

# Supplementary Material: An energy-speed-accuracy relation in complex networks for biological discrimination

Felix Wong,<sup>1,2</sup> Ariel Amir,<sup>1</sup> and Jeremy Gunawardena<sup>2,\*</sup>

<sup>1</sup>*School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA*

<sup>2</sup>*Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA*

We provide proofs here of the mathematical assertions made in the main text.

## I. EQUILIBRIUM ERROR FRACTION FOR THE HOPFIELD MECHANISM

The error fraction,  $\varepsilon$ , for the Hopfield mechanism is given in equation (4) of the main text, and is repeated here for convenience,

$$\varepsilon = \frac{[l'_D(k_D + m') + m'k'_D][(k_C + m')(W + l_C) + mk_C]}{[l'_C(k_C + m') + m'k'_C][(k_D + m')(W + l_D) + mk_D]} \quad (1)$$

The background assumptions, as mentioned in the main text, are  $k'_C = k'_D$ ,  $l'_C = l'_D$  and  $k_C/k_D < 1$ .

If the mechanism is at thermodynamic equilibrium, then detailed balance must be satisfied. The equivalent cycle condition [1] applied to the two cycles in Fig. 1(a) of the main text yields

$$\frac{m'}{m} = \frac{l'_C k_C}{l'_D k'_C} = \frac{l'_D k_D}{l'_C k'_D} \quad (2)$$

Note that  $W$  does not appear in equation (2) since, although the mechanism itself is at thermodynamic equilibrium, the system remains open, with substrate being converted to product. Denote by  $\varepsilon_{eq}$  the value of  $\varepsilon$  under the equilibrium constraint in equation (2). Using equation (2), define  $\alpha, \beta$  so that

$$\alpha = \frac{l_C}{l'_C} = \frac{mk_C}{m'k'_C} \quad \text{and} \quad \beta = \frac{l_D}{l'_D} = \frac{mk_D}{m'k'_D}.$$

Using the background assumptions, define the quantity  $\varepsilon_0$ , given by

$$\varepsilon_0 = \frac{\alpha}{\beta} = \frac{l_C}{l_D} = \frac{k_C}{k_D}, \quad (3)$$

to which a physical interpretation will be given shortly. Substituting  $\alpha$  and  $\beta$  into equation (1) and rewriting, we see that

$$\varepsilon_{eq} = \frac{W.A + \alpha}{W.B + \beta},$$

where

$$A = \frac{k_C + m'}{(k_C + m')l'_C + m'k'_C}, \quad B = \frac{k_D + m'}{(k_D + m')l'_D + m'k'_D}.$$

If  $W = 0$ , then by equation (3),  $\varepsilon_{eq} = \varepsilon_0$ , which shows that  $\varepsilon_0$ , as defined in equation (3), is the error fraction for the closed system at thermodynamic equilibrium as defined in the main text.

We now want to prove that  $\varepsilon_{eq}$  increases from  $\varepsilon_0$  as  $W$  increases, for which it is sufficient to show that  $d\varepsilon_{eq}/dW > 0$ . For this,

$$\frac{d\varepsilon_{eq}}{dW} = \frac{A\beta - B\alpha}{(W.B + \beta)^2},$$

so that  $d\varepsilon_{eq}/dW > 0$  if, and only if,  $A/B > \alpha/\beta$ . We have

$$\frac{A}{B} = \left( \frac{k_C + m'}{k_D + m'} \right) \left( \frac{(k_D + m')l'_D + m'k'_D}{(k_C + m')l'_C + m'k'_C} \right). \quad (4)$$

The following result is straightforward.

**Lemma.** Consider the rational function  $r(x) = (a + x)/(b + x)$ , where  $a, b > 0$ . If  $a/b > 1$ , then  $r(x)$  decreases strictly monotonically from  $r(0) = a/b$  to 1. If  $a/b < 1$ , then  $r(x)$  increases strictly monotonically from  $r(0) = a/b$  to 1.

Applying the Lemma repeatedly to the terms in equation (4), and recalling the background assumptions, we see that

$$\frac{k_C + m'}{k_D + m'} > \frac{k_C}{k_D} \quad \text{and} \quad \frac{(k_D + m')l'_D + m'k'_D}{(k_C + m')l'_C + m'k'_C} > 1.$$

Hence,  $A/B > k_C/k_D = \alpha/\beta$  and so  $d\varepsilon_{eq}/dW > 0$ . It follows that  $\varepsilon_{eq}$  increases strictly monotonically from  $\varepsilon_0$  as  $W$  increases from 0.

## II. DERIVATION OF EQUATION (5) OF THE MAIN TEXT

The non-equilibrium error fraction in equation (1) can be rewritten as  $\varepsilon(m') = u(m')v(m')$ , where

$$u(m') = \frac{[l'_D k_D + (l'_D + k'_D)m']}{[l'_C k_C + (l'_C + k'_C)m']}, \quad \text{and}$$

$$v(m') = \frac{[k_C(l_C + m + W) + (l_C + W)m']}{[k_D(l_D + m + W) + (l_D + W)m']}.$$

Using the background assumptions and the Lemma, we see that, as  $m'$  increases,  $u(m')$  decreases hyperbolically from  $u(0) = k_D/k_C = \varepsilon_0^{-1}$  to 1 while  $v(m')$  increases hyperbolically between

$$v(0) = \varepsilon_0 \left( \frac{l_C + m + W}{l_D + m + W} \right) \quad \text{and} \quad v(\infty) = \left( \frac{l_C + W}{l_D + W} \right).$$

\* Corresponding author: jeremy@hms.harvard.edu

Hence,  $\varepsilon(0) = (l_C + m + W)/(l_D + m + W)$  and  $\varepsilon(m') \rightarrow (l_C + W)/(l_D + W)$  as  $m' \rightarrow \infty$ . Furthermore,

$$\varepsilon(m') > v(0) = \varepsilon_0 \left( \frac{l_C + m + W}{l_D + m + W} \right) > \varepsilon_0^2, \quad (5)$$

as required for equation (5) in the main text.

### III. THE LIMITING ARGUMENT IN KINETIC PROOFREADING

The argument for kinetic proofreading given by Hopfield [2] is based on the non-dimensional quantities,

$$\delta_1 = \frac{l'_D(m' + k_D)}{m'k'_D}, \quad \delta_2 = \frac{m'}{k_C}, \quad \delta_3 = \frac{m}{l_D}, \quad \delta_4 = \frac{W}{l_D},$$

which are to be taken very small. Accordingly, we consider the non-equilibrium error-fraction,  $\varepsilon$ , as defined in equation (1), in the limit as these four quantities  $\rightarrow 0$ . Since  $k_D > k_C$ , we have that

$$\delta_1 > \frac{l'_C(m' + k_C)}{m'k'_C} > 0. \quad (6)$$

If we now divide above and below by  $m'k'_D = m'k'_C$  in the expression  $\varepsilon = uv$  introduced above, we get

$$\varepsilon = \left( \frac{\delta_1 + 1}{l'_C(m' + k_C)/m'k'_C + 1} \right) v.$$

If we take  $\delta_1 \rightarrow 0$  and formally treat  $v$  as a constant, we see from equation (6) that  $\varepsilon \rightarrow v$  as  $\delta_1 \rightarrow 0$ . However, the expression for  $\delta_1$  involves  $m'$  and this is also a parameter in  $v$ . Hence,  $v$  is not constant during the limiting process, which has coupled  $m'$  to the values of the other parameters. If we ignore this coupling, we can divide above and below in  $v$  by  $k_C$  to get

$$v = \frac{(1 + \delta_2)(W + l_C) + m}{(k_D/k_C + \delta_2)(W + l_D) + mk_D/k_C}$$

and if then let  $\delta_2 \rightarrow 0$ , we see that

$$v \rightarrow \varepsilon_0 \left( \frac{W + l_C + m}{W + l_D + m} \right).$$

If we now divide above and below in this expression by  $l_D$ , we get

$$\frac{W/l_D + l_C/l_D + m/l_D}{W/l_D + 1 + m/l_D} \rightarrow \varepsilon_0^2$$

as  $\delta_3 \rightarrow 0$  and  $\delta_4 \rightarrow 0$ . Hence, putting the sequence of limits together, we have, formally,

$$\lim_{\delta_4 \rightarrow 0} \lim_{\delta_3 \rightarrow 0} \lim_{\delta_2 \rightarrow 0} \lim_{\delta_1 \rightarrow 0} \varepsilon = \varepsilon_0^2.$$

This seems to be the interpretation that has been given in the literature to Hopfield's assertion that kinetic proofreading achieves the error fraction of  $\varepsilon_0^2$ .

The coupling noted above specifically affects  $m'$ , which has to satisfy two conditions. On the one hand  $m'$  has to be large, in order that  $u$  should be close to  $u(\infty) = 1$ . That is the role of  $\delta_1 \rightarrow 0$ . On the other hand  $m'$  has to be small, in order that  $v$  should be close to  $v(0)$ . That is the role of  $\delta_2 \rightarrow 0$ . The remaining limits for  $\delta_3$  and  $\delta_4$  are only there to make sure that  $v(0) \rightarrow \varepsilon_0^2$ ; compare equation (5). The consequence of the coupling between the  $\delta_1$  and  $\delta_2$  limits, which arises through  $m'$ , can be seen by rewriting  $\delta_1$ ,

$$\delta_1 = \left( \frac{l'_D}{k'_D} \right) \left( 1 + \left( \frac{k_D}{k_C} \right) \frac{1}{\delta_2} \right).$$

Hence, in the limit as  $\delta_1 \rightarrow 0$  and  $\delta_2 \rightarrow 0$ ,

$$\frac{l'_D}{k'_D} = \frac{l'_C}{k'_C} \rightarrow 0.$$

In order to achieve the proofreading limit, it is necessary for rates other than  $m'$  to change. Specifically, the “on rates” for the first discrimination,  $k'_C = k'_D$ , must become large with respect to those for the second discrimination,  $l'_C = l'_D$ .

### IV. DERIVATION OF EQUATION (7) OF THE MAIN TEXT

Suppose that  $R(x)$  and  $Q(x)$  are allowable functions, as defined in the main text, with  $R(x)/x^\alpha \rightarrow c_1$  and  $Q(x)/x^\beta \rightarrow c_2$ , as  $x \rightarrow \infty$ , where  $c_1, c_2 > 0$ .

Since  $R^{-1}/x^{-\alpha} \rightarrow (c_1)^{-1} > 0$ , it follows that  $R^{-1}$  is allowable and  $\deg(R^{-1}) = -\deg(R)$ . Since  $(RQ)/x^{\alpha+\beta} \rightarrow c_1c_2 > 0$ , it follows that  $RQ$  is allowable and  $\deg(RQ) = \deg(R) + \deg(Q)$ . Finally, suppose, without loss of generality, that  $\max(\alpha, \beta) = \alpha$ , so that  $\alpha \geq \beta$ . Then,

$$\frac{R+Q}{x^\alpha} = \left( \frac{R}{x^\alpha} \right) + \left( \frac{Q}{x^\beta} \right) x^{\beta-\alpha}. \quad (7)$$

The limit of this, as  $x \rightarrow \infty$ , is  $c_1 > 0$ , if  $\alpha > \beta$ , or  $c_1 + c_2 > 0$ , if  $\alpha = \beta$ . In either case, the limit is positive. Hence,  $R+Q$  is allowable and  $\deg(R+Q) = \max(\deg(R), \deg(Q))$ . This proves equation (7) of the main text.

### V. DERIVATION OF EQUATION (14) OF THE MAIN TEXT

Suppose that  $S(x)$  is an allowable function and that  $S/x^\alpha \rightarrow c > 0$ , where  $\alpha = \deg(S)$ . Since  $\ln$  is a continuous function,

$$\ln(S(x)) - \alpha \ln(x) \rightarrow \ln(c).$$

Dividing through by  $\ln(x)$ , we see that

$$\frac{\ln(S(x))}{\ln(x)} \rightarrow \alpha. \quad (8)$$

If  $\deg(S) = \alpha > 0$ , then  $\ln(S) \sim \ln(x)$ , while if  $\deg(S) < 0$  then  $\ln(S) \sim \ln(x^{-1})$ , which proves equation (14) of the main text.

## VI. DERIVATION OF EQUATION (15) OF THE MAIN TEXT

Note that if  $R(x), Q(x), S(x), T(x)$  are functions, not necessarily allowable, and if  $R \sim Q$  and  $S \sim T$ , so that  $R/Q \rightarrow c_1 > 0$  and  $S/T \rightarrow c_2 > 0$ , then  $(RS)/(QT) \rightarrow c_1 c_2 > 0$ , so that  $RS \sim QT$ . We will use this without reference below.

Following the discussion in the main text, consider  $P(i \rightleftharpoons j) = (R - Q) \ln(R/Q)$  where  $R$  and  $Q$  are allowable functions with  $R/x^\alpha \rightarrow c_1 > 0$  and  $Q/x^\beta \rightarrow c_2 > 0$ . There are three cases to consider.

Suppose that  $\alpha \neq \beta$ . If  $\alpha > \beta$ , then the same argument as in equation (7) shows that  $(R - Q) \sim x^\alpha$ . By equation (7) of the main text,  $\deg(R/Q) > 1$ , so that  $\ln(R/Q) \sim \ln(x)$ . Hence,  $P(i \rightleftharpoons j) \sim x^\alpha \ln(x)$ . If  $\alpha < \beta$ , then  $(R - Q) \sim -x^\beta$  and  $\ln(R/Q) \sim -\ln(x)$ , so that  $P(i \rightleftharpoons j) \sim x^\beta \ln(x)$ . This proves case 1.

Suppose that  $\alpha = \beta$  but  $c_1 \neq c_2$ . Then,  $(R - Q)/x^\alpha \rightarrow c_1 - c_2 \neq 0$ . Also,

$$\frac{R}{Q} = \left( \frac{R}{x^\alpha} \right) \left( \frac{x^\alpha}{Q} \right) \rightarrow \frac{c_1}{c_2} > 0.$$

Hence,  $\ln(R/Q) \rightarrow \ln(c_1/c_2)$ . Since  $(c_1 - c_2) \ln(c_1/c_2) > 0$ , it follows that  $P(i \rightleftharpoons j) \sim x^\alpha$ , which proves case 2.

Suppose that  $\alpha = \beta$  and  $c_1 = c_2$ . Then  $(R - Q)/x^\alpha \rightarrow 0$  and  $R/Q \rightarrow 1$ , so that  $\ln(R/Q) \rightarrow 0$ . Hence,  $P(i \rightleftharpoons j)/x^\alpha \rightarrow 0$  as  $x \rightarrow \infty$ , so that  $P(i \rightleftharpoons j) \prec x^\alpha$ , which proves case 3.

## VII. DERIVATION OF EQUATION (20) OF THE MAIN TEXT

For the Hopfield mechanism (Fig. 1(a) of the main text), we described in the main text how the asymptotic error rate of  $\varepsilon \sim x^{-2}$  could be achieved, by assuming that the labels are allowable functions of  $x$  such that:  $\deg(l'_D) = \deg(l'_C) = -1$ ,  $\deg(k_D) = \deg(l_D) = 1$  and  $\deg(k'_C) = \deg(k_C) = \deg(k'_D) = \deg(m') = \deg(m) = \deg(l_C) = \deg(W) = 0$ . Using equations (1) and (3) of the main text, we find that  $\deg(p_1^*) = \deg(p_{2_C}^*) = \deg(p_{3_C}^*) = 0$ ,  $\deg(p_{2_D}^*) = -1$ , and  $\deg(p_{3_D}^*) = -2$ . It is helpful to introduce the notation  $R \approx Q$ , for functions  $R(x), Q(x)$  which may not be allowable, to signify that  $\lim_{x \rightarrow \infty} R/Q = 1$ . We can use this to calculate the asymptotic behaviour of the terms  $P(i \rightleftharpoons j)$  in the entropy production rate (equation (12) of the main text). For instance,

$$P(1 \rightleftharpoons 3_C) = ((l_C + W)p_{3_C}^* - l'_C p_1^*) \ln \left( \frac{(l_C + W)p_{3_C}^*}{l'_C p_1^*} \right),$$

The only term in this expression which depends on  $x$  is  $l'_C$  for which  $\deg(l'_C) = -1$ . Since  $\deg(\ln(R)) \approx \deg(R) \ln(x)$  (equation (8)), it follows that

$$P(1 \rightleftharpoons 3_C) \approx (l_C + W)p_{3_C}^* \ln(x).$$

Similar calculations yield  $P(2_C \rightleftharpoons 3_C) \approx C_1, P(1_C \rightleftharpoons 2_C) \approx C_2, P(1 \rightleftharpoons 3_D) \lesssim C_3, P(2_D \rightleftharpoons 3_D) \lesssim C_4$ , and  $P(1_D \rightleftharpoons 2_D) \approx C_5$ , where  $C_1, C_2, C_3, C_4$ , and

$C_5$  are constants independent of  $x$ . Since  $\deg(\varepsilon) = -2$ ,  $\ln(x) \approx \ln(\varepsilon^{-1})/2$ . Hence,

$$P \approx \frac{(l_C + W)}{2} p_{3_C}^* \ln(\varepsilon^{-1}),$$

so that

$$\lim_{x \rightarrow \infty} \frac{P}{\sigma \ln(\varepsilon^{-1})} = \frac{l_C + W}{2W}. \quad (9)$$

This proves equation (20) of the main text.

## VIII. ADDITIONAL NUMERICAL CALCULATIONS

In Supplementary Figs. 1 and 2, we consider two discrimination mechanisms under the assumptions of equations (21) and (22) of the main text. Supplementary Fig. 1(a) shows a graph for McKeithan's T-cell receptor mechanism [3], while Supplementary Fig. 2(a) shows a graph different from both this and the Hopfield example. We used previously developed, freely-available software [4] to compute the Matrix-Tree formula (equation (1) of the main text) for each mechanism, from which we obtained symbolic expressions for  $P, \varepsilon$ , and  $\sigma$ . The graphs in Supplementary Figs. 1(a) and 2(a) have 441 and 64 spanning trees rooted at each vertex, respectively, underscoring the combinatorial complexity which arises away from equilibrium (main text, Discussion). (If a graph has reversible edges, so that  $i \rightarrow j$  if, and only if,  $j \rightarrow i$ , which is the case for all the graphs discussed here, there is a bijection between the sets of spanning trees rooted at any pair of distinct vertices.) Supplementary Figs. 1(b)-(c) and 2(b)-(c) show numerical plots undertaken in a similar way to those for the Hopfield mechanism (main text, Fig. 2), as described in the main text. Similar vertical and diagonal bounds were found for the symmetric cases, while similar observations regarding the asymmetric cases as those made in the main text apply.

## IX. ASYMPTOTIC RELATION FOR A NON-DISSOCIATION-BASED MECHANISM

We consider a discrimination mechanism having the graph shown in Supplementary Fig. 3(a). Its structure is identical to that of the Hopfield mechanism (Fig. 1(a) of the main text) but its labels differ to reflect the energy landscape illustrated in Supplementary Fig. 3(b). If the labels are allowable functions with  $\deg(l'_D) = -1$  and  $\deg(l'_C) = \deg(l_C) = \deg(l_D) = 0$ , then, if the mechanism reaches thermodynamic equilibrium, it follows from equation (9) of the main text that its equilibrium error fraction satisfies

$$\varepsilon_{eq} \sim x^{-1}. \quad (10)$$

If it is further assumed that  $\deg(m'_D) = \deg(m'_C) = -1$  and  $\deg(m_D) = -\deg(m_C) = 1/2$ , while all other labels have degree 0, then the mechanism is no longer at equilibrium. Using equations (1) and (3) of the main text, we find that

$$\begin{aligned}
\rho_1 &= [(m'_C + k)(l_C + W) + km_C][(m'_D + k)(l_D + W) + km_D] \\
\rho_{2_C} &= [m_C(l'_C + k') + k'(l_C + W)][(m'_D + k)(l_D + W) + km_D] \\
\rho_{3_C} &= [m'_C(k' + l'_C) + kl'_C][(k + m'_D)(l_D + W) + m_Dk] \\
\rho_{2_D} &= [(m'_C + k)(l_C + W) + km_C][m_D(l'_D + k') + k'(l_D + W)] \\
\rho_{3_D} &= [m'_D(k' + l'_D) + kl'_D][(k + m'_C)(l_C + W) + m_Ck].
\end{aligned}$$

It follows that

$$\deg(p_1^*) = \deg(p_{2_C}^*) = \deg(p_{3_C}^*) = \deg(p_{2_D}^*) = 0 \quad (11)$$

and that

$$\deg(p_{3_D}^*) = -3/2, \quad (12)$$

Using equations (7) and (14) of the main text along with equations (11) and (12), we can calculate the asymptotics of the terms in the entropy production rate  $P$  (equation (12) of the main text), assuming, as in the proof of the Theorem, that we are outside the measure-zero subset of parameter space arising

from case 3 of equation (15) of the main text. We find that

$$\begin{aligned}
P(1 \rightleftharpoons 2_C) &\sim 1 \\
P(2_C \rightleftharpoons 3_C) &\sim x^{-1/2} \ln(x) \\
P(1 \rightleftharpoons 3_C) &\sim 1 \\
P(1 \rightleftharpoons 2_D) &\sim 1 \\
P(2_D \rightleftharpoons 3_D) &\sim x^{-1} \ln(x) \\
P(1 \rightleftharpoons 3_D) &\sim x^{-1} \ln(x).
\end{aligned}$$

It follows that  $P \lesssim 1$ , so that the entropy production rate is asymptotically constant or vanishes. Furthermore, it can be shown from equations (11) and (12) that  $\deg(\varepsilon) = -3/2$  and  $\deg(\sigma) = 0$ . Hence, the error rate is asymptotically better than at equilibrium, for which  $\deg(\varepsilon_{eq}) = -1$  (equation (10)), while the speed remains asymptotically constant. This reflects a different asymptotic relation to that in equation (19) of the main text.

- 
- [1] Gunawardena, J. *A linear framework for time-scale separation in nonlinear biochemical systems*, PLoS ONE **7**, e36321 (2012).
  - [2] Hopfield, J. J. *Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity*, Proc. Natl. Acad. Sci. USA **71**, 4135 (1974).
  - [3] McKeithan, T. W. *Kinetic proofreading in T-cell receptor signal transduction*, Proc. Natl. Acad. Sci. USA **92**, 5042 (1995).
  - [4] Ahsendorf, T., Wong, F., Eils, R. & Gunawardena, J. *A framework for modelling gene regulation which accommodates non-equilibrium mechanisms*, BMC Biol. **12**, 102 (2014).
  - [5] Banerjee, K., Kolomeisky, A. B. and Igoshin, O. A. *Elucidating interplay of speed and accuracy in biological error correction*, Proc. Natl. Acad. Sci. USA **114**, 5183 (2017).

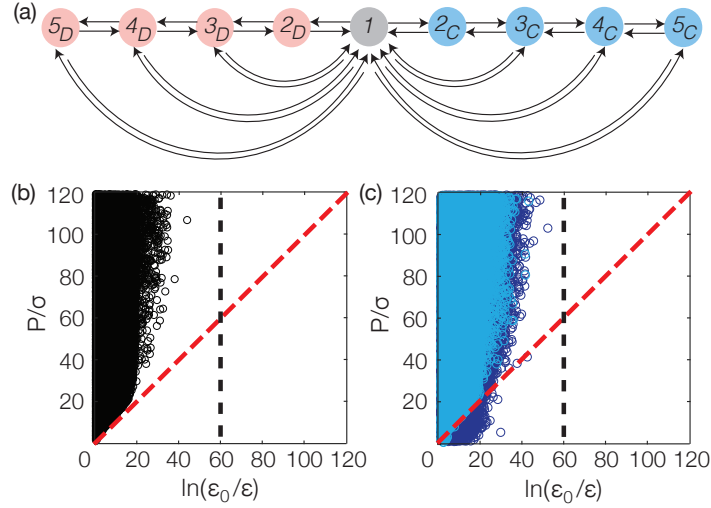


FIG. 1. Numerics for the T-cell receptor mechanism. (a) Graph for an instance of McKeithan's T-cell receptor mechanism [3], with label names omitted for clarity. (b) Plot of  $P/\sigma$  against  $\ln(\epsilon_0/\epsilon)$  for approximately  $10^5$  points, with the labels satisfying equations (21) and (22) of the main text and numerically sampled as described in the main text. The vertical black dashed line corresponds to the bound  $\epsilon > \epsilon_0^4$  for this mechanism (calculation not shown) that is analogous to equation (5) of the main text for the Hopfield mechanism. The diagonal red dashed line corresponds to equation (23) of the main text, as discussed further there. (c) Similar plot to (b) but with internal discrimination between correct and incorrect substrates, as described in the text, with the light blue points having a lower asymmetry range ( $A = 1$ ) and the dark blue points having a higher range ( $A = 5$ ).

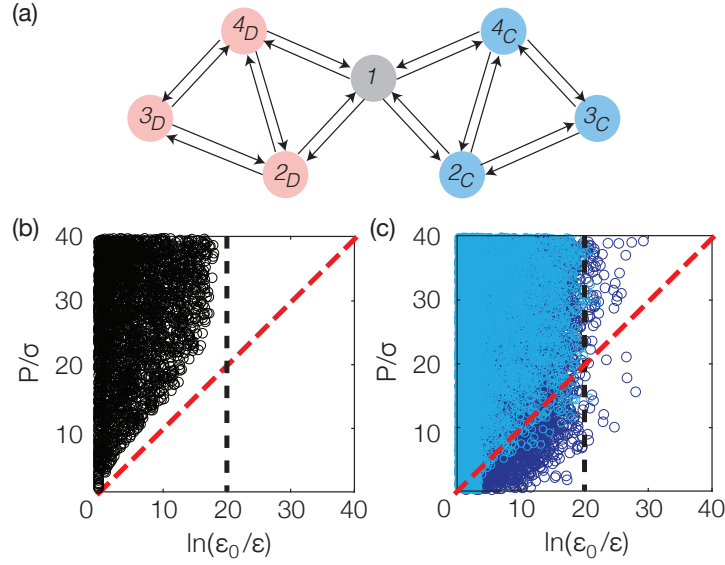


FIG. 2. Numerics for another discrimination mechanism. (a) Graph for a discrimination mechanism that is different from both the Hopfield and McKeithan mechanisms, with label names omitted for clarity. (b) Points plotted as in Supplementary Fig. 1(b). The vertical black dashed line corresponds to the bound  $\epsilon > \epsilon_0^2$  for this mechanism (calculation not shown) that is analogous to equation (5) of the main text for the Hopfield mechanism. The diagonal red dashed line corresponds to equation (23) of the main text, as discussed further there. (c) Similar plot to (b) but with internal discrimination between correct and incorrect substrates, as described in the text, with the light blue points having a lower asymmetry range ( $A = 1$ ) and the dark blue points having a higher range ( $A = 5$ ).

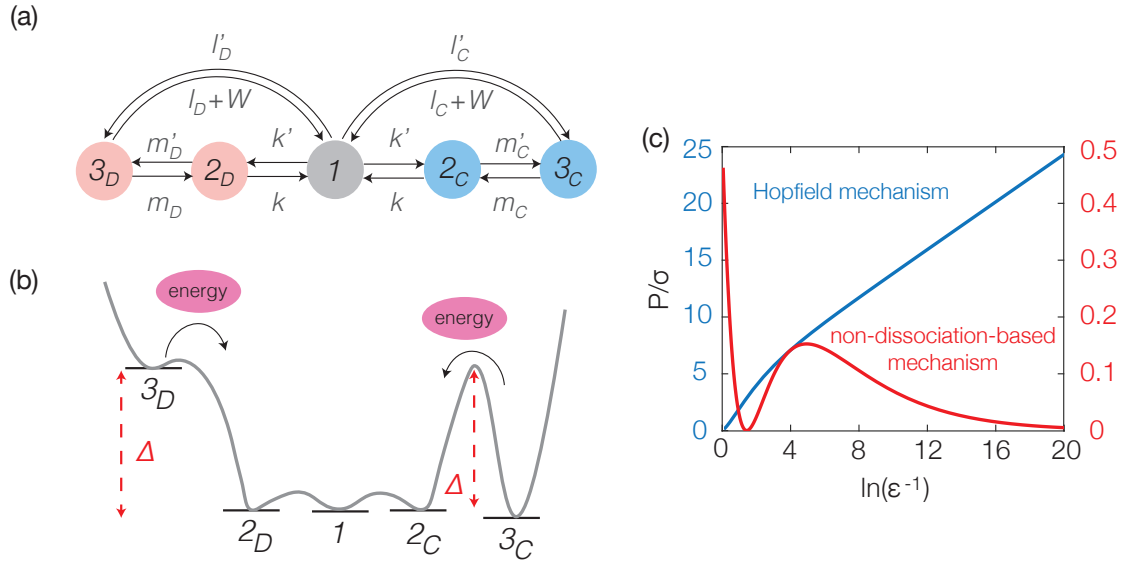


FIG. 3. A non-dissociation-based mechanism. (a) Graph with the same structure as that for the Hopfield mechanism (Fig. 1(a) of the main text) but no discrimination between  $C$  and  $D$  takes place through  $1 \rightleftharpoons 2_X$ , while internal discrimination takes place through  $2_X \rightleftharpoons 3_X$ , as reflected in the label names. (b) Hypothetical energy landscape for the mechanism shown in (a), illustrating where energy may be expended to drive the steps with labels  $m_C$  and  $m_D$ . (c) Plot of  $P/\sigma$  against  $\ln(\epsilon^{-1})$  for a numerical instance of the Hopfield mechanism (Fig. 1(a) of the main text, in blue) and a numerical instance of the non-dissociation-based mechanism in (a) (in red), as  $x$  is varied in the range  $x \in [0, e^{20}]$ . The numerical label values have been determined by taking  $k_D = l_D = x$  and  $l'_C = l'_D = x^{-1}$  for the Hopfield mechanism and  $l'_D = m'_D = m'_C = x^{-1}$ ,  $m_D = x^{1/2}$  and  $m_C = x^{-1/2}$  for the non-dissociation-based mechanism, with all other labels being 1.

label	DNAP	ribosome (wild type)	ribosome (hyperaccurate)	ribosome (error-prone)
$k_C$	900	0.5	0.41	0.43
$k_D$	900	47	46.002	3.999
$k'_C$	0.001	40	27	37
$k'_D$	0.0092	27	25.002	36.001
$m_C$	0.2	0.001	0.001	0.001
$m_D$	2.3	$[4.5 \times 10^{-8}, 21.9]; 10^{-7}$	$[6.0 \times 10^{-8}, 16.6]; 10^{-7}$	$[4.7 \times 10^{-8}, 17.5]; 10^{-7}$
$m'_C$	700	25	14	31
$m'_D$	700	1.2	0.49	3.906
$l_C$	1	0.085	0.048	0.077
$l_D$	$1 \times 10^{-5}$	0.6715	0.4963	0.5891
$l'_C$	250	0.001	0.001	0.001
$l'_D$	0.002	$[1.7 \times 10^{-10}, 0.06]; 0.0272$	$[1.8 \times 10^{-10}, 0.05]; 0.0299$	$[5.8 \times 10^{-9}, 2.1]; 1.0085$
$W_C$	250	8.415	4.752	7.623
$W_D$	0.012	0.0353	0.0035	0.0313

TABLE I. Experimentally measured parameter values, in units of  $s^{-1}$ , for the Hopfield mechanism in Fig. 1(a) of the main text, shown for discrimination during DNA replication by the bacteriophage T7 DNA polymerase (DNAP) and discrimination during mRNA translation by three *E. coli* ribosome variants, as annotated. The values were obtained from Tables S1-S4 of [5]. The labels in the first column correspond to those in Fig. 1(a) of the main text, except that  $m$ ,  $m'$  and  $W$  now have subscripts  $C$  and  $D$ , for the correct and incorrect substrates, respectively, to allow for internal discrimination, as explained in the main text. The values of  $m_D$  and  $l'_D$  were not known for the ribosome variants, so we chose  $m_D$  from  $m_C$  by randomly selecting  $\ln(m_D/m_C)$  from the uniform distribution on  $[-10, 10]$ , which is similar to the asymmetry ranges of the other parameters, and chose  $l'_D$  to satisfy the external chemical potential constraint used by [5], as explained in footnote [45] of the main text. The intervals given for  $m_D$  and  $l'_D$  indicate the range of sampled values. Some samples have  $\epsilon < \epsilon_0$  and these are not shown in Fig. 2(b) of the main text. The values following each interval give the averages of the plotted values in Fig. 2(b) of the main text, as indicated there by asterisks, \*.