5. At the stall force, the numbers N_1 and N_2 fluctuate as in equilibrium so that the principle of detailed balancing is satisfied $p_t(N_1, N_2) \simeq p_t(-N_1, -N_2)$ and the affinities are vanishing $A_1 = A_2 = 0$, as confirmed by Fig. 6.

IV. THE GROWTH OF FIRST-ORDER MARKOV CHAINS

Similar results hold for growing Markov chains.

A. The kinetic process

Here, the attachment and detachment rates may depend on the previously incorporated monomeric unit:

$$m_1 m_2 \cdots m_{l-1} \; + \; m_l \underset{w_{-m_l|m_{l-1}}}{\overset{w_{+m_l|m_{l-1}}}{\rightleftharpoons}} m_1 m_2 \cdots m_{l-1} m_l \, . \tag{45}$$

As before, the rates obey mass action law and are supposed to depend exponentially on the external force F:

$$w_{+n|m} \; = \; k^0_{+n|m} \, \mathrm{e}^{\beta F \delta_{+n|m}} \, C_n \, , \tag{46}$$

$$w_{-n|m} \; = \; k^0_{-n|m} \, \mathrm{e}^{\beta F \delta_{-n|m}} \, , \tag{47}$$

with $m, n = 1, 2, ..., M$ and the concentrations C_n of the monomers n in the surrounding solution. The attachment of the monomeric unit n to the chain ending with the unit m generates the displacement $\delta_{n|m} = \delta_{+n|m} - \delta_{-n|m}$.

The master equation of this process can be solved asymptotically in time by assuming that, in the regime of steady growth, the copolymer sequence forms a first-order Markov chain of distribution

$$\mu_l(m_1 m_2 \cdots m_{l-1} m_l) = \mu(m_1|m_2) \, \cdots \, \mu(m_{l-1}|m_l) \, \mu(m_l) \tag{48}$$

expressed in terms of the conditional probabilities $\mu(m|n)$ to find the monomeric unit m given that the next one is n, and the tip probabilities $\mu(m_l)$ that the chain ends with the monomeric unit m_l [25]. These tip probabilities may differ from the bulk probabilities defined as the stationary probabilities of the Markov chain such that

$$\bar{\mu}(m) = \sum_{n=1}^{M} \mu(m|n) \, \bar{\mu}(n) \, . \tag{49}$$

If r denotes the mean elongation rate, we may introduce partial elongation rates as

$$r_m \equiv r \, \frac{\bar{\mu}(m)}{\mu(m)} \, . \tag{50}$$

These partial elongation rates can be shown to be obtained as the non-negative roots of the following recursive equations:

$$r_m = \sum_{n=1}^{M} \frac{w_{+n|m} \, r_n}{w_{-n|m} + r_n} \, . \tag{51}$$

The tip probabilities are then given by

$$\sum_{m=1}^{M} \frac{w_{+n|m}}{w_{-n|m} + r_n} \, \mu(m) = \mu(n) \, , \tag{52}$$

the conditional probabilities of the Markov chain by

$$\mu(m|n) = \frac{w_{+n|m} \, \mu(m)}{(w_{-n|m} + r_n) \, \mu(n)} \, , \tag{53}$$

and the mean elongation rate by

$$r = \sum_{m=1}^{M} \mu(m) \, r_m \, , \tag{54}$$

as proved in Ref. [25].

In the limit $w_{\pm n|m} = w_{\pm n}$ where the rates no longer depend on the previously incorporated monomeric unit, the conditional, tip, and bulk probabilities coincide, so that the partial elongation rates are equal to the mean one, $r_m = r$ for all $m = 1, 2, ..., M$, and we recover Bernoulli chains.

B. The thermodynamic entropy production

The thermodynamic entropy production (11) is here given in terms of the mean monomeric size

$$\delta = \sum_{m,n=1}^{M} \bar{\mu}(n)\, \mu(m|n)\, \delta_{n|m}\,, \tag{55}$$

the mean chemical free enthalpy per monomeric unit

$$g_c = -k_{\rm B}T \sum_{m,n=1}^{M} \bar{\mu}(n)\, \mu(m|n)\, \ln \frac{k^0_{+n|m}\, C_n}{k^0_{-n|m}}\,, \tag{56}$$

and the Shannon disorder per monomeric unit

$$D = -\sum_{m,n=1}^{M} \bar{\mu}(n)\, \mu(m|n)\, \ln \mu(m|n) \geq 0\,. \tag{57}$$

C. The stall force

The stall force is thus given by the expression (14), which here reads

$$F_{\rm st,copol} = -\frac{k_{\rm B}T}{\delta} \sum_{m,n=1}^{M} \bar{\mu}(n)\, \mu(m|n)\, \ln \frac{k^0_{+n|m}\, C_n}{k^0_{-n|m}\mu(m|n)}\,. \tag{58}$$

Provided that the present process is at equilibrium if the external force takes its stall value, an equivalent way to obtain the stall force is to require that detailed balancing is satisfied and the partial elongation rates are vanishing $r_m = 0$ for all $m = 1, 2, ..., M$. This leads to the condition

$$\det(\mathbf{Z} - \mathbf{1}) = 0\,, \tag{59}$$

where the $M \times M$ matrix $\mathbf{Z} = (z_{n|m})$ is composed of the elements

$$z_{n|m} = \frac{w_{+n|m}}{w_{-n|m}} \tag{60}$$

with $n, m = 1, 2, ..., M$ [25].

D. Illustrative example

The growth of Markov chains composed of $M = 2$ monomeric species is simulated using Gillespie's algorithm [36, 37] with the following rate constants and concentrations

$$\begin{aligned} &k^0_{+1|1} = 2\,, \quad k^0_{+1|2} = k^0_{+2|1} = 1\,, \quad k^0_{+2|2} = 0.5\,, \\ &k^0_{-1|1} = k^0_{-1|2} = k^0_{-2|2} = 0.01\,, \quad k^0_{-2|1} = 0.005\,, \\ &C_1 = 0.003\,, \quad C_2 = 0.01\,, \end{aligned} \tag{61}$$

and $\delta_{\pm m|n} = \pm\delta/2$ for $m, n = 1, 2$.

In the conditions (61), the dependence of different quantities on the external force F is shown in Fig. 7. At the stall force $\beta F_{\rm st,copol}\delta \simeq -0.2823$, the elongation rate r is vanishing together with the affinity A and the entropy production, while the Shannon disorder per monomeric unit takes the value $D \simeq 0.6764$ and the free-energy driving power is negative $\varepsilon = -D \simeq -0.6764$. However, the free-energy driving power vanishes $\varepsilon = 0$ for a positive value $\beta F\delta = \beta g_c \simeq 0.4219$ of the rescaled external force. Between the stall force and this value, there exists the regime of growth driven by sequence disorder.

Figure 8 shows how the affinity A and the elongation rate r, i.e., the growth velocity $v = r\delta$, depend on the rescaled external force $\beta F\delta$ for the copolymer, as well as for the pure polymers grown with different concentrations of

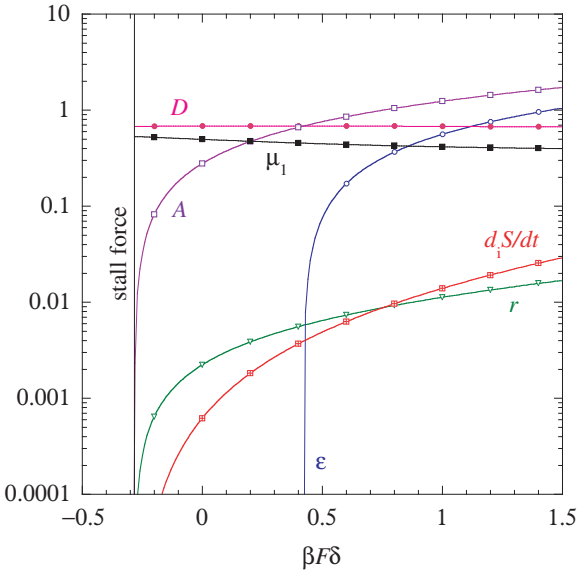

FIG. 7: Growth of first-order Markov chains in the conditions (61) versus the rescaled external force $\beta F\delta$. r is the elongation rate (giving the force-velocity relation), μ_1 the fraction of monomeric units $m = 1$ in the sequence, D the Shannon disorder (57), $\varepsilon = \beta(F\delta - g_c)$ the free-energy driving power, $A = \varepsilon + D$ the affinity, $d_iS/dt = rA$ the entropy production in units with $k_B = 1$. The solid lines are the theoretical predictions by Eqs. (11) and (51)-(57). The symbols are the results of numerical simulations with Gillespie's algorithm.

monomers, in order to illustrate the effects of sequence disorder in the copolymer. As aforementioned, the affinity of pure polymers coincides with their free-energy driving power, $A = \epsilon$, because they have a vanishing sequence disorder $D = 0$.

In Fig. 8a, the concentrations chosen for the growth of pure polymers take the same value as for the growth of the copolymer, but in the absence of their respective comonomers. In this case, the copolymer has a negative stall force, which is opposed to the growth driven by disorder, although the pure polymers have positive stall forces.

In Fig. 8b, the pure polymers are grown with concentrations chosen to satisfy the conditions $\beta F_{\rm st,1}\delta = \beta F_{\rm st,2}\delta = \beta F_{\rm st,copol}\delta = -0.2823$. Now, the concentrations should take the values $C_1 \simeq 0.00663$ and $C_2 = 0$ for the growth of polymer 1, but $C_1 = 0$ and $C_2 \simeq 0.02652$ for the growth of polymer 2. Under such conditions, the pure polymers develop the same stall forces as the copolymer, as seen in Fig. 8b. We notice that the affinity coincides with the free-energy driving power $A = \varepsilon$ for the pure polymers, while they are different for the copolymer because of disorder in its sequence.

E. Multivariate fluctuation relation

The multivariate fluctuation relation can also be applied to the growth of first-order Markov chains. Here, we need to consider the random numbers $\mathbf{N} = \{N_{mn}\}_{m,n=1}^{M}$ of the different monomeric doublets mn found in the growing sequence. For $M = 2$, the probability $p_t(\mathbf{N})$ that $\mathbf{N}$ such doublets have been incorporated into the chain during the time interval $[0, t]$ obey the following multivariate fluctuation relation

$$\frac{p_t(N_{11}, N_{12}, N_{21}, N_{22})}{p_t(-N_{11}, -N_{12}, -N_{21}, -N_{22})} \simeq_{t\to\infty} \mathrm{e}^{A_{11}N_{11}+A_{12}N_{12}+A_{21}N_{21}+A_{22}N_{22}}\,, \tag{62}$$

in terms of the affinities

$$A_{mn} \equiv \ln \frac{w_{+n|m}\,\mu(m)}{w_{-n|m}\,\mu(m|n)\,\mu(n)}\,, \tag{63}$$

with $m, n = 1, 2, ..., M$, as proved in the Appendix A. The overall affinity is the statistical average of these affinities associated with every monomeric doublet: $A = \varepsilon + D = \beta\,F\,\delta - \beta\,g_c + D = \sum_{m,n=1}^{M} \bar{\mu}(n)\mu(m|n)\,A_{mn}$. As in the

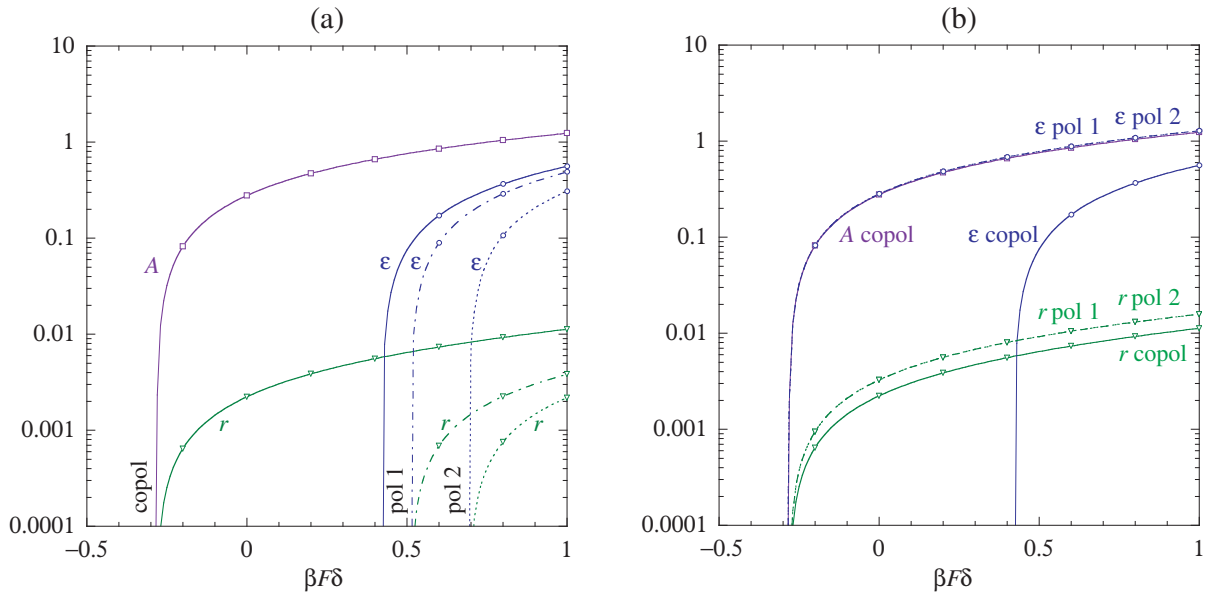

FIG. 8: Growth of first-order Markov chains in the conditions (61) (solid lines and symbols) and of pure polymers with different concentrations: (a) For the polymer 1, the concentrations are $C_1 = 0.003$ and $C_2 = 0$ (dashed-dotted lines and symbols). For the polymer 2, the concentrations are $C_1 = 0$ and $C_2 = 0.01$ (dashed lines and symbols). (b) For the polymer 1, the concentrations are $C_1 = 0.00663$ and $C_2 = 0$ (dashed-dotted lines and symbols). For the polymer 2, the concentrations are $C_1 = 0$ and $C_2 = 0.02652$ (dashed lines and symbols). The elongation rate r (giving the force-velocity relation), the free-energy driving power $\varepsilon = \beta(F\delta - g_c)$, and the affinity $A = \varepsilon + D$ are plotted versus the rescaled external force $\beta F\delta$. The lines are the theoretical predictions by Eqs. (11) and (51)-(57). The symbols are the results of numerical simulations with Gillespie's algorithm.

previous section, the multivariate fluctuation relation can be used in order to determine the affinities (63) by setting to zero all the numbers $N_{m'n'}$ except the one N_{mn} corresponding to the affinity of interest A_{mn}.

This is illustrated for the growth of first-order Markov chains in the conditions (61) without external force in Fig. 9a, and at the stall force in Fig. 9b. In the absence of external force, the four affinities (63) are positive, although they vanish at the stall force, which corresponds to equilibrium conditions. In Fig. 9, we see the agreement between numerical simulations with Gillespie's algorithm and the theoretical expectation (63), which supports the multivariate fluctuation relation (62).

V. THE GROWTH OF COPOLYMERS WITH A TEMPLATE

In this section, an example is considered illustrating copolymerization processes with a template.

A. The kinetic process

The kinetics proceeds by random attachments and detachments of monomeric units to a growing chain in contact with a template so that the rates depend not only on the involved monomer, but also on the corresponding monomeric unit of the template:

$$\begin{array}{lll} \alpha: & n_1\, n_2\, \cdots n_{l-1}\, n_l\, n_{l+1}\, \cdots & \overset{w_{+m_l|n_l}}{\underset{w_{-m_l|n_l}}{\rightleftharpoons}} \quad n_1\, n_2\, \cdots n_{l-1}\, n_l\, n_{l+1}\, \cdots \\ \omega: & m_1 m_2 \cdots m_{l-1} \qquad + m_l & \qquad\quad m_1 m_2 \cdots m_{l-1} m_l \end{array} \tag{64}$$

The attachment and detachment rates are here also supposed to obey mass action law and to have an exponential dependence on the external force F:

$$w_{+m|n} = k^0_{+m|n}\, \mathrm{e}^{\beta F \delta_{+m|n}}\, C_m\,, \tag{65}$$

$$w_{-m|n} = k^0_{-m|n}\, \mathrm{e}^{\beta F \delta_{-m|n}}\,, \tag{66}$$

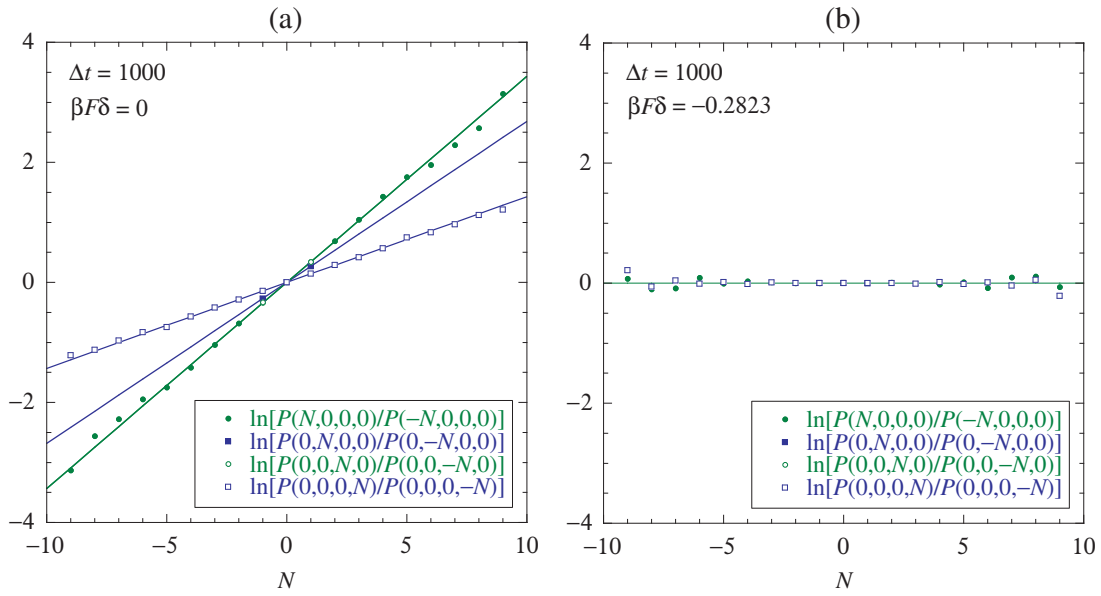

FIG. 9: Multivariate fluctuation relation (62) applied to first-order Markov chains in the conditions (61) over a time interval $[0,t]$ with $t = \Delta t = 1000$: (a) in the absence of external force $\beta F\delta = 0$; (b) at the stall force $\beta F\delta = -0.2823$. The plots show the logarithms of the ratios of the probabilities of opposite fluctuations versus the number $N = N_{mn}$ for $mn = 11$ (filled circles), $mn = 12$ (filled squares), $mn = 21$ (open circles), and $mn = 22$ (open squares). The symbols give the simulation results with Gillespie's algorithm and the lines the predictions of the multivariate fluctuation relation (62) with the expected affinities $A_{11} = 0.343327$, $A_{12} = 0.268250$, $A_{21} = 0.343339$, and $A_{22} = 0.143099$. The statistics is carried out over 10^7 random paths.

where C_m is the concentration of monomers m in the surrounding solution so that their ratios are given by

$$\frac{w_{+m|n}}{w_{-m|n}} = \frac{k^0_{+m|n}}{k^0_{-m|n}}\, \mathrm{e}^{\beta F \delta_{m|n}}\, C_m \qquad \text{with} \quad \delta_{m|n} = \delta_{+m|n} - \delta_{-m|n}\,, \tag{67}$$

for $m, n = 1, 2, ..., M$.

We notice that, for processes with DNA or RNA polymerases, the kinetics is of Michaelis-Menten type, which requires a specific study beyond the aims of the present paper.

The process defined by Eqs. (64)-(66) can be simulated with Gillespie's algorithm [36, 37] to obtain the different quantities of interest given by Eqs. (17)-(23). We consider an example with $M = 2$ monomeric species. The template is taken as a Bernoulli chain of probabilities $\nu(1) = \nu(2) = 1/2$, in which case the copy is itself a Bernoulli chain and the stationary probability distribution of the sequences introduced with Eq. (15) is factorized as $\mu_l(\omega|\alpha) = \prod_{j=1}^{l} \mu(m_j|n_j)$.

B. Illustrative example

The following rate constants and concentrations are chosen for simulations:

$$\begin{aligned} &k^0_{+1|1} = k^0_{+2|2} = 1\,, \quad k^0_{+1|2} = k^0_{+2|1} = 0.5\,, \\ &k^0_{-1|1} = k^0_{-1|2} = k^0_{-2|2} = 0.001\,, \quad k^0_{-2|1} = 0.002\,, \\ &C_1 = 0.001\,, \quad C_2 = 0.0015\,, \end{aligned} \tag{68}$$

with $\delta_{\pm m|n} = \pm\delta/2$ for $m, n = 1, 2$.

The numerical results of simulations with Gillespie's algorithm are shown in Fig. 10 as a function of the rescaled external force $\beta F\delta$. The elongation rate r and, thus, the growth velocity $v = r\delta$ are vanishing at the stall force $\beta F_{\rm st}\delta \simeq -0.5$ given by Eq. (24), together with the affinity $A = \varepsilon + D(\omega|\alpha)$, and the entropy production $\frac{1}{k_{\rm B}}\frac{d_{\rm i}S}{dt} = rA$. In contrast, the free-energy driving power $\varepsilon = \beta(F\delta - g_c)$ vanishes at the positive value $\beta F\delta \simeq 0.1$ of the rescaled external

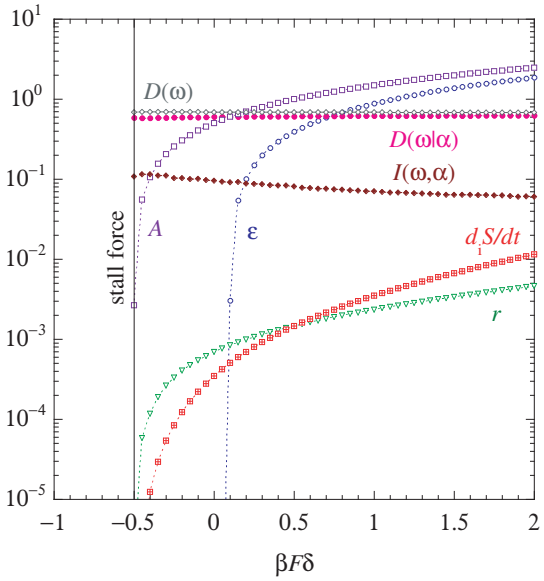

FIG. 10: Copolymerization with a template in the conditions (68) versus the rescaled external force $\beta F\delta$: r is the elongation rate (giving the force-velocity relation), $\varepsilon = \beta(F\delta - g_c)$ the free-energy driving power, $D(\omega|\alpha)$ the conditional Shannon disorder of the copy (18), $A = \varepsilon + D(\omega|\alpha)$ the affinity, $d_iS/dt = rA$ the entropy production (23) in units with $k_B = 1$, $D(\omega)$ the overall Shannon disorder of the copy (19), and $I(\omega, \alpha)$ the mutual information (22) between the copy and the template. The template is a Bernoulli chain of probabilities $\nu(1) = \nu(2) = 1/2$. The symbols connected by dashed lines are the results of numerical simulations with Gillespie's algorithm.

force. Again, the growth is driven by sequence disorder, which is dominant between these two values of the external force. This disorder is characterized by the conditional Shannon disorder (18) and the overall disorder (19), which are depicted in Fig. 10. The difference between the overall and conditional disorders gives the mutual information (22) between the copy and the template. As seen in Fig. 10, the mutual information is significantly lower than the disorders, which means that the copying process is affected by a lot of errors in the present example.

VI. CONCLUSIONS

In the present paper, we have studied how copolymerization processes may depend on an external force exerted on the end of a growing chain. Different processes are considered: the free copolymerization of Bernoulli or first-order Markov chains and the copolymerization with a template, which concerns the biological processes of DNA replication, transcription of DNA into mRNA, and translation of mRNA into proteins.

A stochastic approach is used to describe the random process of attachment and detachment of monomeric units to a copolymer growing on average in a surrounding solution where the monomers are supplied at fixed concentrations. If an external force is exerted on the process, its transition rates have some dependence on this force. Since the rates of opposite transitions have ratios related to the free energy of the states, these ratios generally depend exponentially on the external force. The master equation of the stochastic process rules the time evolution of every quantity of interest and, in particular, the thermodynamic entropy. In this way, the entropy production can be obtained for copolymerization processes. Surprisingly, the entropy production depends not only on the external force and the chemical free energy of monomer binding to the growing chain, but also on the Shannon disorder per monomeric unit in the copolymer sequence. Consequently, the stall force, for which the mean elongation rate vanishes, is influenced by the sequence disorder. This feature is specific to copolymers and does not arise for pure polymers, in which case the stall force is only determined by the chemical free energy. Accordingly, the growth of a copolymer can proceed in two possible regimes driven either by a favorable free-energy landscape, or by a large enough disorder in the growing sequence although the free-energy landscape is adverse. In this latter regime, the growth is driven by the entropic effect of sequence disorder and the force exerted by the growing copolymer would thus be of entropic origin. This concerns free copolymerization, as well as copolymerization with a template, as illustrated with several examples.

For growing Bernoulli chains, the stall force of copolymers is bounded from above by the average value of the stall forces of pure polymers grown in corresponding conditions, the difference being due to the Shannon disorder per

monomeric units, as shown by Eqs. (38) and (39). An illustrative example is presented where the relation between the external force and the growth velocity is calculated theoretically and with numerical simulations by Gillespie's algorithm, together with other quantities of interest including the thermodynamic entropy production. For this example, the stall force is compared between the copolymer and pure polymers grown with different concentrations, demonstrating the influence of sequence disorder in the copolymer. Moreover, the time evolution of copolymerization can be investigated thanks to a multivariate fluctuation relation for the random numbers of monomeric units incorporated in the chain. This fluctuation relation is proved in the Appendix A.

Similar results are obtained for growing first-order Markov chains. Here also, the kinetic equations can be solved analytically in the long-time limit and compared with numerical simulations by Gillespie's algorithm in order to obtain the dependences of the elongation rate and entropy production on the external force. Here also, an example is presented where the stall force of the copolymer is compared with the stall forces of pure polymers grown in different conditions. The growth process can be analyzed in great detail thanks to the multivariate fluctuation relation, which allows us to determine the individual affinities characterizing the incorporation of monomeric doublets in the Markov chain. At the stall force, these affinities are vanishing as in equilibrium conditions.

For copolymerization with a template, simulations with Gillespie's algorithm shows again that sequence disorder may generate forces of entropic origin in an adverse free-energy landscape.

These results could be tested experimentally with optical trapping techniques to exert forces on growing copolymers or on polymerases catalyzing the chain elongation. Knowing the sequence of growing copolymer or the template, the effect of its disorder could be studied as a function of the exerted force.

Acknowledgments

This research is financially supported by the Université Libre de Bruxelles and the Belgian Science Policy Office under the Interuniversity Attraction Pole project P7/18 "DYGEST".

APPENDIX A: THE MULTIVARIATE FLUCTUATION RELATION FOR COPOLYMERIZATION

In this Appendix, the multivariate fluctuation relation is proved for the growth of first-order Markov chains thanks to methods based on previously work [30, 31].

The probability $p_t(\mathbf{N})$ that $\mathbf{N} = \{N_{mn}\}$ monomeric doublets mn have been incorporated in the chain during the time interval $[0, t]$ can be obtained as

$$p_t(\mathbf{N}) = \sum_{m_1 \cdots m_{l-1} m_l} P_t(m_1 \cdots m_{l-1} m_l, \mathbf{N}) \tag{A1}$$

in terms of the probability $P_t(m_1 \cdots m_{l-1} m_l, \mathbf{N})$ that the copolymer has the sequence $\omega = m_1 \cdots m_{l-1} m_l$ and contains $\mathbf{N}$ monomeric doublets at the time t. This latter is ruled by the master equation:

$$\begin{aligned} \frac{d}{dt} P_t(m_1 \cdots m_{l-1} m_l, \mathbf{N}) &= w_{+m_l|m_{l-1}} \, P_t(m_1 \cdots m_{l-1}, \mathbf{N} - \mathbf{1}_{m_{l-1} m_l}) \\ &+ \sum_{m_{l+1}=1}^{M} w_{-m_{l+1}|m_l} \, P_t(m_1 \cdots m_{l-1} m_l m_{l+1}, \mathbf{N} + \mathbf{1}_{m_l m_{l+1}}) \\ &- \left(w_{-m_l|m_{l-1}} + \sum_{m_{l+1}=1}^{M} w_{+m_{l+1}|m_l} \right) P_t(m_1 \cdots m_{l-1} m_l, \mathbf{N}) \end{aligned} \tag{A2}$$

with the notation $\mathbf{N} \pm \mathbf{1}_{mn} = \{N_{11}, N_{12}, ..., N_{mn} \pm 1, ..., N_{MM}\}$. In the long-time limit, the probability factorizes as

$$P_t(m_1 m_2 \cdots m_l, \mathbf{N}) \simeq p_t(\mathbf{N}) \, \mu(m_1|m_2) \, \cdots \, \mu(m_{l-1}|m_l) \, \mu(m_l) \tag{A3}$$

in terms of the conditional and tip probabilities given by Eqs. (51)-(53) [25]. Accordingly, the probabilities $p_t(\mathbf{N})$ are ruled by the equation $dp_t/dt = \hat{L} p_t$ with the linear operator:

$$\hat{L} = \sum_{m,n=1}^{M} \left[w_{+n|m} \, \mu(m) \left(\hat{E}^{-}_{mn} - 1 \right) + w_{-n|m} \, \mu(m|n) \, \mu(n) \left(\hat{E}^{+}_{mn} - 1 \right) \right] , \tag{A4}$$

which is written in terms of the creation-annihilation operators [38]

$$\hat{E}^{\pm}_{mn} f(\mathbf{N}) = \exp\left(\pm \frac{\partial}{\partial N_{mn}}\right) f(\mathbf{N}) = f(\mathbf{N} \pm \mathbf{1}_{mn})\,. \tag{A5}$$

The cumulant generating function of the numbers of doublets incorporated in the chain is defined as

$$Q(\boldsymbol{\lambda}) \equiv \lim_{t\to\infty} -\frac{1}{t} \ln \left\langle \mathrm{e}^{-\boldsymbol{\lambda}\cdot\mathbf{N}} \right\rangle_t \tag{A6}$$

in terms of the counting parameters $\boldsymbol{\lambda} = \{\lambda_{mn}\}$. This generating function can be obtained by solving the eigenvalue problem $\hat{L}_{\boldsymbol{\lambda}}\phi = -Q(\boldsymbol{\lambda})\phi$ for the modified operator

$$\hat{L}_{\boldsymbol{\lambda}} \equiv \mathrm{e}^{-\boldsymbol{\lambda}\cdot\mathbf{N}} \hat{L}\, \mathrm{e}^{+\boldsymbol{\lambda}\cdot\mathbf{N}}\,. \tag{A7}$$

The eigenfunctions of this operator take the form $\phi = \mathrm{e}^{i\mathbf{k}\cdot\mathbf{N}}$. In the limit $\mathbf{k} \to 0$, the leading eigenvalue gives the cumulant generating function:

$$Q(\boldsymbol{\lambda}) = \sum_{m,n=1}^{M} \left[w_{+n|m}\, \mu(m) \left(1 - \mathrm{e}^{-\lambda_{mn}}\right) + w_{-n|m}\, \mu(m|n)\, \mu(n) \left(1 - \mathrm{e}^{+\lambda_{mn}}\right) \right] , \tag{A8}$$

which obeys the symmetry relation

$$Q(\boldsymbol{\lambda}) = Q(\mathbf{A} - \boldsymbol{\lambda}) \tag{A9}$$

in terms of the affinities (63). By using large-deviation theory, we finally infer the multivariate fluctuation relation (62).

In the limit $w_{\pm n|m} = w_{\pm n}$ where the transition rates no longer depend on the previously incorporated monomer, the affinities (63) reduce to those of Bernoulli chains given by Eq. (42): $A_{mn} = A_n$. In this Bernoullian limit, we get

$$\mathbf{A}\cdot\mathbf{N} = \sum_{m,n=1}^{M} A_{mn} N_{mn} = \sum_{n=1}^{M} A_n N_n \qquad \text{with} \quad N_n = \sum_{m=1}^{N} N_{mn}\,, \tag{A10}$$

and the multivariate fluctuation relation (41) is recovered.

[1] A. F. Huxley, Prog. Biophys. **7**, 255 (1957).
[2] T. L. Hill, Proc. Natl. Acad. Sci. USA **78**, 5613 (1981).
[3] T. L. Hill and M. W. Kirschner, Proc. Natl. Acad. Sci. USA **79**, 490 (1982).
[4] A. J. Hunt, F. Gittes, and J. Howard, Biophys. J. **67**, 766 (1994).
[5] M. J. Schnitzer, K. Visscher, and S. M. Block, Nat. Cell Biol. **2**, 718 (2000).
[6] G. Oster, H.-Y. Wang, and M. Grabe, Phil. Trans. R. Soc. Lond. B **355**, 523 (2000).
[7] A. Mogilner and G. Oster, Biophys. J. **84**, 1591 (2003).
[8] J. Xing, H.-Y. Wang, C. von Ballmoos, P. Dimroth, and G. Oster, Biophys. J. **87**, 2148 (2004).
[9] G. Oster and A. Mogilner, in: A. Ciferri, Editor, *Supramolecular Polymers*, 2nd Edition (CRC Press, Boca Raton, 2005) pp. 741-752.
[10] S. Toyabe, T. Watanabe-Nakayama, T. Okamoto, S. Kudo, and E. Muneyuki, Proc. Natl. Acad. Sci. USA **108**, 17951 (2011).
[11] J. Xing, F. Bai, R. Berry, and G. Oster, Proc. Natl. Acad. Sci. USA **103**, 1260 (2006).
[12] Y. Sowa and R. M. Berry, Q. Rev. Biophys. **41**, 103 (2008).
[13] M. Dogterom and B. Yurke, Science **278**, 856 (1997).
[14] A. Mogilner and G. Oster, Eur. Biophys. J. **28**, 235 (1999).
[15] G. Sander van Doorn, C. Tanase, B. M. Mulder, and M. Dogterom, Eur. Biophys. J. **29**, 2 (2000).
[16] J. A. Theriot, Traffic **1**, 19 (2000).
[17] A. B. Kolomeisky and M. E. Fisher, Biophys. J. **80**, 149 (2001).
[18] M. J. Footer, J. W. J. Kerssemakers, J. A. Theriot, and M. Dogterom, Proc. Natl. Acad. Sci. USA **104**, 2181 (2007).
[19] C. Brangbour, O. du Roure, E. Helfer, D. Démoulin, A. Mazurier, M. Fermigier, M.-F. Carlier, J. Bibette, and J. Baudry, PLoS Biol. **9**, e1000613 (2011).
[20] F. Jülicher and R. Bruinsma, Biophys. J. **74**, 1169 (1998).
[21] H.-Y. Wang, T. Elston, A. Mogilner, and G. Oster, Biophys. J. **74**, 1186 (1998).

[22] C. H. Bennett, Biosystems **11**, 85 (1979).
[23] D. Andrieux and P. Gaspard, Proc. Natl. Acad. Sci. USA **105**, 9516 (2008).
[24] D. Andrieux and P. Gaspard, J. Chem. Phys. **130**, 014901 (2009).
[25] P. Gaspard and D. Andrieux, J. Chem. Phys. **141**, 044908 (2014).
[26] C. Jarzynski, Proc. Natl. Acad. Sci. USA **105**, 9451 (2008).
[27] M. Esposito, K. Lindenberg, and C. Van den Broeck, J. Stat. Mech.: Th. Exp. P01008 (2010).
[28] H.-J. Woo and A. Wallqvist, Phys. Rev. Lett. **106**, 060601 (2011).
[29] P. Gaspard, J. Chem. Phys. **120**, 8898 (2004).
[30] D. Andrieux and P. Gaspard, J. Chem. Phys. **121**, 6167 (2004).
[31] D. Andrieux and P. Gaspard, J. Stat. Phys. **127**, 107 (2007).
[32] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, Hoboken, 2006).
[33] M. E. Fisher and A. B. Kolomeisky, Physica A **274**, 241 (1999).
[34] A. B. Kolomeisky and M. E. Fisher, Annu. Rev. Phys. Chem. **58**, 675 (2007).
[35] E. Gerritsma and P. Gaspard, Biophys. Rev. Lett. **5**, 163 (2010).
[36] D. T. Gillespie, J. Comput. Phys. **22**, 403 (1976).
[37] D. T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).
[38] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1981).