

Kinetic proofreading and the limits of thermodynamic uncertaintyWilliam D. Piñeros¹ and Tsvi Tlusty^{1,2,*}¹*Center for Soft and Living Matter, Institute for Basic Science (IBS), Ulsan 44919, Korea*²*Department of Physics, Ulsan National Institute of Science and Technology (UNIST), Ulsan 44919, Korea*

(Received 11 November 2019; accepted 24 January 2020; published 24 February 2020)

To mitigate errors induced by the cell's heterogeneous noisy environment, its main information channels and production networks utilize the kinetic proofreading (KPR) mechanism. Here, we examine two extensively studied KPR circuits, DNA replication by the T7 DNA polymerase and translation by the E. coli ribosome. Using experimental data, we analyze the performance of these two vital systems in light of the fundamental bounds set by the recently discovered thermodynamic uncertainty relation (TUR), which places an inherent trade-off between the precision of a desirable output and the amount of energy dissipation required. We show that the DNA polymerase operates close to the TUR lower bound, while the ribosome operates ~ 5 times farther from this bound. This difference originates from the enhanced binding discrimination of the polymerase which allows it to operate effectively as a reduced reaction cycle prioritizing correct product formation. We show that approaching this limit also decouples the thermodynamic uncertainty factor from speed and error, thereby relaxing the accuracy-speed trade-off of the system. Altogether, our results show that operating near this reduced cycle limit not only minimizes thermodynamic uncertainty, but also results in global performance enhancement of KPR circuits.

DOI: [10.1103/PhysRevE.101.022415](https://doi.org/10.1103/PhysRevE.101.022415)**I. INTRODUCTION**

Fast and accurate processing of molecular information is essential for proper control, growth, and regulation in the living cell. In carrying out essential tasks like protein synthesis, ribosomes are known to operate at error levels of 10^{-3} – 10^{-4} incorrect peptide bindings per cycle [1–3], with even lower error rates for polymerases carrying out RNA transcription (10^{-5} – 10^{-6}) [4–6] and DNA replication (10^{-7} – 10^{-9}) [7–9]. This implies that the involved molecular machines must readily discriminate and preferentially bind the correct substrates over very similar, yet incorrect, competing substrates. However, simple energy discriminant binding models imply prohibitive energy binding differences among the pool of analogous substrates, and fail to predict the high level of fidelity observed. Instead, as originally proposed by Hopfield [10] and Ninio [11], high accuracy may be achieved through kinetic proofreading (KPR), a mechanism that couples an effectively irreversible process to an intermediate reaction step which can then preferentially react or discard substrates—via kinetic discrimination—further down in the reaction pathway. In this manner, the original discrimination step that relies on binding energy differences is amplified through a second round of substrate verification. This mechanism has been confirmed experimentally for a variety of polymerase and ribosome systems [12–14], and later recognized in signal transduction [15,16] and homologous recombination [17].

While KPR facilitates high fidelity synthesis, it imparts a significant energy cost to the overall process. Furthermore, the nanometric scale of these molecular systems renders them vulnerable to strong thermal and active fluctuations from

the cellular environment, suggesting performance limits set by fundamental thermodynamic fluctuation-dissipation trade-offs [18–20]. Indeed, recent work on generic KPR models linked the amount of energy dissipated to the loss of configurational entropy during accurate product formation, and found that the efficiency of this process decreased rapidly for increasing levels of accuracy [21]. More generally, accuracy of the copying process was found to be fundamentally tied by the amount of excess work dissipated by the system, with higher dissipation corresponding to higher accuracy, independent of underlying system topology [22].

In analyzing KPR processes, one typically coarse-grains the full complex biochemical system into a discrete set of states connected by stochastic transitions approximated as a Markov process. However, even under these simplified dynamics, thermodynamics places an inherent energetic cost to the output precision of an observed quantity. This seminal result has been dubbed the thermodynamic uncertainty relation (TUR), which is expressed in terms of the trade-off measure \mathcal{Q} as

$$\mathcal{Q} \equiv \dot{Q}t\epsilon^2(t) \geq 2k_B T, \quad (1)$$

where T is the temperature, k_B is the Boltzmann constant, t is time, \dot{Q} is the energy dissipation rate, and $\epsilon^2(t) = \text{Var}X/\langle X \rangle^2$ is the ratio of mean and variance of a current observable X [23]. In short, Eq. (1) implies that a reduction in the uncertainty of an observable must be accompanied by a matching increase in energy consumption. Optimal trade-off is achieved in the limit of vanishing energy use (i.e., equilibrium) and with normally distributed heat dissipation [24]. The TUR was first shown to hold in the limit of linear response, and later proven to hold for any Markov jump process in the short or long time limits [25,26]. More recently, this relation has been

*tsvitlusty@gmail.com

shown to follow for currents from a generalized fluctuation theorem framework [27].

In the context of enzymatic kinetics, the TUR has been used to infer performance boundaries in molecular systems such as motors [28–30]. For instance, in a study by Hyeon and Hwang [31], the transport efficiency of microtubule protein motors was defined in terms of \mathcal{Q} and analyzed using experimental data, showing that \mathcal{Q} is suboptimized within physiological data, while the studied wild-type kinesin protein operates near a minimum in \mathcal{Q} , the defective mutant was about three times less efficient and did not display a minimum.

Clearly, the TUR not only places fundamental constraints on system performance, but highlights the degree to which present day molecular systems may have accommodated to this limitation. While extensive kinetic analysis of copying-fidelity or proofreading mechanisms have been advanced in various contexts [32–38], direct TUR analysis of experimental KPR systems is surprisingly lacking. To this end, we consider a recent work by Banerjee, Kolomeisky, and Igoshin [39] on the KPR systems of the *E. coli* ribosome and the T7 bacteriophage polymerase, and analyze these circuits in the context of TUR. In [39], the reaction networks were adapted to a standard Hopfield-Ninio KPR model using experimental values of the kinetic rate constants. Taking the physiological state as a reference point, they investigated the speed-accuracy trade-off as a function of the kinetic rates, finding that speed is prioritized over error rate. In a follow-up work on the same systems, they found that speed is also prioritized over energy dissipation and output noise [40].

In this paper, we offer a complementary view on the existing body of analysis, focusing on the fundamental implications of the TUR on the *synthesis process* in the KPR networks of *E. coli* ribosome and T7 DNA polymerase. We show that, in general, decreasing error rates and mean production times coincide with an underlying effective *reduced network* of reactions steps that minimizes the TUR measure \mathcal{Q} of production. Approaching this reduced network not only provides the best energetic trade-off between production precision and energy dissipation through \mathcal{Q} , but also decouples the operational speed from the dispersion of production. As a result, this regime minimizes trade-off constraints between the mean production time and the error rate for a given set of control parameters and fixed energy budget. Together, we show that approaching the reduced network regime corresponds to enhanced global performance of the studied ribosome and polymerase systems.

II. KINETIC PROOFREADING CIRCUITS

We study the standard Hopfield-Ninio KPR model as adapted to the T7 DNA polymerase and the *E. coli* ribosome by Banerjee *et al.* [39], using measured values of the kinetic rate constants. These networks, shown in Fig. 1, capture the main steps of nucleotide ligation in the polymerase and peptide elongation in the ribosome. In particular, these kinetic pathways model the steady-state processive action of the polymerase or ribosome and consist of the following three major steps. For the T7 DNA polymerase, the cycle starts at the polymerase bound to the growing DNA strand in state **E**

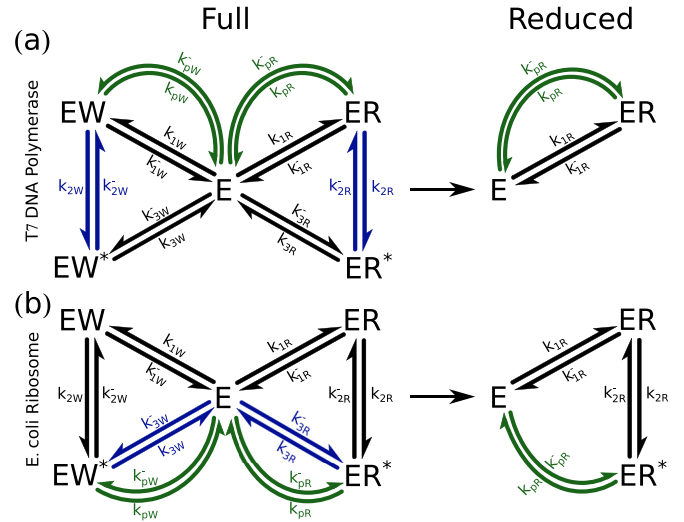


FIG. 1. Chemical reaction networks for (a) T7 DNA polymerase and (b) *E. coli* ribosome. Half-arrows indicate reversible forward and backward paths for both the right (R) and wrong (W) cycles. Kinetic rates are labeled by $k_{i,W/R}^{(-)}$, where $i = 1, 2, 3, p$ for each relevant path. Note that transitions through p in green (curved gray) half-arrows are implied to reset the enzyme to its initial state after product formation. Blue half-arrows (straight gray) indicate the proofreading transitions. Shown to the right of the black arrows are the reduced kinetic cycles (RC) where only relevant paths and rates leading to correct product formation are included.

where it adds either the correct (R) or incorrect (W) deoxy-NTP molecule to the growing strand and forms the **ER/W** state. The system can then ligate another dNTP molecule (path p) to restart the cycle or shift the strand to the polymerase exosite **ER*/W*** where the dNTP is hydrolyzed and removed back to state **E**. Similarly for the ribosome cycle in (b), state **E** represents the mRNA template bound in the growing ribosome polypeptide complex. Binding of the cognate (R) or non-cognate (W) aminoacyl-tRNA, EF-TU elongation factor and GTP leads to the second state **ER/W**. Hydrolysis of the GTP molecule brings the complex to state **ER*/W*** where the amino-acid can be added to the growing polypeptide strand (link p , green) or discarded (link 3, blue) which restarts the cycle. Note that the main difference in the topology of the cycles is that the KPR and production steps follow the first intermediate **ER/W** in the polymerase, whereas in the ribosome these occur only following the second hydrolyzed intermediate **ER*/W***.

All underlying kinetic rates $k_{i,W/R}^{(-)}$ for these cycles are assumed to be reversible and first order (i.e., measured in s^{-1} units) in constant substrate concentration at physiological conditions. Additionally, we maintain constant chemical potential differences $\Delta\mu$ of the underlying chemical reactions in both R and W cycles. This requirement constrains the rates as

$$\prod_{i=1}^N \frac{k_{i,W/R}}{k_{i,W/R}^{-}} = e^{\Delta\mu}, \quad (2)$$

where $\Delta\mu = \Delta\mu_{\text{KPR}}$ for proofreading cycles ($N = 3$) or $\Delta\mu_p$ for production cycle ($N = p$) and chemical potentials are

TABLE I. Values of calculated TUR measure \mathcal{Q} [Eq. (1)] and score ratios to its lower bound \mathcal{Q}_{lh} for a given number of states N and constant $\Delta\mu$ as defined in Eq. (9). For ribosomes, $N = 3$ and $\Delta\mu = \Delta\mu_p = 26k_B T$. For T7 polymerase $N = 2$ and $\Delta\mu = \Delta\mu_p = 11k_B T$.

	$\mathcal{Q} (k_B T)$	$\mathcal{Q}/\mathcal{Q}_{\text{lh}}$
Err ribosome	137	16
WT ribosome	48	5.6
Acc ribosome	28	3.2
T7 polymerase	7.1	1.3

hereafter measured in $k_B T$ units. We take the approximate physiological values of the chemical potential differences, $\Delta\mu_p \sim 26k_B T$ for polypeptide elongation, $\Delta\mu_p \sim 11k_B T$ for nucleotide ligation and $\Delta\mu_{\text{KPR}} \sim 20k_B T$ for the hydrolysis KPR step in both systems [40]. For convenience, we also define discrimination factors f_i which relate the R and rate constants as $f_i^{(-)} = k_{i,W}^{(-)}/k_{i,R}^{(-)}$ and represent the biased discriminant enzyme behavior when bound to either the right or wrong substrate. These factors are similarly constrained through Eq. (2),

$$\prod_{i=1}^N \frac{f_i}{f_i^{(-)}} = 1. \quad (3)$$

Values for all $k_{i,R/W}^{(-)}$ and $f_i^{(-)}$ are as adapted by Banerjee *et al.* [39] from experimental sources [41–44] and listed in Table II below.

For the purpose of our thermodynamic uncertainty analysis, we calculate the TUR measure \mathcal{Q} for *correct product transitions* across the path p in the ribosome and polymerase as follows. The relative uncertainty $\epsilon^2(t)$ across this production path can be found from the mean transition current $J_p^R = \langle X \rangle / t$ and its diffusion constant $D_p^R = \text{Var}X / 2t$ so that

$$\epsilon^2(t) = \text{Var}X / \langle X \rangle^2 = 2D_p^R / (J_p^R)^2 t. \quad (4)$$

TABLE II. Kinetic model parameters as originally reported by Banerjee *et al.* [39]. Kinetic rate constants $k_{i,R}^{(-)}$ reported in s^{-1} and discrimination factors f_i are dimensionless by definition. Constants k_3^- and k_p^- and discrimination factors f_3^- and f_p^- are derived from the constraint relations in Eqs. (3) and (2), respectively.

Parameters	Ribosome WT	Ribosome Acc	Ribosome Err	T7 polymerase
$k_{1,R}$	40	27	37	250
$k_{1,R}^-$	0.5	0.41	0.43	1
$k_{2,R}$	25	14	31	0.2
$k_{2,R}^-$	10^{-3}	10^{-3}	10^{-3}	700
$k_{3,R}$	8.5×10^{-2}	4.8×10^{-2}	7.7×10^{-2}	900
$k_{p,R}$	8.415	4.752	7.623	250
f_1	0.675	0.926	0.973	8×10^{-6}
f_1^-	94	112.2	9.3	1×10^{-4a}
f_2	4.8×10^{-2}	3.5×10^{-2}	0.126	11.5
f_2^-	1	1	1	1
f_3	7.9	10.34	7.65	1
f_p	4.2×10^{-3}	7.4×10^{-4}	4.1×10^{-3}	4.8×10^{-5}

^aValue chosen from same experimental limits to ensure positive values of J_{pW} .

Using the definition for \mathcal{Q} in Eq. (1) and ϵ above we thus get

$$\mathcal{Q} = 2\dot{Q}D_p^R / (J_p^R)^2, \quad (5)$$

where the energy dissipation $\dot{Q} = k_B T \sigma$ is defined in terms of the entropy production rate σ as

$$\sigma = J_{\text{KPR}} \Delta\mu_{\text{KPR}} + J_p \Delta\mu_p. \quad (6)$$

The currents that determine σ in Eq. (6) are $J_p = J_p^R + J_p^W$, the production current for both R and W cycles, and $J_{\text{KPR}} = J_i^R + J_i^W$, the discarded substrate current from kinetic proofreading (for $i = 2$ in polymerase or $i = 3$ ribosome). We also calculate the mean production time τ defined as the average time required to observe one net product addition onto the growing strand. This is given from the production rate as

$$\tau \equiv 1/J_p^R. \quad (7)$$

This definition of time is equivalent to a mean passage time in the limit of irreversible product formation and vanishing incorrect product rate $k_{pW/R}$. Similarly, the error η is defined as the fraction of incorrect substrate units added onto a growing peptide chain or DNA strand,

$$\eta = \frac{J_p^W}{J_p^W + J_p^R}, \quad (8)$$

where J_p^W is the current across the p link in the wrong W cycle. The values of the currents, J_p^R and J_p^W , and the diffusion constant D_p^R , are calculated using Koza's steady-state method [45], as demonstrated extensively in other studies [23,28,40].

In addition to the full reaction networks, it is instructive to consider idealized cycles consisting of only states and paths leading to correct product formation (shown on the right side in Fig. 1). These reduced cycles (RC) represent, by construction, perfect performance of the underlying protein systems. The RC circuits allow direct comparison of current-dependent metrics like \mathcal{Q} and τ between the actual and ideal system. Considering that $k_i \gg k_i^-$ for the systems studied here, the production current J_p^R in these idealized cycles is a simplified function of the forward rate constants k_i of

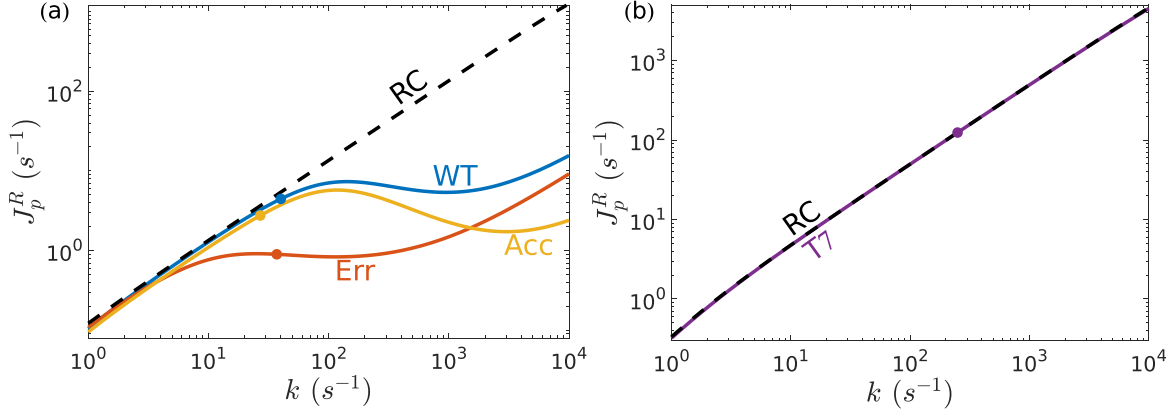


FIG. 2. Product cycle current J_p^R as a function of generalized rate constant k , defined as $k_{1R} = a_1 k$, $k_{2R} = a_2 k$, $k_{pR} = a_p k$ with $a_1 = 1$, $a_2 = k_{2R}^{\text{phys}}/k_{1R}^{\text{phys}}$ and $a_p = k_p^{\text{phys}}/k_{1R}^{\text{phys}}$, where k_i^{phys} are the physiological values. The physiological points are shown as dots. (a) The current of correct substrate production J_p^R for Wild-type ribosome (WT, blue top), more erroneous (Err, red bottom) and more accurate (Acc, yellow middle) mutants and the ideal RC current (dashed line). (b) Same as (a) but for T7 polymerase (solid line). The difference between the actual T7 current and RC is in the order of 1% and not noticeable at this scale.

the form $J_p^R \sim \prod_i^n k_i$. One can thus define the forward rate constants in terms of a single control parameter k as $k_i \equiv a_i k$, where $a_i \equiv k_i^{\text{phys}}/k_1^{\text{phys}}$ are the ratios of the rate constants at physiological values. It follows that current in the idealized circuit $J_p^R \sim k^n \prod_i^n a_i$ is a monotonic function (a power) of k that conserves the rate constant proportionality of the original system. The physiological state is matched when $k = k_1^{\text{phys}}$. As we later show, operating near the regime of the ideal RC cycles implies minimal \mathcal{Q} values, and affords enhanced accuracy/speed trade-off performance.

III. RESULTS AND DISCUSSION

A. The approach of KPR circuits to the TUR limit

We begin by reporting the physiological values of \mathcal{Q} for product transitions in the ribosome and polymerase systems as shown in Table I. Here, the T7 polymerase achieves the lowest value of \mathcal{Q} which is about seven times smaller than the native WT ribosome, with the more accurate mutant Acc closer to the limit than either WT or the less accurate Err. In order to compare these results more meaningfully, however, we must account for the underlying energy cost of cycle operation, which differs between the T7 DNAP and ribosome systems.

This can be achieved by considering the reduced cycles (RCs) consisting of only states and transitions leading to correct product formation as introduced in Sec. II and Fig. 1(b). These RCs were extracted from the full network of states and represent an idealized limit where only the correct substrate is processed in the absence of any competing paths. Operating at the RC limit therefore provides the best overall enzyme performance for a fixed energy budget. In particular, the RC limit implies a unicycle regime for which a lower bound for \mathcal{Q} is known to be

$$\mathcal{Q} \geq \mathcal{Q}_{\text{th}} \equiv 2k_B T \left(\frac{\Delta\mu}{2N} \coth \frac{\Delta\mu}{2N} \right), \quad (9)$$

where N denotes the number of states in the network, and $\Delta\mu$ is the overall change in Gibbs free energy of the underlying

chemical reactions per cycle (in $k_B T$ units) [30,46]. This hyperbolic lower bound \mathcal{Q}_{th} is achieved for a system with uniform forward and backward rate constants and reduces to the minimal value of $2k_B T$ in the vanishing $\Delta\mu$ limit. Thus, the hyperbolic bound is pertinent for far-from-equilibrium driven process such as KPR and represents the best efficiency attainable given an energy input. The ratio of the physiological values of $\mathcal{Q}/\mathcal{Q}_{\text{th}}$ shown in Table I are therefore a normalized optimization score, for a specific energetic constraint, of either the polymerase or ribosome systems. Markedly, the polymerase operates close to the TUR limit at $\mathcal{Q}/\mathcal{Q}_{\text{th}} \simeq 1.3$, while the ribosomes are 3–16 times further away, even after accounting for the specific energy cost of the underlying chemical transitions.

One can obtain a deeper appreciation of the score ratios $\mathcal{Q}/\mathcal{Q}_{\text{th}}$ by comparing the full enzymatic cycles to their respective RC limits. To achieve this, we define a collective rate constant k which governs the product output current $J_p^R \sim k^n$ in RC networks, and from which \mathcal{Q} and other performance metrics are calculated as detailed in Sec. II. Consequently, k serves as a control parameter that allows for direct performance comparison between actual and idealized RC systems. Figure 2 presents J_p^R as a function of k for both RCs and the full ribosome and polymerase systems. As seen, the ideal current is increasing monotonically with k , while the actual current is nonmonotonic for the ribosome systems and nearly indistinguishable from RC for the T7 polymerase. These results are a first indication that the polymerase is indeed working at virtually reduced network conditions, hence the lower \mathcal{Q} value, while the ribosomes only approach this limit at longer operation times (lower currents).

The proximity of these systems to their lower TUR limits motivates a further examination of the full circuits and their corresponding RC limits. To this end, we consider two operating cases that either reduce or preserve the full reaction network, respectively: (a) perfect binding discrimination corresponding to $f_1 \rightarrow 0$ and (b) perfect proofreading discrimination corresponding to $f_3 \rightarrow \infty$. In case (a) the full system gradually reduces to the RC limit by entire omission of the

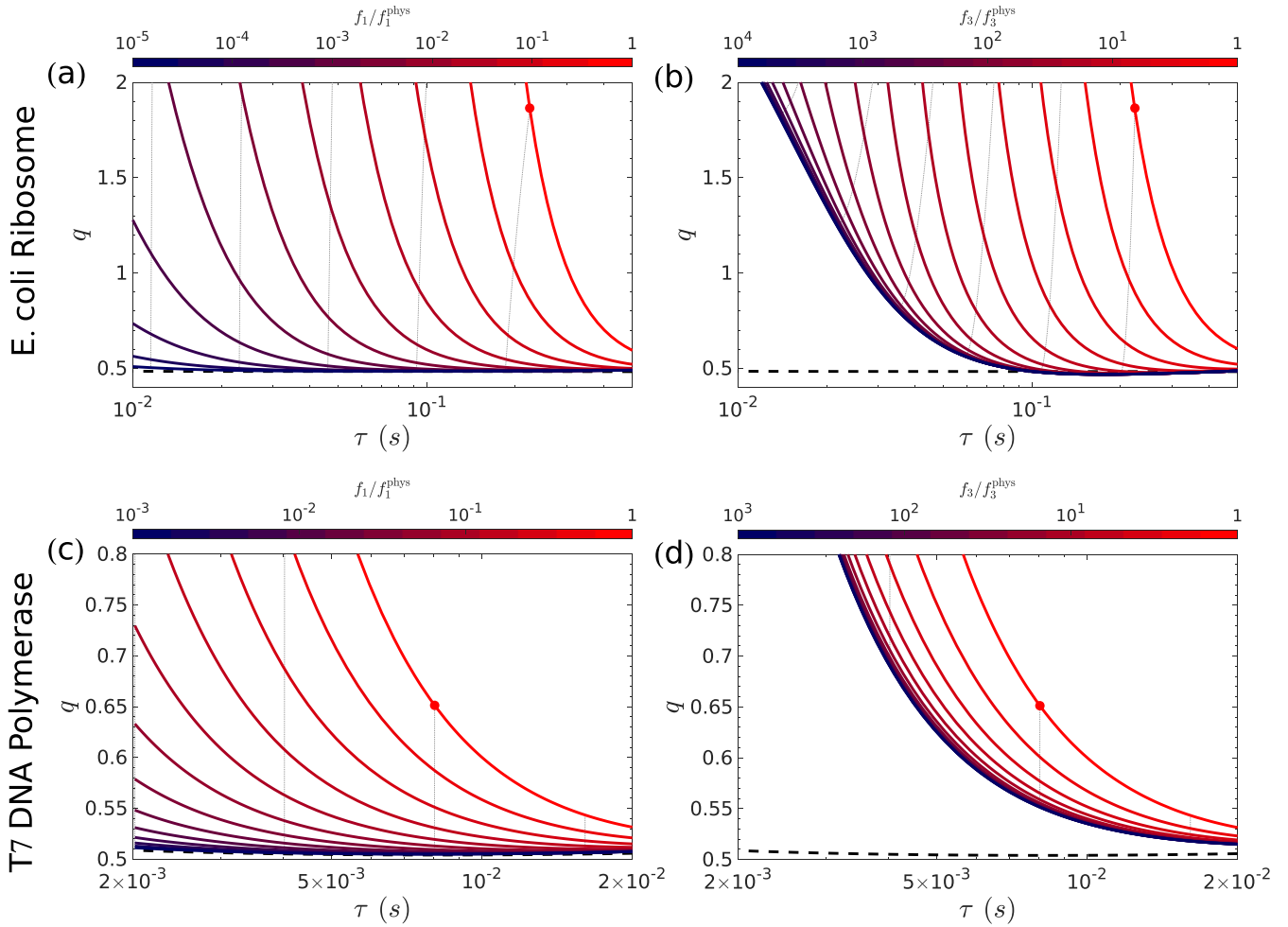


FIG. 3. Parametric plots of the normalized TUR measure $q \equiv Q/\Delta\mu_p$ vs. mean production time τ (s) as a function of k for decreasing values of discrimination factor f_1 (a,c) and increasing values of the proofreading discrimination factor f_3 (b,d) for the wild-type ribosome and DNA polymerase, respectively. Each corresponding factor f_i is scaled from its physiological value f_i^{phys} as shown in the gradient scale on top. Dashed black lines indicate ideal RC limit. Dotted gray lines indicate curves of constant k scaled from its physiological value in powers of 2.

W branch. In contrast, case (b) preserves the overall system topology while minimizing the impact of incorrect synthesis in the W branch. As shown below, comparing Q , the error rate η and the mean production time τ for either case allows us to see how the approach to the RC limit governs the performance of the full systems.

B. Approaching the RC limit decouples the TUR measure Q from the mean production time τ

To allow for normalized comparison of the actual system over its RC cycle, independent of system topology and energetic cost, we define the normalized TUR measure $q \equiv Q/\Delta\mu_p$. Figure 3 shows q for the WT ribosome against the mean production time τ , as a function of k while f_1 [case (a)] or f_3 [case (b)] are parametrically varied. As seen in (a) the WT ribosome displays a clear trade-off between q and τ (red line), but quickly attenuates and decouples as it approaches the RC limit (black dashed line). On the other hand, increasing f_3 (b) maintains the trade-off constraint between q and τ at all points, even when approaching the RC limit. Similar trends are seen for the polymerase in Fig. 3(c) and 3(d). From these

results we find that a system operating near the RC limit may more readily minimize both the product output noise and mean production time without being constrained by a strong trade-off relation.

C. The error rate η decouples from the TUR measure Q in high fidelity regimes

Figure 4 shows $q \equiv Q/\Delta\mu_p$ against the error η for decreasing f_1 (a) and increasing f_3 (b) from the measured physiological values. In both cases, q decreases with decreasing η and becomes decoupled in the low error regime. However, this asymptotic value of Q only matches the RC limit in the vanishing f_1 case (a) but not in case (b) where the Err mutant stays well above the RC value at the $f_3 \rightarrow \infty$ limit. The inset illustrates this asymptotic behavior more clearly where in this case k has been rescaled from the physiological values as $k = 10k_1^{\text{phys}}$ for the WT ribosome, and shows that the RC limit can be reached for $f_1 \rightarrow 0$, but not for $f_3 \rightarrow \infty$.

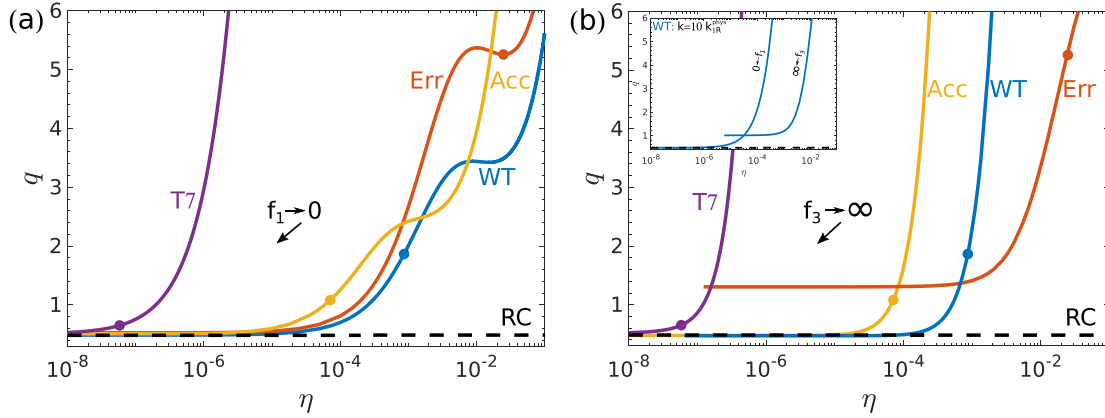


FIG. 4. (a) Parametric plot of $q \equiv Q/\Delta\mu_p$ vs. error η as a function of f_1 . Lines indicate T7 DNA polymerase (purple leftmost), WT ribosome (blue bottom), erroneous (red top), and more accurate (yellow middle) ribosome mutants. Thick points indicate physiological values. Dashed line is a guide to the eye showing the value of Q achieved at the ideal RC limit of WT ribosome which is approximately the same for all systems shown. (b) Same as (a) but for $f_3 \rightarrow \infty$. Inset: WT ribosome scaled from $k = k_{1R}^{\text{phys}}$ to $k = 10k_{1R}^{\text{phys}}$ illustrating that only $f_1 \rightarrow 0$ guarantees Q goes to the RC limit.

D. Approaching the RC limit relaxes the trade-off constraint between error rate and mean production time

It is also instructive to compare the error η to the mean production time τ as a function of k in the context of the idealized RC limit (Fig. 5). While changing either f_1 or f_3 parametrically is not expected to decouple the trade-off between measures, these curves highlight improved performance close to the RC limit in addition of minimizing Q . For instance, while increasing discriminant proofreading f_3 naturally improves the accuracy of the system, it ultimately approaches a best trade-off curve for this parameter variation (a Pareto front). In contrast, reducing f_1 weakens the trade-off relationship (smaller negative derivatives) while moving these trade-off curves arbitrarily close to the origin by construction (incorrect substrate is never bound).

E. The energy cost rate for faster speed of operation is minimized in the RC limit

Lastly, we consider energy dissipation in the limit of the RC cycle. While Q in general provides an efficiency measure of dissipation and product output precision, it is independent of time and hence agnostic to the cost of driving a cycle up to a required speed of operation. In this regard, Mallory *et al.* [40] have shown that the ribosome and T7 polymerase prioritize speed over dissipation, and is therefore interesting to see how dissipation and mean production time vary between physiological systems and their corresponding RC limits. In particular, we calculate the difference in energy dissipation between the actual systems and their RC limits. Figure 6 shows the normalized dissipation rate difference, $\dot{q} - \dot{q}_{RC} \equiv k_B T (\sigma - \sigma_{RC})/\Delta\mu_p$, against the mean production time τ as a

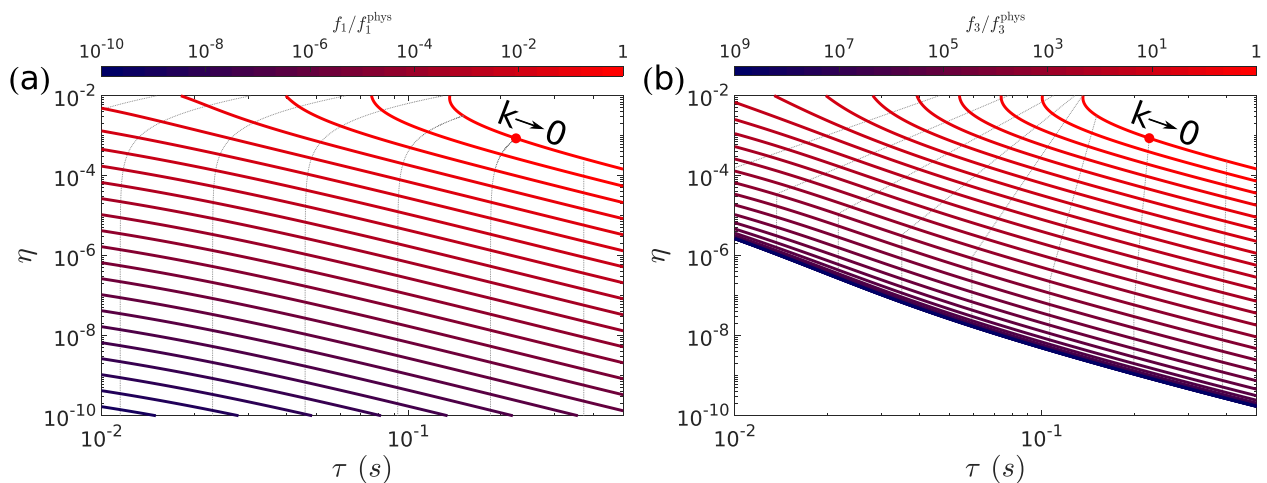


FIG. 5. Parametric plots of the error η versus mean production time τ for wild-type ribosome as a function of k for (a) decreasing values of binding discrimination f_1 , and (b) increasing values of proofreading discrimination f_3 . Each respective factor f_i is scaled from its physiological value f_i^{phys} as shown in the gradient scale on top. Dotted gray lines indicate curves of constant k value scaled from its physiological value in powers of 2.

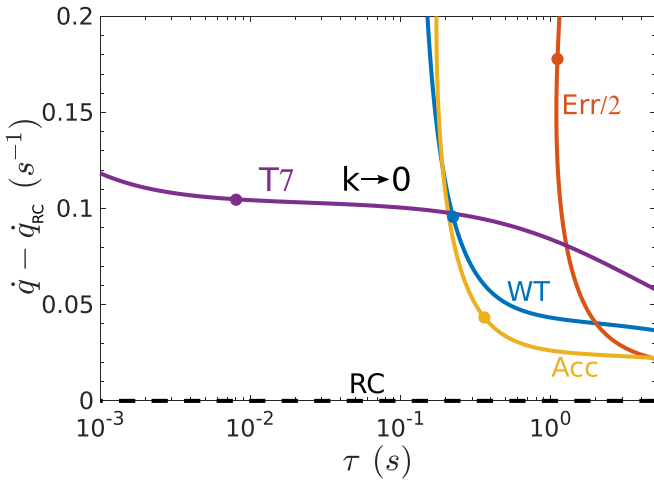


FIG. 6. Parametric plots of normalized dissipation differences between the actual $\dot{q} = \dot{Q}/\Delta\mu_p$ and the ideal RC limit $\dot{q}_{RC} = \dot{Q}_{RC}/\Delta\mu_p$ versus the mean production time τ as a function of k . Lines indicate T7 DNA polymerase (purple left), WT ribosome (blue middle), erroneous (red top), and more accurate (yellow bottom) ribosome mutants. Err mutant plot has been scaled down by a factor of 2 to fit the figure. Points indicate physiological values for each line respectively. Dashed line marks the RC difference which is zero by definition.

parametric function of k for ribosomes and the T7 polymerase. The dissipation rate was normalized by the operating energy cost $\Delta\mu_p$ to allow comparison of different reaction networks. Evidently, while the ribosomes display absolute differences lower than the polymerase, they operate more slowly by two orders of magnitude and with steep energy costs for τ shorter than physiological values. On the other hand, the T7 polymerase maintains a relatively flat profile over many τ decades, ensuring that the energy dissipation rate does not deviate strongly from the ideal RC values, which achieve minimal \dot{Q} by construction.

In closing, by all metrics considered, operating near the the RC limit confers considerable performance advantages to the KPR systems examined. By this measure, it is not surprising the polymerase outperforms the ribosomes given that its binding discrimination factor f_1 is about a million times more restrictive than that of the ribosomes ($f_{1,\text{polymerase}}/f_{1,\text{ribosome}} \sim 10^{-6}$) and places it significantly closer to the underlying RC limit. Note that a low \dot{Q} score *does not* imply by itself the RC limit; low values of \dot{Q} are achieved for certain limiting values of f_3 , and as discussed previously, this does not confer the similar trade-off advantages of approaching the RC limit. For instance, the Acc mutant achieves lower \dot{Q} score due to its enhanced f_3 , but must operate at slower production times than the WT due to steeply increasing energy demands as seen in Fig. 6. As a result, operating near the RC limit not only achieves low \dot{Q}/\dot{Q}_{th} score by definition, but also improves the

overall *global* performance per production cycle in a reaction network given a fixed energy budget.

IV. CONCLUSIONS

The ribosome and the DNA polymerase drive two essential production networks in the cell. The efficiency of these circuits is an important determinant of the organism fitness, and therefore they must be tuned to prioritize product-forming transitions over competing incorrect substrate binding and proofreading cycles. In this work, we have analyzed these circuits in the light of the Thermodynamic Uncertainty Relation (TUR), and found that the TUR measure \dot{Q} for the product current is closer to the lower bound in the polymerase than in the E. coli ribosome system. In particular, we considered a reduced cycle (RC) limit that accounts for paths leading to product formation, and showed that operating near this regime affords minimized values of \dot{Q} for corresponding rate constants. Notably, the polymerase operates very near the RC regime and thereby achieves nearly optimal performance manifested by the proximity of its \dot{Q} measure to the lower bound \dot{Q}_{th} [Eq. (9) and Table I]. Further, operating near RC relaxes the trade-off constraint between accuracy and speed, while decoupling both these measures from \dot{Q} . On the other hand, a similar analysis showed that E. coli ribosomes operate relatively farther away from the RC limit, resulting in stronger coupling across all performance measures and increased energy costs, manifested by larger values of \dot{Q} . That said, the ribosome is not more than one order-of-magnitude away from the TUR bound. The significant difference in the performance of polymerase and ribosome stems from the accuracy of substrate discrimination, which is higher by about six orders of magnitude in the polymerase. This binding selectivity difference, which is not directly addressed here, is linked to the different biochemical mechanisms employed by the polymerase and the ribosome [42,47,48]. As a result, the polymerase is more likely to operate in the regime of correct product cycles than the ribosome, close to the RC limit. The different regimes of performance may also reflect the much more deleterious impact of errors in replication, which are carried through genome heredity, relative to errors in translation that vanish when the protein is degraded. For future studies, it would be interesting to study how distinct reaction pathways in other protein systems, e.g., in signal transduction, adjust to prioritize correct response cycles and whether these imply similar RC limits that optimize the underlying TUR constraint.

ACKNOWLEDGMENTS

The authors thank A. B. Kolomeisky, O. A. Igoshin, and C. Hyeon for helpful discussions. This work was supported by the taxpayers of South Korea through the Institute for Basic Science, Project Code IBS-R020-D1.

[1] P. Edelman and J. Gallant, *Cell* **10**, 131 (1977).
 [2] A. Drummond and C. O. D. Wilke, *Nat. Rev. Genet.* **10**, 715 (2009).

[3] E. B. Kramer and P. J. Farabaugh, *RNA* **13**, 87 (2007).
 [4] J.-F. Gout, W. Li, C. Fritsch, A. Li, S. Haroon, L. Singh, D. Hua, H. Fazelinia, Z. Smith, S. Seeholzer, K.

- Thomas, M. Lynch, and M. Vermulst, *Sci. Adv.* **3**, e1701484 (2017).
- [5] A. Blank, J. A. Gallant, R. R. Burgess, and L. A. Loeb, *Biochemistry* **25**, 5920 (1986).
- [6] C. C. Traverse and H. Ochman, *Proc. Natl. Acad. Sci. USA* **113**, 3311 (2016).
- [7] J. A. S. Charles, S. E. Liberti, J. S. Williams, S. A. Lujan, and T. A. Kunkel, *DNA Repair* **31**, 41 (2015).
- [8] R. M. Schaafer, *J. Biol. Chem.* **268**, 23762 (1993).
- [9] J. W. Drake, B. Charlesworth, D. Charlesworth, and J. F. Crow, *Genetics* **148**, 1667 (1998).
- [10] J. J. Hopfield, *Proc. Natl. Acad. Sci. USA* **71**, 4135 (1974).
- [11] J. Ninio, *Biochimie* **57**, 587 (1975).
- [12] T. A. Kunkel and K. Bebenek, *Annu. Rev. Biochem.* **69**, 497 (2000).
- [13] R. C. Thompson and P. J. Stone, *Proc. Natl. Acad. Sci. USA* **74**, 198 (1977).
- [14] K. B. Gromadski and M. V. Rodnina, *Mol. Cell* **13**, 191 (2004).
- [15] W. S. Hlavacek, A. Redondo, H. Metzger, C. Wofsy, and B. Goldstein, *Proc. Natl. Acad. Sci. USA* **98**, 7295 (2001).
- [16] B. Goldstein, D. Coombs, J. R. Faeder, and W. S. Hlavacek, Kinetic Proofreading Model, in *Multichain Immune Recognition Receptor Signaling: From Spatiotemporal Organization to Human Disease*, edited by A. B. Sigalov (Springer, New York, 2008), pp. 82–94.
- [17] D. Sagi, T. Tlusty, and J. Stavans, *Nucleic Acids Res.* **34**, 5021 (2006).
- [18] F. S. Gnesotto, F. Mura, J. Gladrow, and C. P. Broedersz, *Rep. Prog. Phys.* **81**, 066601 (2018).
- [19] A. I. Brown and D. A. Sivak, *Chem. Rev.* **120**, 434 (2020).
- [20] C. Jarzynski, *Annu. Rev. Condens. Matter Phys.* **2**, 329 (2011).
- [21] R. Rao and L. Peliti, *J. Stat. Mech.: Theory Exp.* (2015) P06001.
- [22] P. Sartori and S. Pigolotti, *Phys. Rev. X* **5**, 041039 (2015).
- [23] A. C. Barato and U. Seifert, *Phys. Rev. Lett.* **114**, 158101 (2015).
- [24] C. Hyeon and W. Hwang, *Phys. Rev. E* **96**, 012156 (2017).
- [25] T. R. Gingrich, J. M. Horowitz, N. Perunov, and J. L. England, *Phys. Rev. Lett.* **116**, 120601 (2016).
- [26] J. M. Horowitz and T. R. Gingrich, *Phys. Rev. E* **96**, 020103(R) (2017).
- [27] Y. Hasegawa and T. Van Vu, *Phys. Rev. Lett.* **123**, 110602 (2019).
- [28] A. C. Barato and U. Seifert, *J. Phys. Chem. B* **119**, 6555 (2015).
- [29] P. Pietzonka, A. C. Barato, and U. Seifert, *J. Stat. Mech.: Theory Exp.* (2016) 124004.
- [30] U. Seifert, *Physica A: Statistical Mechanics and its Applications* **504**, 176 (2018), lecture Notes of the 14th International Summer School on Fundamental Problems in Statistical Physics.
- [31] W. Hwang and C. Hyeon, *J. Phys. Chem. Lett.* **9**, 513 (2018).
- [32] D. S. Tawfik, *Curr. Opin. Chem. Biol.* **21**, 73 (2014).
- [33] A. Murugan, D. A. Huse, and S. Leibler, *Proc. Natl. Acad. Sci. USA* **109**, 12034 (2012).
- [34] A. Murugan, D. A. Huse, and S. Leibler, *Phys. Rev. X* **4**, 021016 (2014).
- [35] T. E. Ouldridge, C. C. Govern, and P. R. ten Wolde, *Phys. Rev. X* **7**, 021004 (2017).
- [36] F. Wong, A. Amir, and J. Gunawardena, *Phys. Rev. E* **98**, 012420 (2018).
- [37] T. Tlusty, R. Bar-Ziv, and A. Libchaber, *Phys. Rev. Lett.* **93**, 258103 (2004).
- [38] R. Bar-Ziv, T. Tlusty, and A. Libchaber, *Proc. Natl. Acad. Sci. USA* **99**, 11589 (2002).
- [39] K. Banerjee, A. B. Kolomeisky, and O. A. Igoshin, *Proc. Natl. Acad. Sci. USA* **114**, 5183 (2017).
- [40] J. D. Mallory, A. B. Kolomeisky, and O. A. Igoshin, *J. Phys. Chem. B* **123**, 4718 (2019).
- [41] H. S. Zaher and R. Green, *Mol. Cell* **39**, 110 (2010).
- [42] I. Wong, S. S. Patel, and K. A. Johnson, *Biochemistry* **30**, 526 (1991).
- [43] M. J. Donlin, S. S. Patel, and K. A. Johnson, *Biochemistry* **30**, 538 (1991).
- [44] S. S. Patel, I. Wong, and K. A. Johnson, *Biochemistry* **30**, 511 (1991).
- [45] Z. Koza, *J. Phys. A: Math. Gen.* **32**, 7637 (1999).
- [46] P. Pietzonka, A. C. Barato, and U. Seifert, *J. Phys. A: Math. Theor.* **49**, 34LT01 (2016).
- [47] K. B. Gromadski, T. Daviter, and M. V. Rodnina, *Mol. Cell* **21**, 369 (2006).
- [48] Y. Savir and T. Tlusty, *Cell* **153**, 471 (2013).