Nonequilibrium Transitions in a Template Copying Ensemble

Arthur Genthon,[1,*] Carl D. Modes,[2,3,4,†] Frank Jülicher,[1,3,4,‡] and Stephan W. Grill[2,3,4,§]
[1]Max Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany
[2]Max Planck Institute for Molecular Cell Biology and Genetics, 01307 Dresden, Germany
[3]Center for Systems Biology Dresden, 01307 Dresden, Germany
[4]Cluster of Excellence, Physics of Life, TU Dresden, 01307 Dresden, Germany

(Received 9 March 2024; accepted 10 January 2025; published 14 February 2025)

The fuel-driven process of replication in living systems generates distributions of copied entities with varying degrees of copying accuracy. Here we introduce a thermodynamically consistent ensemble for investigating universal population features of template copying systems. In the context of copolymer copying, coarse-graining over molecular details, we establish a phase diagram of copying accuracy. We discover sharp non-equilibrium transitions between populations of random and accurate copies. Maintaining a population of accurate copies requires a minimum energy expenditure that depends on the configurational entropy of copolymer sequences.

DOI: 10.1103/PhysRevLett.134.068402

Introduction—The ability to replicate is a hallmark of the living world. Organisms can replicate themselves, as well as cells, and DNA replication, RNA transcription and RNA translation to proteins are examples of polymer template copying [1]. Generating a DNA polymer copy with near-identical sequence to the template DNA requires energy consumption [2,3]. DNA replication is catalyzed by the molecule DNA polymerase, which progressively moves along the template DNA strand as it generates a polymer copy [4]. Generating copies of DNA competes with DNA disassembly, catalyzed for example by DNAses without involvement of a fuel [5]. Detailed models have been used to discuss the key properties of this copy process, typically focusing on individual copies of a template sequence [6–17]. This provided insights onto the fundamental limits and trade-offs associated with template copying. Examples of this include trade-offs (or absence thereof) [11,15,18,19] and correlations [14] between speed, accuracy and cost of copying; links between dissipation, elongation and information transmission [6,8]; and definitions of copying efficiency [13].

Here we investigate the conditions for establishing whole populations of accurate copies of a copolymer template. In order to focus on generic features, we coarse-grain molecular details of copying such as sequential steps of initiation [16], polymer elongation [6,9] and strand separation [13,17] into a one-step stochastic process. We define what we call a template copying ensemble where a single template in presence of reservoirs of fuel and monomers generates a population of stochastic copies. We study the distribution of copying errors as a function of copying specificity and active driving by the fuel. We establish a phase diagram of copying accuracy for the template copying ensemble, and discover sharp transitions between populations of random and accurate copies in the limit of long polymers. Our template copying ensemble allows for a thermodynamic description of nonequilibrium steady-state populations of accurate and random polymer copies.

Template copying ensemble—We consider a system containing one polymer template sequence T of length L in contact with four reservoirs: a pool of monomers of m different types M_i with $i = 1 \ldots m$ ($m = 4$ for DNA and RNA), a bath of fuel molecules F, a bath of waste molecules W, and a heat bath at constant temperature (we measure energy in units of the thermal energy, $k_B T = 1$). We call the setting the template copying ensemble. Sequences S_j with $j = 1 \ldots m^L$ can be generated by copying the template sequence in a process we refer to as templated assembly. The templated assembly process consumes fuel F and generates waste W, to produce copy sequences S_j of the same length as the template without altering the template:

$$\sum_{i=1}^{m} n_{ij} \mathrm{M}_i + \mathrm{T} + L\mathrm{F} \underset{k_j^-}{\overset{k_j^+}{\rightleftharpoons}} \mathrm{S}_j + \mathrm{T} + L\mathrm{W}. \tag{1}$$

Here n_{ij} is a stoichiometric coefficient describing the number of monomers of type i in sequence S_j, with

*Contact author: agenthon@pks.mpg.de
†Contact author: modes@mpi-cbg.de
‡Contact author: julicher@pks.mpg.de
§Contact author: grill@mpi-cbg.de

$\sum_{i=1}^{m} n_{ij} = L$. The templated assembly leading to sequence S_j occurs at a rate k_j^+ and we consider that template availability is not limiting. Copies are error free if $k_j^+ = 0$ for all sequences j that differ from the template, $S_j \neq T$. Copying errors are captured by finite rates $k_j^+ > 0$ for these sequences. Microscopic reversibility implies that for each j the reverse pathway, which we call templated disassembly, also exists with rate k_j^-. However, one may expect spontaneous disassembly at rate k_r^- to be more frequent:

$$S_j \underset{k_r^+}{\overset{k_r^-}{\rightleftharpoons}} \sum_{i=1}^{m} n_{ij} M_i, \tag{2}$$

with k_r^+ denoting a spontaneous assembly rate.

Microreversibility requires that the templated assembly and disassembly rates obey $k_j^+/k_j^- = e^{-(\Delta\mu_r - \Delta\mu_F)L}$, where $\Delta\mu_r = \epsilon_S/L - \mu_M$ is the per-monomer energy associated with the assembly of a single polymer, and $\Delta\mu_F = \mu_F - \mu_W > 0$ is the per-monomer Gibbs free energy provided by the fuel. Because $\Delta\mu_r$ is independent of the template, its dependence on sequence S_j cannot be used to generate accurate copies of the template [12]. The template behaves as a catalyst and kinetic rates and energy barriers depend on template sequence. We therefore choose $\Delta\mu_r$ to be independent of sequence S_j. We write the rates as $k_j^+ = k_j e^{-(\Delta\mu_r - \Delta\mu_F)L}$, and for the reverse rate $k_j^- = k_j$. Sequence dependence of the process enters via the kinetic coefficients k_j according to $k_j = k_0 e^{-aq}$ [20], where $q \leq L$ is the number of incorrectly copied monomers (the Hamming distance between T and S_j), k_0 is a rate prefactor, and parameter a a specificity.

The spontaneous disassembly and assembly rates also obey $k_r^+/k_r^- = e^{-\Delta\mu_r L}$. We write for the rate of spontaneous assembly $k_r^+ = k_r e^{-\Delta\mu_r L}$ and for the rate of spontaneous disassembly $k_r^- = k_r$, with a sequence-independent coefficient k_r.

Note that for now the forward rates k_j^+ and k_r^+ depend on energetics but the backward rates k_j^- and k_r^- do not. In this case, the rate of templated disassembly $k_j^- = k_0 e^{-aq}$ vanishes for large q while the rate of spontaneous disassembly k_r^- is constant. In a later section we will relax this assumption.

A schematic representation of the system and reservoirs is provided in Fig. 1. Our coarse-grained model describes copying as a one-step process. Polymers of length different from L could occur as intermediate states but are not considered at the coarse-grained level.

Statistics of copying errors—We next determine the probability distribution $p(N_S, t)$ to have N_S copies of sequence S at time t which obeys

$$\partial_t p(N_S, t) = k_a p(N_S - 1, t) - (k_a + N_S k_d) p(N_S, t) + (N_S + 1) k_d p(N_S + 1, t), \tag{3}$$

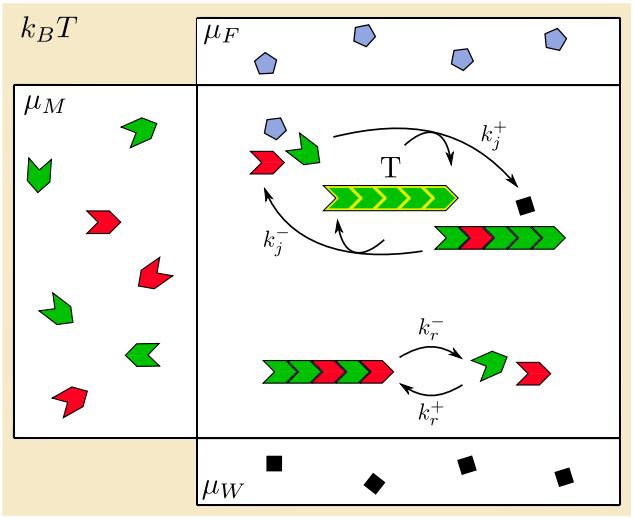

FIG. 1. Schematic of the template copying ensemble with two monomer types ($m = 2$): green and red. In this example, the template T is composed of green monomers, thus red monomers in sequences S_j are incorrectly copied.

$$\partial_t p(0, t) = -k_a p(0, t) + k_d p(1, t), \tag{4}$$

with the total assembly rate $k_a = k_j^+ + k_r^+$ and the total disassembly rate $k_d = k_j^- + k_r^-$. We choose as initial condition $p(N_S, 0) = \delta_{N_s}$, yielding a Poisson distribution $p(N_S, t) = \lambda_q^{N_S} e^{-\lambda_q}/N_S!$ for all times, with $\lambda_q = k_a/k_d[1 - \exp(-k_d t)]$ [20]. The expected number of copies $\langle N_x \rangle$ with a monomer error fraction $x = q/L$ is $\langle N_x \rangle = \lambda_{xL} \Omega_{xL}$ with $\Omega_q = \binom{L}{q}(m-1)^q$ the number of sequences with q wrong monomers. For sufficiently long polymers $L \gg 1$, $\langle N_x \rangle$ is dominated by either random copies with error fraction x_r

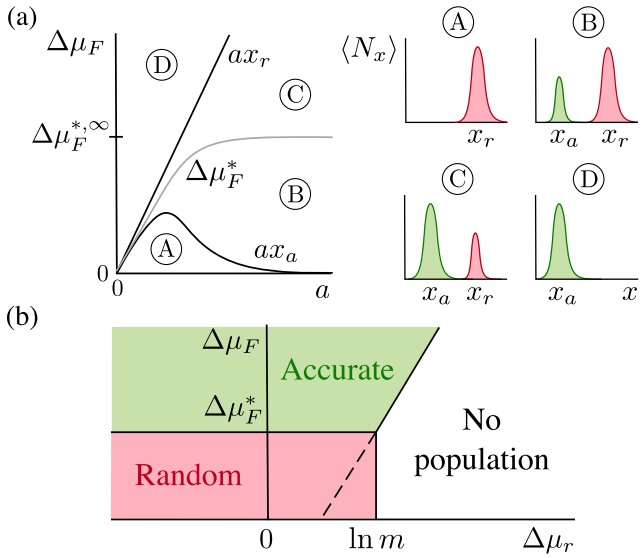

FIG. 2. Schematic phase diagrams for (a) finite L and (b) large L with a fixed value of specificity a. (a) $\Delta\mu_F^{*,\infty} = \ln m + O(L^{-1})$ is the asymptotic value of $\Delta\mu_F^*$ for large specificity a.

or by accurate copies with error fraction x_a, or both [see Fig. 2(a)]. The average fraction of copying errors $\bar{x} = \sum_x x\langle N_x\rangle / \sum_x \langle N_x\rangle$ is used as a measure of copying accuracy. At first order in $1/L$, the error frequencies x_i with index $i \in \{r, a\}$ are given by $x_i = x_i^{(0)} - (1 - 2x_i^{(0)})/(2L) + o(L^{-1})$ with

$$x_r^{(0)} = \frac{m-1}{m}, \tag{5}$$

$$x_a^{(0)} = \frac{1}{1 + e^a/(m-1)}. \tag{6}$$

When $a \to 0$ the templated copying process becomes nonspecific and $x_a^{(0)} \to x_r^{(0)}$. When $a \to +\infty$ the templated copying process becomes precise and $x_a^{(0)} \to 0$.

The distribution of copying errors $\langle N_x\rangle$ depends on the Gibbs free energy provided by the fuel $\Delta\mu_F$. We next consider the copying errors in steady state. If $\Delta\mu_F \leq a x_a$ the number of accurate copies is small, and random sequences dominate. If $\Delta\mu_F \geq a x_r$, the number of random copies is small, and the templated assembly process dominates. Under both conditions $\langle N_x\rangle$ is unimodal with a maximum at $x = x_r$ and $x = x_a$, respectively. These modes correspond to the dominant error fraction of the sequences assembled via the spontaneous and templated processes, respectively. If instead $a x_a < \Delta\mu_F < a x_r$, random and accurate copies coexist, and $\langle N_x\rangle$ is bimodal. In this regime, the number of random copies $\langle N_{x_r}\rangle$ and the number of accurate copies $\langle N_{x_a}\rangle$ are equal when $\Delta\mu_F = \Delta\mu_F^*$, with

$$\Delta\mu_F^* = \ln\left(\frac{m}{1 + (m-1)e^{-a}}\right) + \frac{1}{L}\ln\left(\frac{k_r}{k_0}\sqrt{\frac{x_a^{(0)}\left(1 - x_a^{(0)}\right)}{x_r^{(0)}\left(1 - x_r^{(0)}\right)}}\right) + o\left(\frac{1}{L}\right). \tag{7}$$

These cases are represented in Fig. 2(a). The dependence of $\Delta\mu_F^*$ on specificity a is shown in Fig. 2(a) in gray.

Phase transition—When increasing $\Delta\mu_F$, we observe a smooth transition of the average error fraction $\bar{x}$ from random copies with error fraction x_r (regions A and B) to accurate copies with error fraction x_a (regions C and D). Since the difference between $\langle N_{x_a}\rangle$ and $\langle N_{x_r}\rangle$ grows exponentially with L in regions B and C, this transition becomes sharp at $\Delta\mu_F = \Delta\mu_F^*$ when $L \to \infty$, as illustrated in Fig. 3(a). Therefore, for large L, accurate copies dominate and the average error fraction becomes x_a when $\Delta\mu_F > \Delta\mu_F^*$ where $\Delta\mu_F^*$ depends on specificity a and number of monomer types m. For any value of the specificity, accurate copies dominate in the large L limit when

$$\Delta\mu_F \geq \ln m. \tag{8}$$

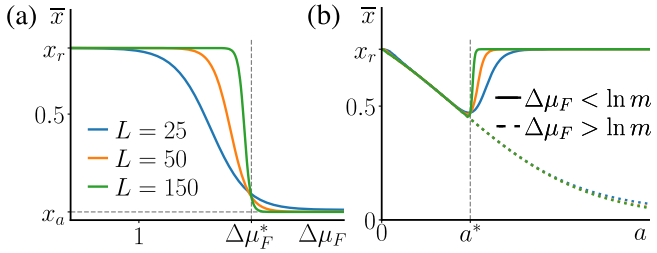

FIG. 3. Average error fraction $\bar{x}$ vs (a) energy drive $\Delta\mu_F$ and (b) specificity a, for different values of the template length L [colors apply to (a) and (b)]. (a),(b) $m = 4$, $k_0 = 1$, $k_r = 0.1$ and $\Delta\mu_r = 0.5$. (a) $a = 3$ and $\Delta\mu_F^* \approx 1.25$. (b) $\Delta\mu_F = 2 > \ln 4$ (dotted lines), $\Delta\mu_F = 0.8 < \ln 4$ (plain lines) and $a^* \approx 1.33$. A 3D plot is shown in [20].

This condition can be interpreted as a Landauer's principle [24,25] for polymer copying: to copy information accurately, the per-monomer externally provided free energy must be larger than the per-monomer entropy of configuration of the polymer $\ln(m^L)/L$ [12]. We refer to $\Delta\mu_F = \ln m$ as the Landauer limit.

Below the Landauer limit ($\Delta\mu_F < \ln m$) the average error fraction $\bar{x}$ decreases as specificity a is increased at low specificity. If specificity a crosses from below a threshold a^* which depends on $\Delta\mu_F$ according to Eq. (7), $\bar{x}$ jumps to the value x_r. For finite L, the transition becomes smooth [see Fig. 3(b)]. This transition occurs because increasing the specificity a reduces the error fraction x_a associated with accurate copies, but also slows down the templated assembly of these copies. Because the kinetics of the spontaneous process is independent of a, random copies will eventually dominate. Above the Landauer limit ($\Delta\mu_F \geq \ln m$) the average error fraction $\bar{x}$ decreases as specificity a is increased, but remains accurate and does not undergo a transition towards random copies.

In the large L limit, the transition from random to accurate copies is a first order phase transition. The gray line shown in Fig. 2(a) becomes a first order phase transition line in the limit of large L.

Population size—We next ask if the number of copies that participate in this phase transition can vanish in the large L limit. The number of accurate and random copies, $\langle N_{x_a}\rangle$ and $\langle N_{x_r}\rangle$, depend on the competition between the energetic and entropic contributions: $\ln\langle N_x\rangle/L = \ln\lambda_{xL}/L + \ln\Omega_{xL}/L$. If $\Delta\mu_r > \ln m$, the population of random copies $\langle N_{x_r}\rangle$ goes extinct for large L, with the energy difference $\ln\lambda_{x_rL}/L = -\Delta\mu_r$ and the entropy of configuration $\ln\Omega_{x_rL}/L = \ln m$. Similarly, if $\Delta\mu_F < \Delta\mu_r - \ln[1 + (m-1)e^{-a}]$, the population of accurate copies $\langle N_{x_a}\rangle$ vanishes, with $\ln\lambda_{x_aL}/L = -(\Delta\mu_r - \Delta\mu_F) - a x_a$ and $\ln\Omega_{x_aL}/L = \ln[1 + (m-1)e^{-a}] + a x_a$ [20].

We can now draw a phase diagram in the large L limit as a function of $\Delta\mu_F$ and $\Delta\mu_r$, for a given specificity a, see Fig. 2(b). This diagram contains three regions: a region

where accurate copies dominate, a region where random copies dominate, and a region where the population vanishes. The two boundary lines of the region of vanishing population are given by the two conditions of extinction discussed above. The boundary line between random and accurate copies occurs at $\Delta\mu_F = \Delta\mu_F^*$ where $\Delta\mu_F^* = \ln(m/[1+(m-1)e^{-a}])$ in the large L limit. The three regions meet at a triple point $(\Delta\mu_r = \ln m, \Delta\mu_F = \Delta\mu_F^*)$.

Non equilibrium current—We now investigate the nonequilibrium nature of the phase diagrams discussed above. In steady state, total assembly and disassembly are balanced for any sequence S: $k_a = k_d\langle N_S\rangle$. However, the templated and spontaneous processes are not balanced individually, which is associated with a nonzero net average fuel current from the fuel bath to the waste bath $\langle J\rangle = L\sum_{j=1}^{mL}[k_j^+ - \langle N_{S_j}\rangle k_j^-]$. In the large L limit $\langle J\rangle \sim Lk_0\exp[-(\Delta\mu_r - \Delta\mu_F - \ln(1+(m-1)e^{-a}))L]$ [20], where $\sim$ describes asymptotic equality in the large L limit.

We now distinguish three regions of the phase diagram which differ in the transduction of fuel energy into useful information. When $\Delta\mu_F < \Delta\mu_r - \ln[1+(m-1)e^{-a}]$, the fuel current vanishes in the large L limit. This region is delimited by the tilted line (both dashed and solid) in Fig. 2(b). In this case, no fuel is consumed and no accurate copies are produced. The other two regions are located above this titled line, where the nonvanishing fuel current maintains the system in a nonequilibrium steady state, and are delimited by the horizontal line in Fig. 2(b). If $\Delta\mu_F > \ln(m/[1+(m-1)e^{-a}])$, accurate copies dominate so fuel energy is efficiently converted into information, with $\langle J\rangle \sim Lk_r\langle N_{x_a}\rangle$. If $\Delta\mu_F < \ln(m/[1+(m-1)e^{-a}])$, random copies dominate in the large L limit. In this case, fuel is burnt but no useful information is transmitted.

Kinetic proofreading—Fuel-driven error-correction mechanisms could increase copying accuracy by modifying the kinetics. For example, kinetic proofreading feeds on fuel energy to undo copy errors at the expense of slowing the copy process [18,26,27]. In our description of template copying, assembly and disassembly kinetics depend on sequence, but energy differences $\Delta\mu_r$ and $\Delta\mu_F$ do not. Hence fuel-driven error-correction mechanisms could be implicitly accounted for in our model.

Proofreading can take different pathways that can be associated with different amounts of fuel consumption. Hence, when explicitly coarse-graining over these pathways the microreversibility condition is broken [28] but effective kinetic rates can still be defined [20]. In general, these rates are more complex than those introduced in the text after Eq. (2). In some scenarios however, for example error correction by single state backtracking [27], an effective kinetic prefactor, an effective specificity and an effective fuel free energy difference can be identified [20]. In such cases the error fraction x_a of accurate copies will decrease, which in turn shifts the boundaries ax_r, $\Delta\mu_F^*$, and ax_a of the phase diagram [Fig. 2(a)], since these depend on a.

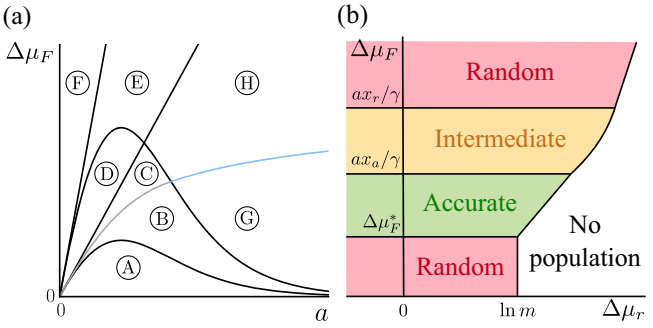

FIG. 4. Schematic phase diagrams for the generalized reaction rates and $\gamma > 0$, for (a) finite L and (b) large L with a fixed value of a such that $\ln(m/[1+(m-1)e^{-a}])/(1+\gamma) < ax_a/\gamma$.

Generalized reaction rates—We now relax the assumption that the backward rates do not depend on the energies by introducing a fuel-dependent energy barrier in the templated rates: $k_j^+ = k_j e^{-[\Delta\mu_r-(1+\gamma)\Delta\mu_F]L}$ and $k_j^- = k_j e^{\gamma\Delta\mu_F L}$ with $\gamma\in\mathbb{R}$. This changes the ratio of the timescales associated with the templated and the spontaneous process in an L-dependent manner.

We show in Fig. 4 the finite and large L phase diagrams for $\gamma > 0$, which speeds up the kinetics of the templated process compared to $\gamma = 0$, and allows templated disassembly to dominate over spontaneous disassembly for a range of error fractions x. The regions A–H and the extinction conditions are discussed in [20]. Similar to the example shown in Fig. 2(b), for small $\Delta\mu_F$, the system generates random copies. The transition from random to accurate copies occurs at the threshold $\Delta\mu_F^*$, which is reduced by a factor $1+\gamma$ relative to the value given in Eq. (7). Further increasing $\Delta\mu_F$ leads to a continuous increase in the average error fraction $\bar{x} = \gamma\Delta\mu_F/a$ (intermediate region) until the system reenters a regime of random copying. This is because both assembly and disassembly are dominated by the templated process in this regime, which results in an equilibrium distribution of random copies. There is thus a range of values $\Delta\mu_F^* < \Delta\mu_F < ax_a/\gamma$ where a maximum accuracy at error fraction x_a is achieved.

Discussion—In this Letter we study populations of copolymer copies and their accuracy in a thermodynamically consistent template copying ensemble. We find sharp transitions between populations of random and accurate copies as a function of fuel driving and copying specificity. Our coarse-grained approach reveals generic features of stochastic copying processes that are independent of many molecular details. It allows for an evaluation of the role of fuel driving, proofreading, and energy barriers on the population of accurate copies. We identify for given specificity the minimal cost of Gibbs free energy $\Delta\mu_F^*$ required to maintain a population of accurate copies. The minimal Gibbs free energy cost to be in a regime of accurate copying regardless of the specificity is given by

the per monomer configurational entropy. In analogy to the Landauer principle of information erasure we refer to this as the Landauer limit.

It will be interesting to extend this framework in several directions. First, by allowing copies to serve as template themselves in an autocatalytic manner, perhaps offering new means to investigate the link between the statistical mechanics of replicating systems and the evolutionary process [29–32]. This might also allow for a reinvestigation of Eigen's paradox of achieving high fidelity copies of large genomes [33,34]. More generally, other problems associated with the origin-of-life question could perhaps be freshly investigated within the template copying ensemble [35–41].

Second, the free energy differences $L\Delta\mu_r$ and $L\Delta\mu_F$ might not scale with template length L and the specificity a may depend on the error fraction x, as observed for example in kinetic proofreading [20]. Hence, the kinetic rates defined in this work are the first order terms in the Taylor expansions of more general rates. In principle, next order terms could be systematically determined in future work.

Third, in the one-step copy process discussed above intermediate stages of polymer growth are not represented. It is an interesting direction of future work to refine our description and take into account such intermediate states.

It will be interesting to test our results in experiments. For example, template copying could be realized with *in vitro* transcription assays where a DNA template is used to generate RNA copies [42]. Furthermore, polymerase-chain-reaction (PCR) experiments using primers that attach to only one of the two strands of the DNA template would generate complementary DNA copies of the template strand and provide another example of template copying [43]. Performing such DNA copying experiments and measuring distributions of error fractions in the produced sequences could provide experimental tests of our results and infer in which regimes these processes operate.

While inspired by the copy process of biological DNA polymers [20], our formalism more generally applies to systems of information transfer from a template to a copy of the template. For example, our approach equally applies to the copy process from RNA to protein (translation, with 20 "monomer" types), or more generally for copy processes with any number of monomer types. Our formalism allows for the discussion of a trade-off that arises when increasing the number m of monomer types, for large L. The energetic cost required to be in a regime of accurate copying is high for low values of m as compared to large values of m, but the accuracy of copies is higher. This is because an increase of m allows for the use of shorter polymers for encoding the same amount of information $\Omega = m^L$, while the more complex encoding is more prone to mistakes [Eq. (6)]. Indeed, the minimal energy to copy a sequence with information content Ω with specificity a is given by $E^*_{\mathrm{tot}} = L(m)\Delta\mu^*_F \sim (1 - \ln[1 + (m-1)e^{-a}]/\ln m)\ln\Omega$, which decreases with m. It is interesting to speculate that this cost-accuracy trade-off is relevant from an evolutionary point of view, manifested in the choice of $m = 4$ for maintaining the genome in the DNA form at high fidelity, and for copying genomic information to protein peptide sequences with $m = 20$ at lower energetic costs and reduced requirements on fidelity.

Acknowledgments—We thank B. Qureshi and T. E. Ouldridge for a critical reading of our manuscript and for providing a preprint of their work [19].

[1] A. Kornberg and T. A. Baker, *DNA Replication*, 2nd ed. (W. H. Freeman and Company, New York, 1992).
[2] I. Lieberman, A. Kornberg, and E. S. Simms, Enzymatic synthesis of nucleoside diphosphates and triphosphates, J. Biol. Chem. **215**, 429 (1955).
[3] G. T. Yarranton and M. L. Gefter, Enzyme-catalyzed DNA unwinding: Studies on *Escherichia coli rep* protein, Proc. Natl. Acad. Sci. U.S.A. **76**, 1658 (1979).
[4] I. Lehman, M. J. Bessman, E. S. Simms, and A. Kornberg, Enzymatic synthesis of deoxyribonucleic acid, J. Biol. Chem. **233**, 163 (1958).
[5] W. Yang, Nucleases: Diversity of structure, function and mechanism, Q. Rev. Biophys. **44**, 1 (2011).
[6] D. Andrieux and P. Gaspard, Nonequilibrium generation of information in copolymerization processes, Proc. Natl. Acad. Sci. U.S.A. **105**, 9516 (2008).
[7] D. Andrieux and P. Gaspard, Molecular information processing in nonequilibrium copolymerizations, J. Chem. Phys. **130**, 014901 (2009).
[8] P. Sartori and S. Pigolotti, Kinetic versus energetic discrimination in biological copying, Phys. Rev. Lett. **110**, 188101 (2013).
[9] P. Gaspard and D. Andrieux, Kinetics and thermodynamics of first-order Markov chain copolymerization, J. Chem. Phys. **141**, 044908 (2014).
[10] P. Sartori and S. Pigolotti, Thermodynamics of error correction, Phys. Rev. X **5**, 041039 (2015).
[11] R. Rao and L. Peliti, Thermodynamics of accuracy in kinetic proofreading: Dissipation and efficiency trade-offs, J. Stat. Mech. (2015) P06001.
[12] T. E. Ouldridge and P. R. ten Wolde, Fundamental costs in the production and destruction of persistent polymer copies, Phys. Rev. Lett. **118**, 158103 (2017).
[13] J. M. Poulton, P. R. ten Wolde, and T. E. Ouldridge, Nonequilibrium correlations in minimal dynamical models of polymer copying, Proc. Natl. Acad. Sci. U.S.A. **116**, 1946 (2019).
[14] D. Chiuchiú, Y. Tu, and S. Pigolotti, Error-speed correlations in biopolymer synthesis, Phys. Rev. Lett. **123**, 038101 (2019).
[15] M. Sahoo, Arsha N., P. R. Baral, and S. Klumpp, Accuracy and speed of elongation in a minimal model of DNA replication, Phys. Rev. E **104**, 034417 (2021).

[16] J. M. Poulton and T. E. Ouldridge, Edge-effects dominate copying thermodynamics for finite-length molecular oligomers, New J. Phys. **23**, 063061 (2021).
[17] J. Juritz, J. M. Poulton, and T. E. Ouldridge, Minimal mechanism for cyclic templating of length-controlled copolymers under isothermal conditions, J. Chem. Phys. **156**, 074103 (2022).
[18] C. H. Bennett, Dissipation-error tradeoff in proofreading, BioSystems **11**, 85 (1979).
[19] We would like to point the reader to the preprint B. Qureshi, J. M. Poulton, and T. E. Ouldridge, arXiv:2404.02791v1 that became available during the review stage of this manuscript, which discusses thermodynamic limits on copying accuracy.
[20] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.134.068402 for additional derivations, which includes Refs. [21–23].
[21] M. F. Goodman, S. Creighton, L. B. Bloom, J. Petruska, and T. A. Kunkel, Biochemical basis of DNA replication fidelity, Crit. Rev. Biochem. Mol. Biol. **28**, 83 (1993).
[22] M. J. Thomas, A. A. Platas, and D. K. Hawley, Transcriptional fidelity and proofreading by RNA polymerase II, Cell **93**, 627 (1998).
[23] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, J. Phys. Chem. **81**, 2340 (1977).
[24] R. Landauer, Irreversibility and heat generation in the computing process, IBM J. Res. Dev. **5**, 183 (1961).
[25] D. Andrieux and P. Gaspard, Information erasure in copolymers, Europhys. Lett. **103**, 30004 (2013).
[26] J. J. Hopfield, Kinetic proofreading: A new mechanism for reducing errors in biosynthetic processes requiring high specificity, Proc. Natl. Acad. Sci. U.S.A. **71**, 4135 (1974).
[27] M. Depken, J. M. Parrondo, and S. W. Grill, Intermittent transcription dynamics for the rapid production of long transcripts of high fidelity, Cell Rep. **5**, 521 (2013).
[28] T. E. Ouldridge, C. C. Govern, and P. R. ten Wolde, Thermodynamics of computational copying in biochemical systems, Phys. Rev. X **7**, 021004 (2017).
[29] N. Goldenfeld and C. Woese, Life is physics: Evolution as a collective phenomenon far from equilibrium, Annu. Rev. Condens. Matter Phys. **2**, 375 (2011).
[30] L. A. Demetrius, Boltzmann, Darwin and directionality theory, Phys. Rep. **530**, 1 (2013).
[31] J. L. England, Statistical physics of self-replication, J. Chem. Phys. **139**, 121923 (2013).
[32] A. Genthon, R. García-García, and D. Lacoste, Branching processes with resetting as a model for cell division, J. Phys. A **55**, 074001 (2022).
[33] M. Eigen, Selforganization of matter and the evolution of biological macromolecules, Naturwissenschaften **58**, 465 (1971).
[34] I. Leuthäusser, Statistical mechanics of Eigen's evolution model, J. Stat. Phys. **48**, 343 (1987).
[35] A. V. Tkachenko and S. Maslov, Spontaneous emergence of autocatalytic information-coding polymers, J. Chem. Phys. **143**, 045102 (2015).
[36] Y. J. Matsubara and K. Kaneko, Optimal size for emergence of self-replicating polymer system, Phys. Rev. E **93**, 032503 (2016).
[37] Y. J. Matsubara and K. Kaneko, Kinetic selection of template polymer with complex sequences, Phys. Rev. Lett. **121**, 118101 (2018).
[38] A. V. Tkachenko and S. Maslov, Onset of natural selection in populations of autocatalytic heteropolymers, J. Chem. Phys. **149**, 134901 (2018).
[39] J. H. Rosenberger, T. Göppel, P. W. Kudella, D. Braun, U. Gerland, and B. Altaner, Self-assembly of informational polymers by templated ligation, Phys. Rev. X **11**, 031055 (2021).
[40] P. W. Kudella, A. V. Tkachenko, A. Salditt, S. Maslov, and D. Braun, Structured sequences emerge from random pool when replicated by templated ligation, Proc. Natl. Acad. Sci. U.S.A. **118**, e2018830118 (2021).
[41] T. Göppel, J. H. Rosenberger, B. Altaner, and U. Gerland, Thermodynamic and kinetic sequence selection in enzyme-free polymer self-assembly inside a non-equilibrium RNA reactor, Life **12**, 567 (2022).
[42] R. Losick, *In vitro* transcription, Annu. Rev. Biochem. **41**, 409 (1972).
[43] K. Mullis, F. Faloona, S. Scharf, R. Saiki, G. Horn, and H. Elrich, Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction, Cold Spring Harbor Symp. Quant. Biol. **51**, 263 (1986).