


Role of thermal fluctuations in biological copying mechanismsMoupriya Das^{✉*} and Holger Kantz[†]*Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Strasse 38, 01187 Dresden, Germany* (Received 22 July 2020; revised 16 December 2020; accepted 18 February 2021; published 10 March 2021)

During transcription, translation, or self-replication of DNA or RNA, information is transferred to the newly formed species from its predecessor. These processes can be interpreted as (generalized) biological copying mechanism as the new biological entities like DNA, RNA, or proteins are representing the information of their parent bodies uniquely. The accuracy of these copying processes is essential, since errors in the copied code can reduce the functionality of the next generation. Such errors might result from perturbations on these processes. Most important in this context is the temperature of the medium, i.e., thermal noise. Although a reasonable amount of experimental studies have been carried out on this important issue, theoretical understanding is truly sparse. In the present work, we illustrate a model study which is able to focus on the effect of the temperature on the process of biological copying mechanisms, as well as on mutation. We find for our paradigmatic models that, in a quite general scenario, the copying processes are most accurate at an intermediate temperature range; i.e., there exists an optimum temperature where mutation is most unlikely. This allows us to interpret the observations for some biological species with the aid of our model study.

DOI: [10.1103/PhysRevE.103.032110](https://doi.org/10.1103/PhysRevE.103.032110)**I. INTRODUCTION**

During some essential processes in biology, such as DNA or RNA replication, transcription, translation, synthesis of DNA from RNA, biological information is transferred from one biological entity to its successor. As it is commonly accepted, the sequence of amino acids of DNA and RNA encodes the information for all building blocks of living matter. More precisely, DNA or RNA strands work as templates and the base sequence in the template strand provides signals to synthesize the desired product. For example, in the course of the DNA or RNA replication process a copy of the parent DNA or RNA is formed by following the sequence of bases in the originator strand. In another mechanism, known as transcription, messenger RNA (mRNA) is produced by reading the information about the arrangement of bases from a particular segment of the template DNA strand. The biological information stored in the newly synthesized mRNA is utilized to generate the protein in the cell during the translation process. Transcription and translation are two fundamental steps for protein synthesis in living organisms and these two processes together are termed as gene expression. During all of these functions, biological information flows from the initiator biological macromolecule to the product body. Therefore, these processes can be seen as copying of biological

information [1–6]. The characteristics of the lineage of species and proper biological functions in a cell which also affects the nature of the next generations, depend decisively on the correctness of these biological copying mechanisms. For example, DNA replication during cell division is an essential part of the growth and reproduction of a biological species. Also, the appropriate proteins need to be synthesized in the cells to properly regulate cell divisions and growth of the organisms. Both of these processes, DNA replication and protein synthesis, involve copying of information from the parent body to its successor.

Error in the copying process of DNA results in a wrong sequence of information that would lead to mutation, i.e., an alteration in the appropriate genetic arrangement of the species [7]. The study of mutation is a very active area of research since the introduction of the science of evolution [8] and the idea of natural selection [9] in early 19th Century and has immense implication in contemporary science [10–18] as it constructs the framework of the basic understanding of the evolution of species. Mutation can be both harmful and beneficial [19–27]. If the mutation is harmful, then the mutated species has less chance to survive. However, beneficial mutation in organisms is supposed to favor the survival of the species in the present environment. However, beneficial mutations might be lost or harmful mutations might be rectified because of the random fluctuations associated with the process. Nevertheless, mutation contributes to the evolution of species to fit in the current environment. Therefore, it is very crucial to understand the effect of the surroundings on mutation, or more accurately on the biological copying processes, the error in which causes mutation. Although there are extensive studies on mutation in diverse directions, the influence of the environmental parameters on mutation is comparatively less explored [28–30]. While a fair extent of attention has

*Corresponding author: mdas@pks.mpg.de†kantz@pks.mpg.de

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Open access publication funded by the Max Planck Society.

been drawn towards the experimental studies [31–39] on this subject, theoretical investigations are essentially countable [24–26]. Here, in the present work we address this issue to establish a theoretical interpretation of this important matter.

Since all chemical reaction rates are temperature dependent, as well as the biochemical processes in a more general sense [40], we presume that the temperature is an elementary factor that can directly impact mutation. Moreover, the molecular machinery which performs the copying processes lives on the nanoscale: Individual amino-acids are transported and get attached to the strand, where molecular motors act stepwise. This whole process is stochastic, and the intensity of the noise is also related to the environmental temperature. Actually, for purely mechanical nanosystems, the theory of stochastic thermodynamics [41] has been developed, which allows us to describe such isothermal nonequilibrium systems.

Here, we aim to analyze the effect of temperature on mutations, or more precisely, on the process of biological copying. We propose a generalized scheme as a model for such biological copying processes and study the correctness or error in the procedures as a function of temperature. Errors in the copying process, if it is for DNA, cause mutations in the species. It has been found that defects in translation can also give rise to mutations [42]. Therefore, at the end of Sec. III B, we compare the temperature dependence of our model for the copying process to empirical findings on the mutation rates in real biological species.

In real biological systems, there occurs another additional step along with replication of the biological memory, which is known as the proofreading procedure [43–53]. This step reduces the error in the copying processes. Here, we disregard this error-correcting step [2] and only concentrate on the efficiency of the biological copying process as a function of temperature. Our main result to be presented below is that its error rate depends in a nontrivial nonlinear way on temperature. The success rate of the biological copying processes exhibits a maximum at an intermediate value of temperature, i.e., there is an optimum temperature for maximal accuracy and hence minimal mutation rate.

II. TWO MODEL SYSTEMS AND THEIR DYNAMICS

For convenience, we will in the following only consider the replication or transcription of DNA, hence we speak about biological information which is encoded by only four different symbols, namely, by the bases Adenine (A) or a Guanine (G) or a Cytosine (C) or a Thymine (T). Also for convenience, we have here suppressed the base Uracil (U) which might occur as an alternative to Thymine T with the same information content. Our models can be straightforwardly extended to the problem of translation, where 20 different codons (triples of these bases) encode the information for amino acids and hence the information needed for protein synthesis.

We propose two general models for biological copying processes to understand the effect of the environmental noise on these mechanisms. We summarize external noises under the notion of temperature, and hence investigate the influence of temperature on these essential biological functions. Both of our models will represent dynamical input-output relations. We interpret the sequence of bases which is being replicated

or transcribed or translated as an input signal (as they dictate the outcome of the process in a finite time) and the complementary base which comes and binds to the growing strand during the replication and transcription process and the amino acid which joins the elongated chain in case of the translation process to be the output of the given procedure. In the course of copying a long sequence, the system is out of equilibrium the whole time. These biological processes occur in systems of very small scale. Therefore, the role of thermal fluctuations becomes important to consider. Consequently, we propose stochastic dynamics for the output of the copying process, where we ignore all biochemical details of the microscopic processes.

A. Langevin model

In our first model, we use the dynamics of a Brownian particle in an appropriate potential to symbolize the output as a function of the input. The input signal is assumed to be a random sequence of the four symbols A, C, G, T which is a sufficiently good approximation to coding DNA strands: By eye, no regularities in the sequence of bases can be recognized, and all bases occur with roughly the same frequency [54]. To exploit this symmetry, we will replace the letters A, C, G, T by the index $i \in \{1, 2, 3, 4\}$.

We introduce a stochastic variable x which represents the output of the copying process. In other words, the value of x signifies the base which joins to the growing strand during the copying mechanism. The dynamics of the output x is modeled by the following over-damped Langevin equation in a potential landscape $V(x, S_i)$, in dimensionless form,

$$\dot{x} = -V'(x, S_i) + \sqrt{D}\xi(t). \quad (1)$$

The first term on the right-hand side of the Eq. (1) expresses the force derived from the deterministic potential $V(x, S_i)$, where S_i is a time dependent control parameter which eventually will be the input. Before we explain this in detail, we discuss the noise term $\xi(t)$. This represents zero-mean, Gaussian, white noise and obeys the fluctuation-dissipation relation. This noise term signifies the thermal fluctuations driving the dynamics. The properties of $\xi(t)$ are as follows:

$$\begin{aligned} \langle \xi(t) \rangle &= 0, \\ \langle \xi(t)\xi(t') \rangle &= 2\delta(t - t'). \end{aligned} \quad (2)$$

Equation (1) represents the over-damped Brownian dynamics in dimensionless form. In the course of making the actual Langevin dynamics dimensionless, the quantity D appears as a ratio of the temperature of the environment and a reference temperature [55]. Therefore, D stands for the scaled temperature of the system and is the measure of the strength of the thermal fluctuations.

The potential landscape $V(x, S_i)$ has as many minima as there are possible symbols for the output, i.e., four in the case of DNA replication or transcription. For fixed S_i , the invariant probability distribution, $\rho_0(x) \propto \exp[-V(x, S_i)/2D]$, has a maximum at each minimum of $V(x, S_i)$, so that under coarse-graining of x we can associate an output symbol to x . Here, the potential energy minima do not correlate to the free energies of the bases or the amino acids. Instead we impose

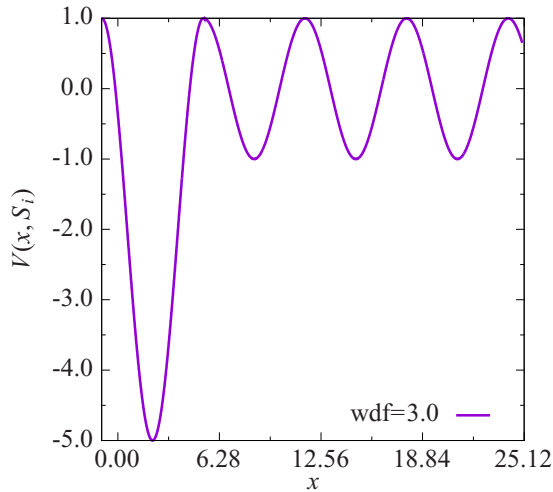


FIG. 1. The modified potential landscape for the replication or transcription process when the system reads the signal $i = 1$. This potential is considered in the Langevin model.

some symmetry in i : For given input S_i , there is one global deepest minimum with energy V_c which corresponds to the biologically correct output of the copying process, and all other minima with the depth V_w are shallower. Therefore, the probability to observe x in the “correct” state is higher than to find it in one of the “wrong” states. When the input signal changes, the previously deepest minimum becomes shallower and another minimum assumes the value V_c .

To be specific and to perform numerical simulations, we chose the following potential: We concatenate four sinusoidal minima on the real axis, where the i th minimum has the value $2a$ given by

$$V(x, S_i) = a \cos(x) - (a - 1), \quad x \in [2(i - 1)\pi, 2i\pi], \quad (3)$$

while the others have a depth of 2 given by

$$V(x, S_j) = \cos(x), \quad j \neq i, \quad x \in [(2(j - 1)\pi, 2j\pi]. \quad (4)$$

The parameter a determines the ratio of energies of the local and the global minimum. We call a the well-depth factor (wdf). The term $(a - 1)$ in Eq. (3) has to be subtracted to have a smooth potential. At $x = 0$ and $x = 8\pi$ we impose reflecting boundaries. We illustrate the potential for a given input signal $i = 1$ in Fig. 1. In Fig. 2, we translate this into the scenario of replication or transcription of DNA.

The last ingredient to our model is the pacing: We assume that the input string $\{S_i\}$ is shifted at time intervals of T_{step} by one step, i.e., at multiples of T_{step} the parameter S_i is replaced by S_{i+1} and the potential is changed correspondingly. Therefore, the process is out of equilibrium, and the motion of $x(t)$ is transient. To study this model, we will numerically integrate the Langevin dynamics Eq. (1) subject to a given input signal $\{S_i\}$ and find the time evolution of the stochastic variable $x(t)$. We let the system evolve for a considerably long time T_{step} compared to its relaxation time τ and define this time duration as the evolution time (t_{ev}). In the ideal situation, at time T_{step} , which we consider as a single step of the copying process, $x(t)$ has settled down in the deepest well and hence represents the correct output. A sequence of such steps mimics

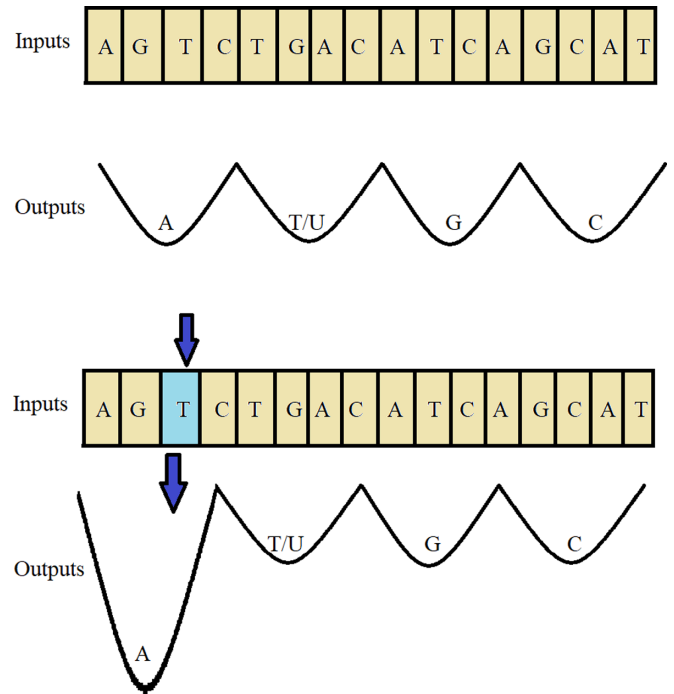


FIG. 2. The schematic diagram of the modification of the potential in the Langevin model representing replication or transcription. In this particular example the copying machinery detects signal T and the well corresponding to the correct output A gets lowered.

the process of replication or transcription or translation. The number of steps correlates to the amount of information that is being copied in the developing entity from the parent DNA or mRNA chain, or in other words the length of the predecessor template strand. Although the accurate copying of information implies a value of x in the appropriate well, x being a stochastic variable, there is a chance that it ends up in a wrong well. At the end of the each step, we examine whether $x(t)$ has a value that signifies the correct output well corresponding to the input signal S_i . This scheme stands for a direct stochastic algorithm that can be considered as a general representation of biological copying mechanisms. To examine whether the initial condition of x has any effect on the measurement, we consider random initial condition for the dynamical variable x , for an ensemble of trajectories, subject to random input signals and estimate the quantities of our interest taking an average over that ensemble.

We solve the over-damped Langevin dynamics Eq. (1) numerically using the improved Euler algorithm or Heun’s method [56]. This is essentially a second order Runge-Kutta method to solve ordinary differential equations with a noise term. In this method, two evaluations of the function is required at each step. The integration time step δt has been taken to be equal to 10^{-3} . The Gaussian white noise term has been generated by the Box-Muller algorithm.

A recent study [57] possesses some similarity in terms of the model as compared to our current research, so that we want to emphasize here the fundamental differences. The work by Chiuchiu *et al.* [57] considers heteropolymerization in terms of reaction coordinates which follow a Langevin dynamics similar to our Eq. (1). Their potential represents the

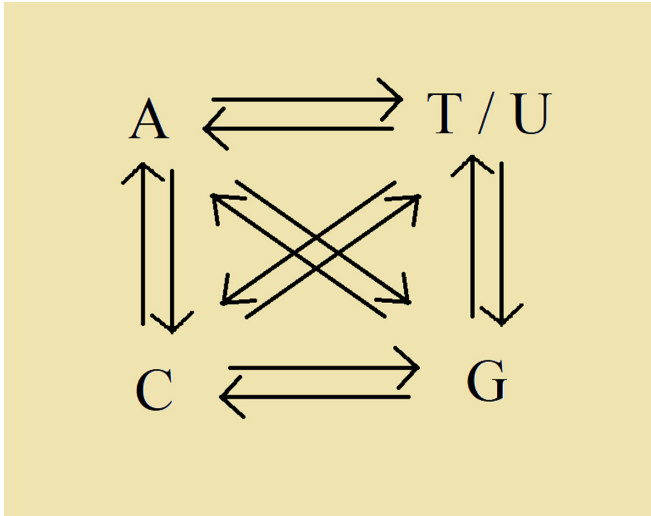


FIG. 3. Schematic diagram of the four-state model representing replication or transcription.

Gibbs free-energy landscape which is also a one-dimensional potential depicted in terms of the reaction coordinate. At the reaction network node, three reactions compete; incorporation of the right monomer, incorporation of the wrong monomer and removal of the lastly added monomer. The last step is not a part of the biological copying process and therefore not present in our model. And unlike in this paper, our coordinate x cannot be interpreted as a reaction coordinate but instead, after coarse-graining, signifies the output value of the copying process.

B. Four-state model

The above Langevin model is a one-dimensional model that is based on an energy landscape along a line. If the input flips from 1 to 4 and we assume that the particle was in the first well, then the trajectory has to pass through wells 2 and 3 to produce the correct output at the end of the next t_{ev} . To study whether this might have some undesired effects, we introduce a second model. This model has a discrete state variable which can be in one of four states, defined by the symbols A , T or U , G and C . The model is illustrated in Fig. 3.

The states again denote the “output” of the copying process, i.e., they indicate which base will be attached to the growing strand in the next time step. During this process, the free base pairs in the pool experience some sort of nonbonded attraction (van der Waals) towards the growing strand. So, for example, say A and G have some interaction energy with the growing strand. Although there is a difference in energy due to their difference in structure, we assume this to be so small that in the model there is no distinction. However, when the copying machinery detects the signal T , it develops a bonding interaction with A . So, there occurs a lowering of the potential well for A , while for G , the interaction still remains of nonbonded type. So, that energy gap remains unaltered. Hence, the proper interaction will have the lowest energy that will make escape from that state difficult and transition from other states to that state easier.

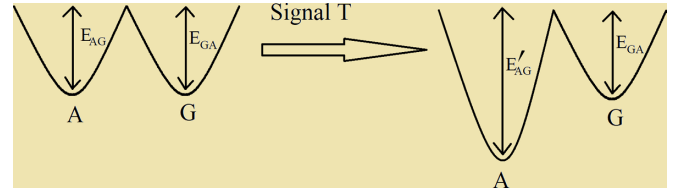


FIG. 4. Schematic diagram of the modification of the energy barriers between the states of a four-state model when the system receives a particular signal.

In our model dynamics, the susceptibility towards jumping from one state to another will be determined by the associated Kramers rate [58–61].

When no input signal is present, we consider that the rate of transition from one state to another is equal for all four states without any kind of bias. In this case, the new DNA strand will be a random sequence. However, the energy barriers get modified when the system reads a distinct input signal from the parent strand, since then the additional binding force to the complementary base of the parent strand favors the attachment of exactly one specific base. We assume that the potential well for the correct state is lowered by a factor b , whereas for simplicity we assume that it remains unchanged for the others. So b has a similar interpretation of the wdf as discussed in the case of the Langevin model. In Fig. 4 we illustrate this for DNA replication, where the system reads the input T and correspondingly lowers the potential well for state A . The potential well of state A being lowered, the transition from A to any other state becomes less probable, since the potential barrier to overcome is higher. Conversely, the probability to jump in between any two other states remains unchanged. So the trapping in the correct state is strengthened, but the escape from the “wrong” states is not.

Our model equations are then describing the rates of change of the probabilities of being in state A , T or U , G , and C , i.e., the dynamics of P_A , P_T , P_G , and P_C , respectively. The total probability is normalized and, therefore,

$$P_A + P_T + P_G + P_C = 1. \quad (5)$$

As an example, the rate equation for P_A can be written as

$$\frac{dP_A}{dt} = k_{TA}P_T + k_{GA}P_G + k_{CA}P_C - k_{AT}P_A - k_{AG}P_A - k_{AC}P_A. \quad (6)$$

Here, k_{TA} , k_{GA} , and k_{CA} represent Kramers rate of transition from the other three states to state A and k_{AT} , k_{AG} , and k_{AC} denote the same for passage from state A to the other three states. These rates are proportional to the exponential of the negative of the energy difference between the states scaled by the factor $k_B T$ where k_B is the Boltzmann constant and T is the temperature of the environment. For example, k_{AG} and k_{GA} can be expressed as [62]

$$k_{AG} \sim \exp\left(-\frac{E_{AG}}{k_B T}\right), \quad k_{GA} \sim \exp\left(-\frac{E_{GA}}{k_B T}\right). \quad (7)$$

Similarly, we can write down the expressions for all other rate pairs and consequently, the rate equations for all other probabilities, such as, $P_{T/U}$, P_G , and P_C . The time evolution

equations for these probabilities are coupled differential equations guided by the Kramers rate of transition between the states. Without any input signal, these rates are equal and when we solve the system of rate equations for the probabilities, asymptotically we retrieve equal occupancy $P = 1/4$ for each state.

Our model is calibrated in the following way: As said, the depths of the energy wells can be interpreted as binding energies of bases to the growing strand and to the complementary base of the “input.” The thermodynamic energy function which corresponds to the well depth in our model is the difference in Gibbs free energy of binding of a specific base with two individuals. In aqueous medium, this ranges from 0 to -2 kcal/mol at room temperature, obtained experimentally and also by molecular dynamics studies [63]. Converted into Joule per molecule, the actual energy differences which enter the Kramers rates are then of the order of $k_B T_{\text{room}}$ where $T_{\text{room}} = 300$ K. This allows us to calibrate the noise strength D in terms of the temperature of the environment T in our the Langevin model. The potential barriers ΔV in the energy landscape of $V(x, S_i)$ can be considered to represent the Gibbs free energy of binding of a given base with all the types of nucleic acid bases present in the system. We then have

$$\frac{\Delta V}{2D} \approx \frac{\Delta G}{k_B T} = \frac{\Delta G}{k_B T_{\text{room}}} \times \frac{T_{\text{room}}}{T}. \quad (8)$$

Now, we have found that the factor $\Delta G/(k_B T_{\text{room}})$ is of the order of 1 and ΔV is set to 2 by us for the unperturbed potential. Therefore, we get, $D \approx T/T_{\text{room}}$. This analysis clarifies our definition of D as the scaled temperature of the system where the reference temperature that proportionates the temperature of the environment, as mentioned in Sec. II A, is the room temperature. Actually, from the values of binding energies of different base pairs, we also conclude that a realistic well-depth factor is $b \approx 4$ and more generally between 1 and 10.

Let us now consider the example when the system reads the signal T such that the potential well of A deepens by the factor $b > 1$. Hence, all energy barriers from A to the other states increase by a factor b . This reduces all Kramers rates for the transition from A to the other states, exemplified here for $A \rightarrow G$:

$$k'_{AG} \sim \exp\left(-\frac{bE_{AG}}{k_B T}\right) = k_{AG}^b, \quad (9)$$

whereas k_{GA} remains the same. All other energy gaps and the corresponding Kramers rate change accordingly. Therefore, we can rewrite Eq. (6) as

$$\begin{aligned} \frac{dP_A}{dt} &= k_{TA/UA}P_{T/U} + k_{GA}P_G + k_{CA}P_C \\ &\quad - k'_{AT/AU}P_A - k'_{AG}P_A - k'_{AC}P_A. \end{aligned} \quad (10)$$

The other three rate equations can be altered in the similar way by considering appropriate transition rates. Now, the set of four rate equations for the probability of states serve as the representation of the dynamics of the system when it reads a particular signal. We analytically solve these differential equations to get the probability in the long time limit. For the ideal copying process, the correct state is expected to be occupied with a probability close to 1 as the modified rates introduce proper bias into the transition probabilities between the states.

However, the precise value of this probability depends on the parameters influencing the dynamics, and most prominently from the environmental temperature T .

Analytical solution of the four-state model

It is straight forward to derive the analytical solution of the rate equations for the probability of the states. Let us first consider Eq. (6), the situation without external input. All the transition rates between the states are equal, i.e.,

$$\begin{aligned} k_{AT/AU} = k_{TA/TU} = k_{AG} = k_{GA} = k_{AC} = k_{CA} \\ = k. \end{aligned} \quad (11)$$

Introducing this condition Eq. (11) in Eq. (6), we arrive at the following equation:

$$\frac{dP_A}{dt} = k(P_{T/U} + P_G + P_C) - 3kP_A. \quad (12)$$

Incorporating the normalization condition Eq. (5) in Eq. (12), we obtain

$$\frac{dP_A}{dt} + 4kP_A = k. \quad (13)$$

Multiplying Eq. (13) with the integrating factor $e^{\int 4kdt}$, we integrate the equation from initial time at $t = 0$ to any arbitrary time t and get the following form of P_A as a function of t ,

$$P_A(t) = P_A(0)e^{-4kt} + \frac{1}{4}[1 - e^{-4kt}], \quad (14)$$

where $P_A(0)$ is the initial value of the probability for state A . Equation (14) shows that in the asymptotic limit, i.e., for $t \rightarrow \infty$, P_A becomes $\frac{1}{4}$. The same is true for the probability of any other state. This is physically understandable that in absence of any external bias all possible states have the equal probability of occupancy.

Next, we consider the copying process when the system receives a particular signal and the transition rates between the states get modified accordingly. In our example, as the system reads signal T , the minima corresponding to the state A gets lowered. Taking into account Fig. 4 and Eq. (9), we can understand that $k_{TA/TU}$, k_{GA} , and k_{CA} will remain unaltered, whereas $k_{AT/AU}$, k_{AG} , and k_{AC} will take larger values. At this point, we define the transition rates for this circumstance as follows:

$$\begin{aligned} k_{TA/TU} = k_{GA} = k_{CA} = k, \\ k'_{AT/AU} = k'_{AG} = k'_{AC} = k'. \end{aligned} \quad (15)$$

Considering the above condition Eq. (15) and the normalized total probability, we can rewrite Eq. (10) as

$$\frac{dP_A}{dt} + (k + 3k')P_A = k. \quad (16)$$

We solve Eq. (16) with the aid of the integrating factor $e^{\int (k+3k')dt}$ to get the following form of $P_A(t)$:

$$P_A(t) = P_A(0)e^{-(k+3k')t} + \frac{k}{k+3k'}[1 - e^{-(k+3k')t}]. \quad (17)$$

In the long time limit, i.e., for $t \rightarrow \infty$, $P_A(t)$ can be expressed as

$$P_A(\infty) = \frac{k}{k+3k'}. \quad (18)$$

The above expression Eq. (18) shows that when the transition rates from the correct state to the other states are much smaller compared to the transition rates from the other states to the correct state, i.e., $k' \ll k$, which happens to be the condition in the ideal situation, we can ignore $3k'$ with respect to k in the denominator of the expression at the right-hand side of Eq. (18) and this consideration implies that $P_A(t)$ will tend to value 1 in the asymptotic limit. This indicates the fact that for the correct copying process, the steady state probability of the appropriate state will attain a value close to 1. As the transition rates depend on the barrier height of the potential and the temperature, we will study the success of the biological copying process in dependence on these two parameters.

III. RESULTS AND DISCUSSIONS

We want to measure the accuracy of the copying mechanisms as a function of the system parameters. For this purpose, we define a quantity which we call success rate for the process. Here, we explain our interpretation for success rate considering the Langevin and the four-state model.

A. Success rate

For a given input, we know the desired output for each of our two models, namely, the symbol which matches the input. If the output is indeed identical to the desired one, then we call this success. We define the success rate S as the fraction of correct output among all input-output relations which we test. As inputs, all four symbols should be taken with the same probability. The details of how the success rates are calculated depend on the model used. In both cases we study the success rate as a function of temperature which is encoded in the diffusion constant D , where additionally the influence of other model parameters needs to be studied. These are the wdf and the t_{ev} , i.e., the time after which we evaluate the current state and present the next input symbol and hence change the potential.

1. Langevin model

We solve Eq. (1) numerically using a random string of 10^6 input symbols. Depending on the input, the potential gets adjusted to make the appropriate minimum which represents the correct output, the deepest one. A trajectory is allowed to evolve for a fixed t_{ev} which is considerably longer than the intrinsic relaxation time of the system. At the end of the evolution time, we examine whether the trajectory is in the correct potential well with respect to the input signal, which is the global minimum. If so, then we say that the information has been copied properly into the developing entity from its parent, and we count this as success. Each such time evolution is regarded as a single step in the copying mechanism, indicating that a single base sequence is encountered by the copying machinery and consequently, a base joins the developing DNA or mRNA strand to expand its length. Starting from the value of the random variable at the end of one evolution time, we continue the stochastic path in the “new” potential in the next evolution time interval. Due to the randomness of the new input, this guarantees that we probe all possible transitions between potential minima. This is relevant, since we have

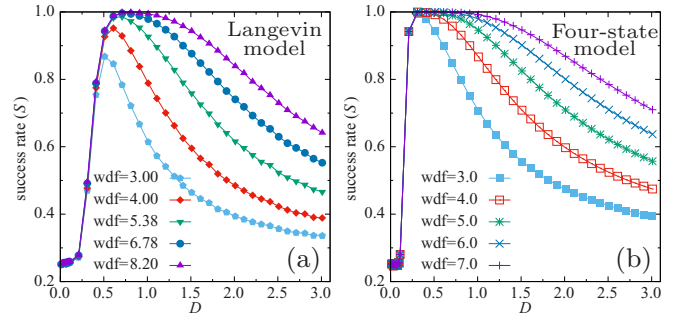


FIG. 5. Variation of the success rate S against noise strength D for different values of the wdf for the (a) Langevin and the (b) four-state model.

a spatial ordering of the minima, so that, e.g., a transition from the leftmost to the rightmost one means to overcome three potential barriers which might require more time or be less probable in the given time interval t_{ev} than a transition between adjacent minima. Each of these stochastic steps can produce correct or incorrect outcomes. The success rate is the total number of successful steps out of 10^6 trials, divided by the number of trials. The initial condition $x(0)$ is chosen randomly, however, since it is exponentially suppressed it does not play any role.

Such a simulation run can be interpreted as the propagation of the copying machinery along a template strand recognizing a base in it at each step. Therefore, the success rate which we calculate, reflects the exactness of the whole process in which a DNA or an mRNA is created exploiting the information stored in the precursor strand.

2. Four-state model

We have presented the analytical solutions of the four-state model in Eq. (17), which is the time dependent probability to be in the deepest well, i.e., to find the correct output state. In analogy with the Langevin model, we insert a t_{ev} in Eq. (17) and read the success rate which is essentially represented by the probability P_i . Due to symmetry, the success rate is the same for all possible input symbols and for all possible initial states when starting a new evolution time. For a close to optimal copying process, the value of $P_i(t_{ev})$ for the correct state i has to achieve a value very close to 1, while the other three should be close to 0.

B. The effect of temperature on the success rate

Here we present the results for the dependence of the accuracy of the biological copying mechanisms on the temperature of the surroundings. For this purpose, we explore the variation of the success rate for the processes as a function of the diffusion coefficient D which is present in the dynamics as a representation of the temperature. In both models, the success rate of the copying processes depends significantly on temperature and it shows a maximum with respect to the variation of the diffusion coefficient D . Quantitative results are shown in Figs. 5(a), 5(b) and 6(a), 6(b) considering both the Langevin and the four-state model. Figures 5(a) and 6(a) represent the results obtained from the Langevin model and

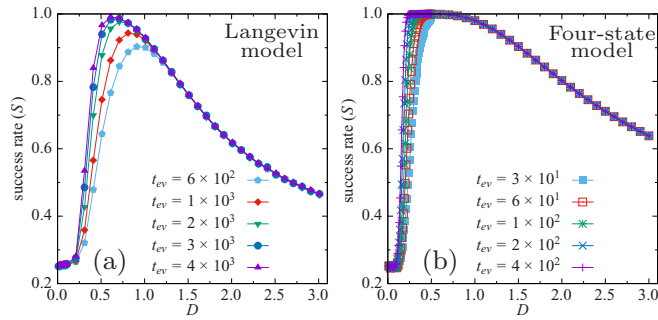


FIG. 6. Variation of the success rate S against noise strength D for different values of the t_{ev} for the (a) Langevin and the (b) four-state model.

Figs. 5(b) and 6(b) exhibit the same for the four-state model. There are two parameters in the dynamics that can influence the variation of the success rate against D ; one is the wdf and another is the t_{ev} . It is evident that the time to escape from a well, the first passage time, is related to the Kramers rate and hence is the shorter, the lower the barrier and the stronger the noise. However, for strong noise or low barriers, there is also a larger probability to escape from the deeper well. Hence, if t_{ev} is small compared to the inverse Kramers rate, then the particle has no chance to escape from the well where it starts, while if the t_{ev} is larger than the inverse Kramers rate to jump back from the deep well into a flat well, t_{back} , the particle also might be in the wrong well at the end. Both Kramers times are controlled by the noise strength, and the time t_{back} in addition by the well-depth factor. This comparability between the two timescales in the dynamics, t_{ev} and the Kramers time, justifies the occurrence of the maxima for the success rate when plotted against the scaled temperature D .

First, we keep t_{ev} fixed and observe the dependence of the success rate S on D for different values of wdf. The timescale of the four-state model is smaller as compared to the Langevin model. Therefore, the constant values of t_{ev} are chosen as 3000 for Langevin model and 300 for the four-state model, respectively. Both of these time duration are sufficiently longer with respect to the intrinsic relaxation timescale of the systems. The results from this study have been presented in Figs. 5(a) and 5(b) for the Langevin and the four-state model, respectively. Although the wdf have similar interpretation signifying the relative difference in the depth of the deepest well compared to others, their same numerical values do not represent the identical difference in energy levels between the global and local minima for the two models. This can be understood from our definition of the models. We compare the findings obtained from the two models and see that both of them produce qualitatively similar kind of dependence of the success rate S on the diffusion coefficient D . The success rate S first increases and then decreases exhibiting a maximum (close to or at 1) when D is varied. The value of the success rate $S = 1$ implies perfect copying of the biological information. The maximum in the curve suggests that the information gets copied most efficiently at an optimum value of the temperature. For different values of the wdf at a fixed t_{ev} , the rise of the success rate S to the maximum as a function of D is more or less the same. However, the success rate S remains near the

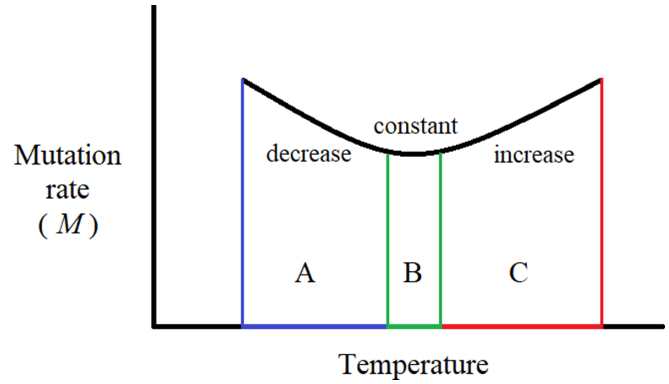


FIG. 7. Schematic representation of the variation of mutation rate against temperature as described in Ref. [31]. The temperature range has been split into three different regions to indicate three characteristic variation of the mutation rate with respect to temperature. The portion of the curve at region A illustrates decrease of mutation rate as a function of temperature as described in Ref. [35], the right-most section of the curve at region C supports the observed increase of mutation rate with respect to temperature as found in the studies [33,34] and the flatness of the curve within region B depicts that mutation rate can remain constant with alteration of temperature as mentioned in Refs. [37,38].

maximum for a greater range of D when the value of the wdf is higher. This indicates that the greater stability of the correct state keeps the accuracy of the copying process to persist for a wider range of physical conditions.

Next, we fix the wdf and study the variation of the success rate S against D for different t_{ev} . For the Langevin model the wdf has been kept fixed at 5.38 and the value is 6.0 for the four-state model. The success rate S shows a maximum against D and it is observed that for higher value of the t_{ev} , the maximum of the success rate S appears at a lower value of noise strength [Figs. 6(a) and 6(b) for the Langevin and the four-state model, respectively]. This points toward the fact that, if the t_{ev} is longer, then the correct minimum of the potential can be occupied with a lower level of noise, i.e., the correct sequence can be copied at a lower temperature. However, above a certain value of the noise strength the t_{ev} does not have any effect on the success rate S .

The above observation which depicts that the biological information can be copied most efficiently at an optimum temperature, is the central result of the present study. The success rate S in our model is related to the mutation rate M as $M = 1 - S$, since a fraction M of the information would not be copied correctly. Hence, our model predicts a minimum mutation rate at an optimal temperature. This result is motivating because it supports some observations in real biological systems. It is reported that the spontaneous mutation rate to respiration deficiency in *saccharomyces* yeast shows a minimum with respect to temperature variation [31]. This has been represented schematically in Fig. 7. It has also been found that the mutation rate can increase [33,34], decrease [35], or remain constant [37,38] with increasing temperature for different organisms. In a recent experiment, it has been observed that the mutation rate in an *Escherichia coli* strain first remains constant and then increases with increasing temperature [28].

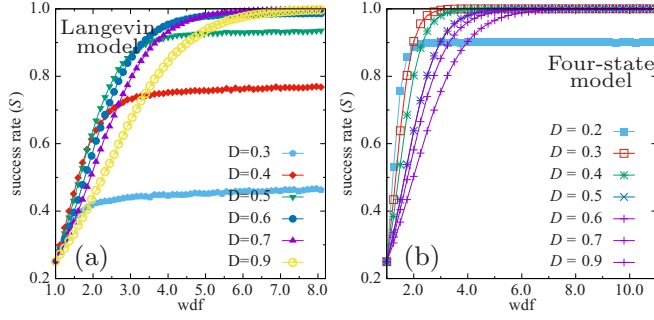


FIG. 8. Variation of the success rate S against the wdf for different values of noise strength D for the (a) Langevin and the (b) four-state model.

These observations can also be explained with the curve obtained from the simulations of our models. We see that the success rate initially rises and then decreases with growing temperature. Therefore, the first trend can be described to support the drop of the mutation rate with increasing temperature (temperature region A in Fig. 7). Similarly, the right-most portion of the curve (temperature region C in Fig. 7) can be considered to represent the increase of the mutation rate as a function of temperature. We also detect some flatness in the curves (temperature region B in Fig. 7) depending on the parameter values which can be thought to correspond to the constant mutation rate with respect to temperature. We can argue that different organisms and their copying machinery are only stable in a given temperature range and they show the characteristic variation of the mutation rate of that particular region. Therefore, we can say that our model, although it is a minimal model with the lowest thinkable set of free parameters, is flexible enough to yield explanations for all observed scenarios of the observed temperature dependence of the mutation rates.

C. The effect of wdf and t_{ev} on the success rate

To complete the study of our models, we will explore in more detail the influence of other parameters, namely, the wdf and the t_{ev} on the success rate S of the copying mechanism.

With increasing wdf the success rate S tends to increase monotonically and saturates beyond a certain value. This implies that the selections are more accurate when the local minima and the global minimum of the potential have larger energy difference and the success rate of the copying process saturates after a certain value of the energy difference between the correct and the incorrect state is reached. This has been illustrated in Figs. 8(a), 8(b) and Figs. 9(a), 9(b) for both the Langevin and the four-state model. The saturation value of the success rate S and the value of wdf at which the saturation sets in, depend on the noise strength (Figs. 8(a) and 8(b) for the Langevin and the four-state model, respectively) for a fixed value of the $t_{ev} = 3000$ for the Langevin model and $t_{ev} = 300$ for the four-state model. The dependence of the saturation value of the success rate on the noise strength is understandable from the variation of the success rate S with respect to D . These two factors, the position of the saturation point and the value of the saturated success rate also show

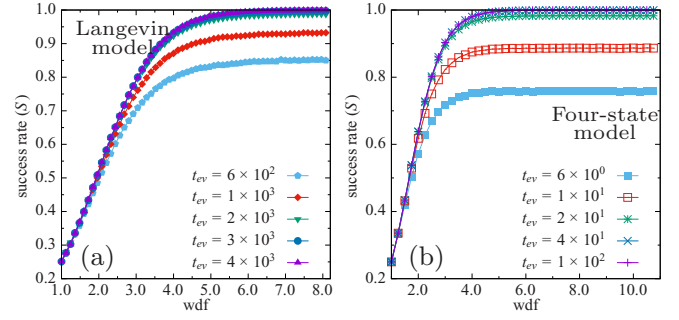


FIG. 9. Variation of the success rate S against wdf for different values of the t_{ev} for the (a) Langevin and the (b) four-state model.

dependence on the t_{ev} , for a fixed D [Figs. 9(a) and 9(b) for the Langevin and the four-state model, respectively]. Here, D has been kept fixed at 0.7 for the Langevin model and at 0.6 for the four-state model. The higher saturation value of the success rate for longer t_{ev} suggests that the copying process becomes more successful when the dynamics evolves for a larger time. This can be expected from the behavior of a stochastic trajectory. If we let the system propagate for a longer time, then it is more probable that the trajectory ends up in the correct minimum at the completion of the process. Actually, one might ask whether the trajectory then also has a larger chance to escape again from the correct well. Indeed, this escape from the deepest well exists, but due to the fact that with probability close to $3/4$ the initial state of a new time step is located in a different minimum, the probability to be in the correct well starts with a value close to $1/4$ and increases until it has reached its equilibrium value if the evolution time were infinite.

To better visualize the dependence of the success rate S on the t_{ev} , we choose t_{ev} as the variable and noise strength D and the wdf as the parameters. In agreement with the previous observations, we see that the S increases with increasing t_{ev} and then saturates after a certain value [Figs. 10(a), 10(b) and Figs. 11(a), 11(b)]. The process of the saturation and the saturated value of the success rate S depends significantly on the noise strength D [Figs. 10(a) and 10(b) for the Langevin and the four-state model] for a given wdf. Here the wdf equals 5.38 and 6.0 for the Langevin and the four-state model, respectively. The saturation of the success rate S with respect to

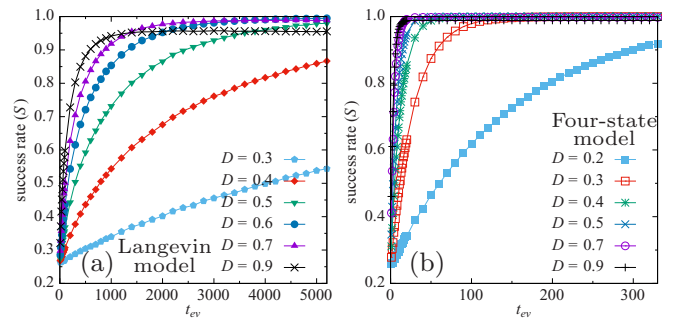


FIG. 10. Variation of the success rate S against t_{ev} for different values of noise strength D for the (a) Langevin and the (b) four-state model.

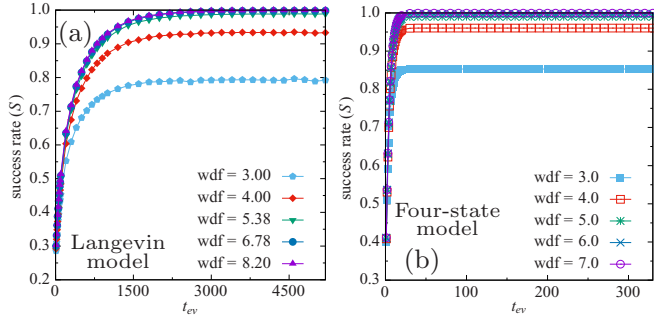


FIG. 11. Variation of the success rate S against the t_{ev} for different values of the wdf for the (a) Langevin and the (b) four-state model.

t_{ev} occurs faster for higher values of D . However, if we look at the absolute saturation value for individual noise strengths, then we can see it is maximum at an intermediate value of the noise strength. This is consistent with Figs. 5 and 6 where we saw a maximum in the success rate S when plotted against D . The other parameter wdf also has some influence on the success rate versus t_{ev} curve for a fixed value of D ($D = 0.7$ and 0.6 for the Langevin and the four-state model, respectively). The approach to the saturation of the success rate with respect to t_{ev} is more or less independent of the wdf. However, the absolute value of the success rate in this variation depends on the wdf [Figs. 11(a) and 11(b) for the Langevin and the four-state model, respectively]. The greater is the wdf, the larger becomes the value of the success rate, implying better accuracy in the copying mechanism for larger energy difference between the correct and the incorrect states. The success rate reaches the maximum after certain value of the wdf and then becomes independent of it.

D. Speed of the copying process and the efficiency

The efficiency of the copying process is not only determined by its success rate S , but also by the time it takes to generate a copy of a strand of given length, i.e., by the speed of the copying mechanism. In our model, we have the t_{ev} which determines the duration of a single copying step. For a chosen high success rate close to 1, the process with the shortest t_{ev} is the most efficient one. From Fig. 11(a) one would conclude that with a well-depth factor of $wdf \geq 5$ and $D = 0.7$ an evolution time of ≈ 2000 is the most efficient choice of parameters, for the Langevin model. Similarly, Fig. 11(b) suggests that for $wdf \geq 5$ and $D = 0.6$, $t_{ev} \approx 30$ would produce the most efficient copying steps corresponding to the four-state model.

To interpret the efficiency of the copying process, we intend to estimate the quantity by defining it in terms of the success rate S or the mutation rate M and the t_{ev} . We consider the success rate S versus t_{ev} plots and note down the smallest value of t_{ev} at which the success rate S saturates and denote it by the symbol t_{ev}^s . We measure the mutation rate M from the success rate S by employing the relation $M = 1 - S$. Now, if we take a product of the mutation rate M and t_{ev}^s , we will get an estimate about the extent of the time extensive mutation in the copying process. If we evaluate the reciprocal of this factor, then we will be able to measure the overall efficiency

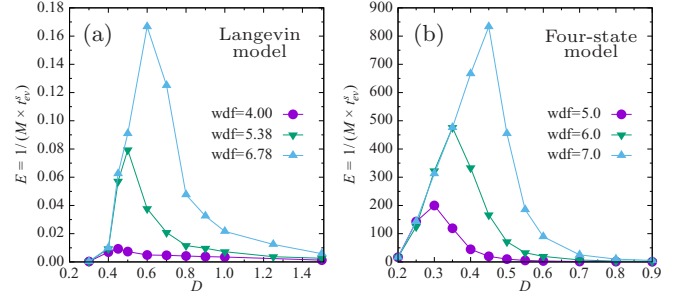


FIG. 12. Variation of the efficiency E with respect to the noise strength D for different values of the “well-depth factor” for the (a) Langevin and the (b) four-state model.

of the copying mechanism which takes into account both the correctness of the process as well as its speed. Therefore, the efficiency for the copying procedure can be explicitly expressed as

$$E = 1/(M \times t_{ev}^s). \quad (19)$$

We observe the variation of the quantity E against the noise strength D . The efficiency shows a maximum with respect to the variation of D . This finding confirms the fact that biological information gets copied most precisely at an intermediate value of temperature. This has been presented in Figs. 12(a) and 12(b) for the Langevin and the four-state model, respectively. We have carried out the analysis for different well-depth factors and shown three representative curves corresponding to the variation of the efficiency against temperature, for both the models. Here, we see that the peak value of the efficiency in the curves increases with increasing wdf, suggesting the occurrence of a more efficient copying process when the energy difference between the correct and the incorrect interactions is larger. However, this peak value becomes constant when a certain range of wdf is reached. This is because the saturation pattern of the success rate against time becomes independent of the wdf after a definite threshold of the wdf value. This has been pointed out while discussing Figs. 11(a) and 11(b).

While concluding the discussions about our findings, we would like to mention that we started our intended study with the one-dimensional Langevin model which is probably the simplest possible representation of the copying mechanism. To examine the fact whether the dimensionality of the model has any effect on the results, we introduced the four-state model which takes four discrete states into account and does not depend anyway on the dimensionality of the potential energy landscape. In all of our parameter studies, we have found that both the models produce same qualitative results with only quantitative difference in the timescale. This confirms the fact that the details of the dimensionality of the model does not leave any impact on the observations.

IV. CONCLUSIONS

We summarize those biological processes, such as replication, transcription, translation where information is transferred to the nascent species from its predecessor, as biological copying mechanisms. We propose two minimal physical

models for this mechanism, namely, a general Langevin model and a four-state model that can represent such biological copying processes. We explore the accuracy of these copying procedures as a function of the parameters influencing the system and dynamics. The most important factor that introduces the effect of the noisy environment directly into the dynamics is temperature. We observe interesting dependence of the correctness of the copying processes on temperature. The success rate of the copying mechanisms shows a peak, and hence the mutation rate has a minimum, when plotted against temperature implying that these processes are most efficient at an optimum value of the temperature. Therefore, the temperature dependence differs from what is known from general considerations of the conventional reaction rate theory which concerns the occurrence of collisions between the reactants. In the common scenario the reaction rate increases with increasing temperature as the probability of encounter between the reacting particles becomes higher at an elevated temperature. In transcription, translation or replication process, selection of appropriate species or reactant is important and our models take that into account. The mechanisms introduced in our models result in nonmonotonic behavior of the successful outcome of the process against temperature instead of a steady rise of the rate. To be more precise, we analyze the copying process by considering the nonequilibrium nature of the mechanism where input is transformed into output. There exist cooperation between the forward and backward rates which is modulated by the temperature and competition between the rates of correct and incorrect transitions which depends on the relative energy difference between the appropriate and inappropriate interactions. These rates are also affected by the finite-time duration of the copying steps. These distinct features and the nonequilibrium transient nature of the dynamics are reflected in the characteristic variation of the success rate with respect to temperature.

Our measure of the correctness of the copying procedures, the success rate, is complementary to the mutation rate. We recalled the findings in some real biological systems [28,31,33–35,37,38] regarding the variation of the mutation rate with respect to temperature. In some biological experiments, mutagenesis, the alteration of genetic information causing mutation, has been found to be favored at higher temperature [28–30,64]. In case of certain species in nature, it is also observed that the mutation rate increases with increasing temperature during the global climate change [65,66]. This rise in mutation rate with respect to temperature could be fully explained by reaction rate theory. However, such increase is only one of different behaviors which have been observed as a function of temperature variation in biological experiments

and in nature [28,31,33–35,37,38,67,68]. It is the advantage of our models that they can explain all the different scenarios which have been found.

Here we emphasize that in the present study, we only consider the information copying step for the DNA replication process and disregard the proofreading mechanism. Error-correcting steps also occur in case of transcription or translation [69], however, the mechanisms differ from DNA proofreading. We also do not take into account this defect rectification associated with transcription or translation. Therefore, the current results illustrate the dependence of the accuracy of the copying step on the significant parameters linked to the system and dynamics. However, we have carried out a simple analysis to understand the incorporation of the proofreading procedure for the DNA replication process. We have modeled the proofreading mechanism with an erasure protocol in which the system is driven to the correct state by an external drive if it ends up in the wrong state after the copying process. Interestingly, the study shows that the efficiency of the erasure mechanism also exhibits a maximum at an intermediate value of the temperature. So, even if we couple the two steps, copying and erasure, our interpretation about the role of thermal fluctuations on the replication process remains unaltered. So, we can say, the biological processes involving information transfer, such as, replication or transcription or translation, occur most efficiently at an optimum value of temperature.

In reality, replication, transcription and translation processes are highly accurate. The percentage of errors in these mechanisms depends on the type of species involved. In most cases, the transcription and translation processes have error frequency 1 in 10^4 . DNA replication process has higher accuracy because of the proofreading mechanism [70,71]. We consider that the observed perfection appearing in the biological information copying processes might be a result of the survival of the organisms in the physical zone where the success rate of the processes is very high and the energy difference between the correct and incorrect interactions between the base pairs is remarkably different. We also define another quantity, the efficiency of the process, taking into account the mutation rate and the speed of the copying mechanism, for our models. This estimate is somewhat dynamic in nature as it considers the effective evolution time of a copying step. The variation of the efficiency as a function of temperature also suggests maximum efficacy of the copying procedures at an intermediate value of temperature. Finally, we can say that our proposed models represent a general scheme for the biological copying processes and can explain the observations concerning the environmental dependence of mutation with a simple approach.

-
- [1] C. H. Bennett, *IBM J. Res. Dev.* **17**, 525 (1973).
 - [2] C. H. Bennett, *Int. J. Theor. Phys.* **21**, 905 (1982).
 - [3] P. Sartori and S. Pigolotti, *Phys. Rev. Lett.* **110**, 188101 (2013).
 - [4] A. C. Barato and U. Seifert, *Phys. Rev. Lett.* **114**, 158101 (2015).
 - [5] R. Rao and M. Esposito, *Phys. Rev. X* **6**, 041064 (2016).
 - [6] T. E. Ouldridge and P. Rein ten Wolde, *Phys. Rev. Lett.* **118**, 158103 (2017).
 - [7] L. Loewe, *Nature Education* **1**, 113 (2008).
 - [8] E. Sober (Editor), *Conceptual Issues in Evolutionary Biology: An Anthology* (MIT Press, Cambridge, MA, 1984).
 - [9] R. A. Fisher, *The Genetical Theory of Natural Selection* (Oxford University Press, Oxford, 1930).

- [10] P. D. Stenson, M. Mort, E. V. Ball, K. Evans, M. Hayden, S. Heywood, M. Hussain, A. D. Phillips, and D. N. Cooper, *Hum. Genet.* **136**, 665 (2017).
- [11] H. Ma *et al.*, *Nature* **548**, 413 (2017).
- [12] R. M. Stone *et al.*, *N. Engl. J. Med.* **377**, 454 (2017).
- [13] R. Sanjuàn and P. Domingo-Calap, *Cell. Mol. Life Sci.* **73**, 4433 (2016).
- [14] R. Rahbari *et al.*, *Nat. Genet.* **48**, 126 (2016).
- [15] T. A. Hopf *et al.*, *Nat. Biotechnol.* **35**, 128 (2017).
- [16] M. L. Mugnai, C. Hyeon, M. Hinczewski, and D. Thirumalai, *Rev. Mod. Phys.* **92**, 025001 (2020).
- [17] Q.-S. Li, P.-D. Zheng, Y.-G. Shu, Z.-C. Ou-Yang, and M. Li, *Phys. Rev. E* **100**, 012131 (2019).
- [18] D. Hathcock and S. H. Strogatz, *Phys. Rev. E* **100**, 012408 (2019).
- [19] L. Loewe and W. G. Hill, *Phil. Trans. R. Soc. B* **365**, 1153 (2010).
- [20] J. B. Coe, Y. Mao, and M. E. Cates, *Phys. Rev. E* **70**, 021907 (2004).
- [21] W. Peng, H. Levine, T. Hwa, and D. A. Kessler, *Phys. Rev. E* **69**, 051911 (2004).
- [22] N. Guttenberg, S. M. Ali Tabei, and A. R. Dinner, *Phys. Rev. E* **84**, 031932 (2011).
- [23] J. Otwinowski and S. Boettcher, *Phys. Rev. E* **84**, 011925 (2011).
- [24] I. Bena, M. Droz, J. Szwabiński, and A. Pekalski, *Phys. Rev. E* **76**, 011908 (2007).
- [25] A. S. Ribeiro, *Phys. Rev. E* **78**, 061902 (2008).
- [26] N. Chia and N. Goldenfeld, *Phys. Rev. E* **83**, 021906 (2011).
- [27] M. O. Lavrentovich and D. R. Nelson, *Phys. Rev. E* **100**, 042406 (2019).
- [28] X. L. Chu, B. W. Zhang, Q. G. Zhang, B. R. Zhu, K. Lin, and D. Y. Zhang, *BMC Evol. Bio.* **18**, 126 (2018).
- [29] D. Berger, J. Stångberg, J. Baur, and R. J. Walters, *Proc. R. Soc. B* **288**, 20203094 (2021).
- [30] D. Berger, J. Stångberg, K. Grieshop, I. Martinossi-Allibert, and G. Arnqvist, *Proc. R. Soc. B* **284**, 20171721 (2017).
- [31] M. Ogur, S. Ogur, and R. St. John, *Genetics*. **45**, 189 (1960).
- [32] D. Lindgren, *Hereditas*. **70**, 179 (1972).
- [33] H. J. Muller, *Genetics*. **13**, 279 (1928).
- [34] H. H. Plough, *Cold Spring Harbor Symposia Quant. Biol.* **9**, 127 (1941).
- [35] G. H. Beale and A. C. Fabergè, *Nature* **147**, 356 (1941).
- [36] M. M. Rhoades, *Cold Spring Harbor Symposia Quant. Biol.* **9**, 138 (1941).
- [37] M. Demerec, *Proc. Natl. Acad. Sci. USA* **18**, 430 (1932).
- [38] E. M. Witkin, *Proc. Natl. Acad. Sci. USA* **39**, 427 (1953).
- [39] A. M. Garcia, R. B. Calder, M. E. T. Dollè, M. Lundell, P. Kapahi, and J. Vijg, *Plos Genet.* **6**, e1000950 (2010).
- [40] M. I. O'Connor, J. F. Bruno, S. D. Gaines, B. S. Halpern, S. E. Lester, B. P. Kinlan, and J. M. Weiss, *Proc. Natl. Acad. Sci. USA* **104**, 1266 (2007).
- [41] U. Seifert, *Rep. Prog. Phys.* **75**, 126001 (2012).
- [42] D. A. Drummond and C. O. Wilke, *Nat. Rev. Genet.* **10**, 715 (2009).
- [43] J. J. Hopfield, *Proc. Natl. Acad. Sci. USA* **71**, 4135 (1974).
- [44] R. R. Freter and M. A. Savageau, *J. Theor. Biol.* **85**, 99 (1980).
- [45] A. Murugan, D. A. Huse, and S. Leibler, *Proc. Natl. Acad. Sci. USA* **109**, 12034 (2012).
- [46] J. Ninio, *Biochimie* **57**, 587 (1975).
- [47] P. Sartori and S. Pigolotti, *Phys. Rev. X* **5**, 041039 (2015).
- [48] M. Voliotis, N. Cohen, C. Molina-París, and T. B. Liverpool, *Phys. Rev. Lett.* **102**, 258101 (2009).
- [49] P. Gaspard, *Phys. Rev. E* **93**, 042420 (2016).
- [50] P. Gaspard, *Phys. Rev. E* **96**, 042403 (2017).
- [51] F. Wong, A. Amir, and J. Gunawardena, *Phys. Rev. E* **98**, 012420 (2018).
- [52] H. Schiessel and R. Blossey, *Phys. Rev. E* **101**, 040401(R) (2020).
- [53] W. D. Piñeros and T. Tlusty, *Phys. Rev. E* **101**, 022415 (2020).
- [54] J. F. Gentleman and R. C. Mullin, *Biometrics* **45**, 35 (1989).
- [55] M. Das, D. Das, D. Barik, and D. S. Ray, *Phys. Rev. E* **92**, 052102 (2015).
- [56] R. W. Hamming, *Numerical Methods for Scientists and Engineers* (McGraw Hill, New York, NY, 1973).
- [57] D. Chiuchìù, J. Ferrare, and S. Pigolotti, *Phys. Rev. E* **100**, 062502 (2019).
- [58] H. A. Kramers, *Physica (Utrecht)* **7**, 284 (1940).
- [59] R. Landauer and J. A. Swanson, *Phys. Rev.* **121**, 1668 (1961).
- [60] P. Hänggi, P. Talkner, and M. Borkovec, *Rev. Mod. Phys.* **62**, 251 (1990).
- [61] P. Reimann, G. J. Schmid, and P. Hänggi, *Phys. Rev. E* **60**, R1(R) (1999).
- [62] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, Springer Series in Synergetics, 3rd ed. (Springer-Verlag, Berlin, 2004).
- [63] P. Cieplak and P. A. Kollman, *J. Am. Chem. Soc.* **110**, 3734 (1988).
- [64] C. LeBlanc, F. Zhang, J. Mendez, Y. Lozano, K. Chatpar, V. F. Irish, and Y. Jacob, *Plant J.* **93**, 377 (2018).
- [65] M. B. Davis, R. G. Shaw, and J. R. Etterson, *Ecology* **86**, 1704 (2005).
- [66] S. U. Pauls, C. Nowak, M. Bálint, and M. Pfenninger, *Mol. Ecol.* **22**, 925 (2013).
- [67] B. L. Sheldon, *Australian J. Biol. Sci.* **11**, 85 (1958).
- [68] C. Held, *Polar Biol.* **24**, 497 (2001).
- [69] M. Depken, J. M. R. Parrondo, and S. W. Grill, *Cell Rep.* **5**, 521 (2013).
- [70] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell*, 4th ed. (Garland Science, New York, 2002).
- [71] F. Bernardi and J. Ninio, *Biochimie* **60**, 1083 (1978).