Supporting Information: The energy cost and optimal design of networks for biological discrimination

Qiwei Yu,[1,2] Anatoly B. Kolomeisky,[1,3,4,5] and Oleg A. Igoshin[1,3,6,7,*]

[1]*Center for Theoretical Biological Physics, Rice University, Houston, TX 77005*
[2]*Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544*
[3]*Department of Chemistry, Rice University, Houston, TX 77005*
[4]*Department of Chemical and Biomolecular Engineering, Rice University, Houston, TX 77005*
[5]*Department of Physics and Astronomy, Rice University, Houston, TX 77005*
[6]*Department of Bioengineering, Rice University, Houston, TX 77005*
[7]*Department of Biosciences, Rice University, Houston, TX 77005*
(Dated: January 18, 2022)

CONTENTS

* Corresponding author: igoshin@rice.edu

I. FLUX-BASED FORMALISM OF THE ORIGINAL HOPFIELD SCHEME

This section provides the detailed derivation of the results for the original Hopfield scheme (Fig. 1A in main text), including the flux-based formalism (Fig. 2A) and the error-cost bound. We also consider how finite γ introduces a small increase to the minimum energy cost.

A. Deriving the normalized steady-state fluxes for the flux-based formalism

As shown in Fig. 2A in the main text, the normalized fluxes in the correct half of the network are denoted by $j_{\pm 1}$, $j_{\pm 2}$, and $\beta_{\pm 1}$. The normalized fluxes in the incorrect half of the network are denoted by their primed counterparts, which we shall derive in terms of the correct fluxes.

First, the (normalized) fluxes originating from the free enzyme E are given by

$$j_1' = \frac{k_1 P_{\mathrm{E}}}{J_R} = j_1, \quad \beta_{-1}' = \frac{k_{-3} P_{\mathrm{E}}}{J_R} = \beta_{-1}, \tag{S1}$$

which are exactly equal to their counterparts in the correct half of the network. As mentioned in the main text, an additional error rate η_0 is defined as the ratio of the flux from EW to EW* to the flux from ER to ER*:

$$\eta_0 = \frac{k_2 P_{\mathrm{EW}}}{k_2 P_{\mathrm{ER}}} = \frac{k_2 P_{\mathrm{EW}}}{J_R} \frac{J_R}{k_2 P_{\mathrm{ER}}} = \frac{j_2'}{j_2}. \tag{S2}$$

Hence we have $j_2' = j_2 \eta_0$. In addition, j_{-1}' can be related to j_{-1} through j_2':

$$\frac{j_{-1}'}{j_2'} = \frac{f k_{-1}}{k_2} = f \frac{j_{-1}}{j_2} \implies j_{-1}' = f j_{-1} \frac{j_2'}{j_2} = f \eta_0 j_{-1}. \tag{S3}$$

Following the same line of thinking, the (normalized) fluxes originating from the activated state EW* are given by:

$$j_p' = \frac{k_p P_{\mathrm{EW}^*}}{J_R} = \frac{J_W}{J_R} = \eta, \tag{S4}$$

$$j_{-2}' = \frac{k_{-2} P_{\mathrm{EW}^*}}{J_R} = \frac{k_{-2}}{k_p} j_p' = j_p' \frac{k_{-2} P_{\mathrm{ER}^*}}{k_p P_{\mathrm{ER}^*}} = \eta j_{-2}, \tag{S5}$$

$$\beta_1' = \frac{f k_3 P_{\mathrm{EW}^*}}{J_R} = f \frac{k_3}{k_p} y_p' = f y_p' \frac{k_3 P_{\mathrm{ER}^*}}{k_p P_{\mathrm{ER}^*}} = f \eta \beta_1, \tag{S6}$$

where $j_p' = \eta$ was the normalized flux for incorrect formation directly given in Fig. 2A. Thus, all expressions given in the box in Fig. 2A have been derived.

B. Deriving the error-cost bound

The stationary conditions for the fluxes for states ER, EW, ER*, and EW* are:

$$j_1 + j_{-2} = j_{-1} + j_2, \tag{S7}$$

$$j_2 + \beta_{-1} = j_{-2} + \beta_1 + 1, \tag{S8}$$

$$j_1 + \eta j_{-2} = f \eta_0 j_{-1} + \eta_0 j_2, \tag{S9}$$

$$\eta_0 j_2 + \beta_{-1} = \eta j_{-2} + f \eta \beta_1 + \eta. \tag{S10}$$

The stationary condition for E is guaranteed if the stationary conditions for all other states are satisfied. From the first two equations, we eliminate $j_{\pm 2}$ and find $j_1 - j_{-1} = \beta_1 - \beta_{-1} + 1$. From the last two equations, we find $j_1 - f \eta_0 j_{-1} = \eta f \beta_1 - \beta_{-1} + \eta$. Subtracting these two relations yields

$$(f \eta_0 - 1) j_{-1} = (1 - \eta) + (1 - \eta f) \beta_1. \tag{S11}$$

Since the normalized fluxes are positive by definition and the error we consider falls within the range $\eta < f^{-1} < 1$ (error rates larger than f^{-1} can be achieved without any proofreading), the right hand side (RHS) must be positive.

Thus, the left hand side (LHS) is also positive, leading to $\eta_0 > f^{-1}$. Indeed, η_0 only approaches its minimum f^{-1} in the limit $j_{-1} \to +\infty$. j_1 would also diverge to infinity in this limit, which corresponds to the fast equilibrium condition in the $j_{\pm 1}$ step.

Recall the cost (Eq. 9 in main text):

$$C = \frac{(1 + \eta f)\beta_1 - 2\beta_{-1}}{1 + \eta}. \tag{S12}$$

From the second and fourth stationary condition:

$$\beta_1 = j_2 - j_{-2} + \beta_{-1} - 1 = \frac{1}{f\eta}[\eta_0 j_2 - \eta j_{-2} + \beta_{-1} - \eta] \Rightarrow j_2 = \frac{(f-1)\eta(1 + j_{-2}) + (1 - \eta f)\beta_{-1}}{\eta f - \eta_0}. \tag{S13}$$

Thus, β_1 can be eliminated from the expression for the cost:

$$
\begin{aligned}
C &= \frac{(1 + \eta f)\beta_1 - 2\beta_{-1}}{1 + \eta} \\
&= \frac{(1 + \eta_0)j_2 - (1 + \eta)(1 + j_{-2})}{1 + \eta} \\
&= \frac{1 + \eta_0}{1 + \eta} \frac{(f-1)\eta(1 + j_{-2}) + (1 - \eta f)\beta_{-1}}{\eta f - \eta_0} - (1 + j_{-2}) \\
&= \frac{(\eta_0 - \eta)(1 + \eta f)}{(1 + \eta)(\eta f - \eta_0)}(1 + j_{-2}) + \frac{(1 + \eta_0)(1 - \eta f)}{(1 + \eta)(\eta f - \eta_0)}\beta_{-1},
\end{aligned} \tag{S14}
$$

which only depends on η_0, j_{-2}, and β_{-1}. Since $\eta < f^{-1} < \eta_0$, the coefficients $\frac{(\eta_0 - \eta)(1 + \eta f)}{(1+\eta)(\eta f - \eta_0)}$ and $\frac{(1 + \eta_0)(1 - \eta f)}{(1+\eta)(\eta f - \eta_0)}$ are both positive. The cost decreases monotonically with η_0, j_{-2} and β_{-1}. The cost is minimized in the limit:

$$\eta_0 \to f^{-1}, \quad j_{-2} \to 0, \quad \beta_{-1} \to 0. \tag{S15}$$

The minimum cost reads

$$C_{\min} = \frac{(f_{-1} - \eta)(1 + \eta f)}{(1 + \eta)(\eta f - f^{-1})} = \frac{1 - \eta^2 f^2}{(1 + \eta)(\eta f^2 - 1)}, \tag{S16}$$

which gives Eq. 10 in the main text. In the optimal system, other fluxes are given by the stationary condition:

$$j_2 = \frac{\eta f(f - 1)}{\eta f^2 - 1}, \quad \beta_1 = \frac{1 - \eta f}{\eta f^2 - 1}, \quad j_{\pm 1} \to +\infty. \tag{S17}$$

C. Effect of the thermodynamic constraint

Reaching the minimum cost derived above requires vanishing j_{-2} and β_{-1}. Namely, these two reactions need to be irreversible. However, complete irreversibility is impossible due to the thermodynamic constraint:

$$\gamma = e^{\beta \Delta \mu_{\text{futile}}} = \frac{k_1 k_2 k_3}{k_{-1} k_{-2} k_{-3}} = \frac{j_1 j_2 \beta_1}{j_{-1} j_{-2} \beta_{-1}}, \tag{S18}$$

where $\Delta\mu_{\text{futile}}$ is the chemical potential difference for the futile cycle. For finite γ, the fluxes j_{-2} and β_{-1} are positive, which would cause the minimum cost to increase (i.e. introduce a positive correction term). Since the bound becomes exact at infinite γ, we shall calculate the positive correction term to the first order in the large γ limit. This is also motivated by the fact that γ is usually sufficiently large in real biological proofreading networks due to the hydrolysis of energy-rich molecules coupled to the futile cycle.

In the optimal network derived before, the fast equilibrium in the $j_{\pm 1}$ step leads to $j_1/j_{-1} \to 1$. Thus, the thermodynamic constraint reduces to $j_{-2}\beta_{-1} = \gamma^{-1} j_2 \beta_1$. The energy cost is

$$C = \frac{(\eta_0 - \eta)(1 + \eta f)}{(1 + \eta)(\eta f - \eta_0)}\left(1 + j_{-2} + \frac{(1 - \eta f)(1 + \eta_0)}{(1 + \eta f)(\eta_0 - \eta)}\beta_{-1}\right) \geq \frac{(\eta_0 - \eta)(1 + \eta f)}{(1 + \eta)(\eta f - \eta_0)}\left(1 + 2\sqrt{\frac{(1 - \eta f)(1 + \eta_0)}{(1 + \eta f)(\eta_0 - \eta)}}\sqrt{j_{-2}\beta_{-1}}\right) \tag{S19}$$

To obtain the first order correction, we substitute with $\eta_0 = f^{-1}$ and

$$j_{-2}\beta_{-1} = \gamma^{-1}j_2\beta_1 \approx \gamma^{-1} \cdot \frac{\eta f(f-1)}{\eta f^2 - 1} \cdot \frac{1 - \eta f}{\eta f^2 - 1} = \frac{\eta f(f-1)(1-\eta f)}{(\eta f^2 - 1)^2}\gamma^{-1}, \tag{S20}$$

where j_2 and β_1 are evaluated at the infinite γ limit. The cost reads

$$C_{\min} = \frac{1 - \eta^2 f^2}{(1+\eta)(\eta f^2 - 1)}\left(1 + 2\sqrt{\frac{\eta f(1-\eta f)(f^2-1)}{(1+\eta f)(\eta f^2 - 1)^2}}\gamma^{-1/2} + O(\gamma^{-1})\right) \tag{S21}$$

Therefore, the thermodynamic constraint introduces a correction term of order $O(\gamma^{-1/2})$, which is negligible in realistic cases where $\gamma \sim e^{20}$.

II. FLUX-BASED FORMALISM OF THE n-STAGE DISSOCIATION-BASED-DISCRIMINATION SCHEME

In this section, we establish the flux-based formalism and derive the error-cost bound for the n-stage dissociation-based-discrimination (DBD) scheme. The reaction scheme is illustrated in Fig. 1B of the main text with discrimination factors given in Eq. 7 and related text. The notation for the flux-based formalism is given in Fig. 2B of the main text.

A. Deriving the normalized steady-state fluxes for the wrong half of the network

First, we recall the definition of intermediate error rates η_m as the forward flux ratio going from EW_m (ER_m) to EW_{m+1} (ER_{m+1}):

$$\eta_0 = \frac{f_2 k_2 P_{\text{EW}_0}}{k_2 P_{\text{ER}_0}} = \frac{j_2'}{j_2}, \quad \eta_m = \frac{f_{2m+2}k_{2m+2}P_{\text{EW}_m}}{k_{2m+2}P_{\text{ER}_m}} = \frac{\alpha_m'}{\alpha_m} \ (m = 1, 2, \ldots, n). \tag{S22}$$

Since the rate discrimination only appears in dissociation steps, we have $f_2 = f_4 = \cdots = f_{2m+2} = 1$. Following the derivation in the original Hopfield scheme, the (normalized) fluxes in the first two steps (i.e. $\text{E} \leftrightarrow \text{EW}_0 \leftrightarrow \text{EW}_1$) are

$$j_1' = j_1, \quad j_2' = \eta_0 j_2, \tag{S23}$$

$$j_{-1}' = f\eta_0 j_{-1}, \quad j_{-2}' = \eta_1 j_{-2}. \tag{S24}$$

Next, we consider the fluxes associated with the m-th intermediate state ER_m/EW_m, which are given by

$$\alpha_m' = \eta_m \alpha_m, \tag{S25}$$

$$\alpha_{-(m-1)}' = \frac{k_{-2m}P_{\text{EW}_m}}{k_{2m+2}P_{\text{EW}_m}}\alpha_m' = \frac{\alpha_{-(m-1)}}{\alpha_m}\alpha_m' = \eta_m \alpha_{-(m-1)}, \tag{S26}$$

$$\beta_m' = \frac{fk_{2m+1}P_{\text{EW}_m}}{k_{2m+2}P_{\text{EW}_m}}\alpha_m' = f\frac{\beta_m}{\alpha_m}\alpha_m' = f\eta_m\beta_m, \tag{S27}$$

$$\beta_{-m}' = \beta_{-m}. \tag{S28}$$

For the product forming steps, we have $\alpha_n = 1$ and $\alpha_n' = \eta_n = \eta$. Thus, we have derived the (normalized) fluxes presented in Fig. 2B for the n-stage DBD scheme.

B. Deriving the error-cost bound

The cost in the n-stage DBD scheme is given by

$$C_n = \frac{1}{1+\eta}\sum_{m=1}^{n}\left(\beta_m + \beta_m' - \beta_{-m} - \beta_{-m}'\right) = \frac{1}{1+\eta}\sum_{m=1}^{n}[(1+\eta_m f)\beta_m - 2\beta_{-m}], \tag{S29}$$

where $\{\beta\}$ are the (normalized) stationary fluxes. Due to the stationary conditions, summing up the net proofreading fluxes is equivalent to calculating the difference of the total fluxes coming out of the E (free enzyme) state and the total fluxes that lead to products:

$$C_n = \frac{(1 + \eta_0)j_2 - (1 + \eta_1)j_{-2}}{1 + \eta} - 1. \tag{S30}$$

To derive the lower bound of $[(1 + \eta_0)j_2 - (1 + \eta_1)j_{-2}]$, we first prove the following recursive relation:

$$\alpha_m > \frac{\eta_{m+1}(f - 1)}{\eta_{m+1}f - \eta_m}[\pi_{m+2} + \alpha_{-m}], \quad (m = 1, 2, \ldots, n - 1) \tag{S31}$$

where $\pi_m = \prod_{k=m}^{n} \frac{\eta_k(f-1)}{\eta_k f - \eta_{k-1}}$ and $\pi_{n+1} = 1$. The equality condition for Eq. S31 is $\alpha_{-k} = 0$ for $k \geq m + 1$ and $\beta_{-k} = 0$ for $k \geq m$. We will also prove the following relation for the error rates:

$$\eta_{m+1}f > \eta_m, \quad (m = 1, 2, \ldots, n - 1) \tag{S32}$$

The relations Eq. S31 and Eq. S32 are derived inductively via the following steps:

- **Step 1.** For $m = n - 1$, the stationary conditions for states ER_n and EW_n read

$$\alpha_{n-1} - \alpha_{-(n-1)} = 1 + \beta_n - \beta_{-n}, \tag{S33}$$
$$\eta_{n-1}\alpha_{n-1} - \eta_n\alpha_{-(n-1)} = \eta_n + \eta_n f \beta_n - \beta_{-n} \tag{S34}$$

  where $\eta_n = \eta$ is the final error. Elimination of β_n yields

$$(\eta_n f - \eta_{n-1})\alpha_{n-1} = \eta_n(f - 1)(1 + \alpha_{-(n-1)}) + (1 - \eta_n f)\beta_{-n}. \tag{S35}$$

  The coefficient $(1 - \eta_n f)$ is positive since we are considering error $\eta < \eta_{eq} = f^{-1}$. Therefore, RHS is positive. On the other hand, α_{n-1} is positive. For LHS to also be positive, we must have

$$\eta_{n-1} < \eta_n f, \tag{S36}$$

  which recovers Eq. S32 for $m = n - 1$. Since $\beta_{-n} > 0$, we have

$$(\eta_n f - \eta_{n-1})\alpha_{n-1} > \eta_n(f - 1)(1 + \alpha_{-(n-1)}) \Rightarrow \alpha_{n-1} > \frac{\eta_n(f - 1)}{\eta_n f - \eta_{n-1}}(1 + \alpha_{-(n-1)}), \tag{S37}$$

  which recovers Eq. S31 for $m = n - 1$ (since $\pi_{n+1} = 1$). The equality condition is $\beta_{-n} = 0$.

- **Step 2.** For any $m = 1, 2, \ldots, n - 2$, we prove Eq. S31 and Eq. S32 given the condition that they both hold for $l = m + 1$, i.e.

$$\eta_{m+2}f > \eta_{m+1}, \quad \alpha_{m+1} \geq \frac{\eta_{m+2}(f - 1)}{\eta_{m+2}f - \eta_{m+1}}[\pi_{m+3} + \alpha_{-(m+1)}], \quad \text{where } \pi_{m+3} = \prod_{k=m+3}^{n} \frac{\eta_k(f - 1)}{\eta_k f - \eta_{k-1}}. \tag{S38}$$

  Consider the stationary conditions for states ER_{m+1} and EW_{m+1}

$$\alpha_m - \alpha_{-m} = \alpha_{m+1} - \alpha_{-(m+1)} + \beta_{m+1} - \beta_{-(m+1)}, \tag{S39}$$
$$\eta_m\alpha_m - \eta_{m+1}\alpha_{-m} = \eta_{m+1}\alpha_{m+1} - \eta_{m+2}\alpha_{-(m+1)} + \eta_{m+1}f\beta_{m+1} - \beta_{-(m+1)}. \tag{S40}$$

  Eliminating β_{m+1}, we have

$$(\eta_{m+1}f - \eta_m)\alpha_m = \eta_{m+1}(f - 1)(\alpha_{-m} + \alpha_{m+1}) + (\eta_{m+2} - \eta_{m+1}f)\alpha_{-(m+1)} + (1 - \eta_{m+1}f)\beta_{-(m+1)}. \tag{S41}$$

  Since $\beta_{-(m+1)} > 0$ and $1 - \eta_{m+1}f > 0$, we obtain the lower bound:

$$(\eta_{m+1}f - \eta_m)\alpha_m > \eta_{m+1}(f - 1)(\alpha_{-m} + \alpha_{m+1}) + (\eta_{m+2} - \eta_{m+1}f)\alpha_{-(m+1)}. \tag{S42}$$

Plugging in the lower bound for α_{m+1} given in Eq. S38:

$$
\begin{aligned}
(\eta_{m+1}f - \eta_m)\alpha_m &> \eta_{m+1}(f-1)\left(\alpha_{-m} + \frac{\eta_{m+2}(f-1)}{\eta_{m+2}f - \eta_{m+1}}\left[\pi_{m+3} + \alpha_{-(m+1)}\right]\right) + (\eta_{m+2} - \eta_{m+1}f)\alpha_{-(m+1)} \\
&= \eta_{m+1}(f-1)\left(\alpha_{-m} + \pi_{m+2} + \frac{\eta_{m+2}(f-1)}{\eta_{m+2}f - \eta_{m+1}}\alpha_{-(m+1)}\right) + (\eta_{m+2} - \eta_{m+1}f)\alpha_{-(m+1)} \\
&= \eta_{m+1}(f-1)(\alpha_{-m} + \pi_{m+2}) + \frac{(\eta_{m+1} - \eta_{m+2})^2 f}{\eta_{m+2}f - \eta_{m+1}}\alpha_{-(m+1)} \\
&> \eta_{m+1}(f-1)(\alpha_{-m} + \pi_{m+2}).
\end{aligned}
$$

(S43)

Since RHS is positive, LHS must also be positive. Thus we have

$$\eta_{m+1}f > \eta_m. \tag{S44}$$

We can divide both sides by $(\eta_{m+1}f - \eta_m)$ which has been shown to be positive. This leads to

$$\alpha_m > \frac{\eta_{m+1}(f-1)}{\eta_{m+1}f - \eta_m}(\alpha_{-m} + \pi_{m+2}). \tag{S45}$$

As a result of the mathematical induction, Eq. S31 and Eq. S32 holds for $m = 1, 2, \cdots, n-1$. Specifically, the relation for $m = 1$ is

$$\eta_1 < f^{n-1}\eta_n, \quad \alpha_1 > \frac{\eta_2(f-1)}{\eta_2 f - \eta_1}[\pi_3 + \alpha_{-1}]. \tag{S46}$$

We repeat the same derivation for states EW_0 and ER_0. The only difference from repeating the above derivation for $m = 0$ is the notation: $\alpha_{\pm 0}$ is now replaced by $j_{\pm 2}$. This gives us

$$\eta_0 < f\eta_1 < f^n\eta, \quad j_2 > \frac{\eta_1(f-1)}{\eta_1 f - \eta_0}[\pi_2 + j_{-2}]. \tag{S47}$$

The total cost is

$$
\begin{aligned}
C_n &= \frac{(1+\eta_0)j_2 - (1+\eta_1)j_{-2}}{1+\eta} - 1 \\
&> \frac{1+\eta_0}{1+\eta}\pi_1 + \frac{1}{1+\eta}\left(\frac{\eta_1(f-1)(1+\eta_0)}{\eta_1 f - \eta_0} - (1+\eta_1)\right)j_{-2} - 1 \\
&= \frac{1+\eta_0}{1+\eta}\pi_1 + \frac{(\eta_0 - \eta_1)(1+\eta_1 f)}{(1+\eta)(\eta_1 f - \eta_0)}j_{-2} - 1.
\end{aligned}
\tag{S48}
$$

The coefficient $\frac{(\eta_0 - \eta_1)(1+\eta_1 f)}{(1+\eta)(\eta_1 f - \eta_0)}$ is positive since $f^{-1}\eta_0 < \eta_1 < \eta_0$. We also recall that error before proofreading $\eta_0 > \eta_{\mathrm{eq}} = f^{-1}$, with the lower bound reached in the limit of fast equilibrium. Therefore, the minimum cost for the n-stage DBD scheme for given intermediate error rates $\{\eta_m\}$ is

$$C_n > \bar{C}_n = \frac{1+\eta_{\mathrm{eq}}}{1+\eta}\pi_1 - 1 = \frac{(1+f^{-1})(f-1)^n}{1+\eta}\prod_{m=1}^{n}\frac{\eta_m}{\eta_m f - \eta_{m-1}} - 1. \tag{S49}$$

The minimum cost $\bar{C}_n$ is reached in the limit

$$\alpha_{-m} \to 0, \ \beta_{-m} \to 0, (m = 1, 2, \ldots n); \quad j_{-2} \to 0; \quad \eta_0 \to \eta_{\mathrm{eq}} = f^{-1}. \tag{S50}$$

The last condition implies $j_1/j_{-1} \to 1$ and $j_{\pm 1} \to +\infty$. These were results reported in Eq. 12 of the main text.

Next, the minimum cost $\bar{C}_n$ can be further optimized with respective to the intermediate error rates $\{\eta_m\}$. In the main text, a symmetry argument is used to illustrate that $\{\eta_m\}$ must form a geometric series for the dissipation to be optimized. Here, we provide the mathematical proof that it is indeed the unique minimum of the energy cost. From Eq. S49, we define

$$\tilde{C}_n = \ln\left[\frac{1+\eta}{(1+f^{-1})(f-1)^n}(\bar{C}_n + 1)\right] = \ln\left[\prod_{m=1}^{n}\frac{\eta_m}{\eta_m f - \eta_{m-1}}\right] = \sum_{m=1}^{n}(\ln\eta_m - \ln(\eta_m f - \eta_{m-1})). \tag{S51}$$

For any fixed error rate η, $\tilde{C}_n$ is apparently a monotonically increasing function of $\bar{C}_n$. Hence, finding the minimum energy cost is equivalent to minimizing $\tilde{C}_n$ with respect to variables η_m ($m = 1, 2, \ldots, n-1$), which is done by simply taking the derivative:

$$\frac{\partial \tilde{C}_n}{\partial \eta_m} = \frac{1}{\eta_m} - \frac{f}{\eta_m f - \eta_{m-1}} + \frac{1}{\eta_{m+1} f - \eta_m} = \frac{f\left(\eta_m^2 - \eta_{m+1}\eta_{m-1}\right)}{\eta_m(\eta_{m+1} f - \eta_m)(\eta_m f - \eta_{m-1})}. \tag{S52}$$

Setting the first derivative to zero, we get $\eta_m^2 = \eta_{m+1}\eta_{m-1}$, i.e. the intermediate error rates indeed form a geometric series. With the first term $\eta_0 = f^{-1}$ and the last term $\eta_n = \eta$, all the other error rates can be determined as

$$\eta_m = f^{-1}(\eta f)^{m/n}. \tag{S53}$$

It can be verified that the optimal error rates satisfy $\eta_m \in \left(f^{-1}\eta_{m-1}, \eta_{m-1}\right)$, which is consistent with Eq. S32. To verify that this solution indeed correspond to a minimum of the cost, we calculate the second derivative:

$$\begin{aligned}
\left.\frac{\partial^2 \tilde{C}_n}{\partial \eta_m^2}\right|_{\eta_m = f^{-1}(\eta f)^{m/n}} &= \left.\left(-\frac{1}{\eta_m^2} + \frac{f^2}{(\eta_m f - \eta_{m-1})^2} + \frac{1}{(\eta_{m+1} f - \eta_m)^2}\right)\right|_{\eta_m = f^{-1}(\eta f)^{m/n}} \\
&= \left.\eta_m^{-2} \cdot \left(-1 + \frac{1}{\left(1 - \frac{\eta_{m-1}}{\eta_m f}\right)^2} + \frac{1}{\left(\frac{\eta_{m+1} f}{\eta_m} - 1\right)^2}\right)\right|_{\eta_m = f^{-1}(\eta f)^{m/n}} \\
&= \left[f^{-1}(\eta f)^{m/n}\right]^{-2} \cdot \left(-1 + \frac{1}{\left(1 - f^{-1}(\eta f)^{-1/n}\right)^2} + \frac{1}{\left(f(\eta f)^{1/n} - 1\right)^2}\right) \\
&= f^2(\eta f)^{-2m/n} \frac{2f(\eta f)^{1/n}}{\left(f(\eta f)^{1/n} - 1\right)^2} > 0.
\end{aligned} \tag{S54}$$

Thus, the solution found above is a minimum of the energy cost. Moreover, it is a global minimum. The minimum cost is given by

$$\begin{aligned}
C_{n,\min} = \left.\bar{C}_n\right|_{\eta_m = f^{-1}(\eta f)^{m/n}} &= \frac{\left(1 + f^{-1}\right)(f - 1)^n}{1 + \eta} \prod_{m=1}^{n} \frac{f^{-1}(\eta f)^{m/n}}{(\eta f)^{m/n} - f^{-1}(\eta f)^{(m-1)/n}} - 1 \\
&= \frac{(1 + f)(f - 1)^n \eta}{(1 + \eta)\left(f(\eta f)^{1/n} - 1\right)^n} - 1.
\end{aligned} \tag{S55}$$

This is the minimum energy cost reported in Eq. 13 in main text.

C. Analysing the minimum cost

The minimum cost (Eq. S55) vanishes in the limit $\eta \to \eta_{\text{eq}} = f^{-1}$ but diverges in the limit $\eta \to \eta_{\min} = f^{-(n+1)}$. Here, we analyse how the minimum cost depends on the discrimination factor f.

$$\begin{aligned}
C_{n,\min} &= \frac{(f - 1)^n (f + 1)\eta}{\left[f^{1 + \frac{1}{n}}\eta^{\frac{1}{n}} - 1\right]^n (1 + \eta)} - 1 = \frac{\left(1 - f^{-1}\right)^n \left(1 + f^{-1}\right)}{\left[1 - (f^{n+1}\eta)^{-1/n}\right]^n (1 + \eta)} - 1 \\
&= \frac{\left(1 - f^{-1}\right)^n \left(1 + f^{-1}\right) - \left[1 - (f^{n+1}\eta)^{-1/n}\right]^n (1 + \eta)}{\left[1 - (f^{n+1}\eta)^{-1/n}\right]^n (1 + \eta)} \\
&= \frac{\left[1 - (n-1)f^{-1} + O\left(f^{-2}\right)\right] - \left[1 + \eta - n\left(f^{n+1}\eta\right)^{-1/n} + O\left[\left(\eta^{1/n} f^{(n+1)/n}\right)^{-2}\right]\right]}{\left[1 - (f^{n+1}\eta)^{-1/n}\right]^n (1 + \eta)} \\
&= \frac{n\left(f^{n+1}\eta\right)^{-1/n} - (n-1)f^{-1} + \text{h.o.t.}}{\left[1 - (f^{n+1}\eta)^{-1/n}\right]^n (1 + \eta)}.
\end{aligned} \tag{S56}$$

In the intermediate error range $f^{-(n+1)} \ll \eta \ll f^{-1}$, the numerator is dominated by the first term which is proportional to $f^{-(n+1)/n}$, and the denominator is approximately 1. Therefore, the minimum cost decreases with f following a power law:

$$C_{\min} \propto f^{-\frac{n+1}{n}}, \quad \left(\eta^{-1} \ll f \ll \eta^{-1/(n+1)} \right). \tag{S57}$$

The power-law exponent $\frac{n+1}{n}$ is verified in Fig. 3A (main text) for $n = 3$. More importantly, the power law relation between the minimum cost and the discrimination factor indicates that increasing f leads to a non-diminishing benefit in cost reduction (see main text for detailed discussion).

D. Partition between proofreading and catalytic fluxes

The derivation of the minimum energy cost in the above section suggests that in the energetically optimal system, the normalized fluxes in the right half of the network are given by

$$\alpha_m = \pi_{m+1}, \quad \beta_m = \alpha_{m-1} - \alpha_m = \frac{\eta_{m-1} - \eta_m}{\eta_m f - \eta_{m-1}} \pi_{m+1}, \tag{S58}$$

where $\pi_m = \prod_{k=m}^{n} \frac{\eta_k(f-1)}{\eta_k f - \eta_{k-1}}$. On the other hand, these fluxes are related to the steady-state probability $P_{\mathrm{ER_m}}$, reaction rates k_{2m+1}, k_{2m+2}, and the correct product formation flux J_R by

$$\alpha_m = \frac{k_{2m+2} P_{\mathrm{ER_m}}}{J_R}, \quad \beta_m = \frac{k_{2m+1} P_{\mathrm{ER_m}}}{J_R}. \tag{S59}$$

The ratio of these two fluxes is

$$\frac{\beta_m}{\alpha_m} = \frac{k_{2m+1}}{k_{2m+2}} = \frac{\eta_{m-1} - \eta_m}{\eta_m f - \eta_{m-1}} = \frac{1 - (\eta f)^{1/n}}{f(\eta f)^{1/n} - 1}, \tag{S60}$$

which has taken into account the optimal error rates $\eta_m = f^{-1}(\eta f)^{m/n}$. This is the partition ratio given in Eq. 14 in the main text. The reaction rates k_{2m+1}, k_{2m+2} can be expressed in terms of the energy levels of the discrete states and the energy barriers:

$$k_{2m+1} = k_{2m+1}^0 \exp\left(\epsilon_m - \epsilon_{m,p}^\dagger \right), \quad k_{2m+2} = k_{2m+2}^0 \exp\left(\epsilon_m - \epsilon_{m,m+1}^\dagger \right). \tag{S61}$$

k_{2m+1}^0 and k_{2m+2}^0 are prefactors independent of the energy levels. ϵ_m is the energy level of $\mathrm{ER_m}$. $\epsilon_{m,p}^\dagger$ and $\epsilon_{m,m+1}^\dagger$ is the energy level of the transition state (energy barrier) between $\mathrm{ER_m}$ and $\mathrm{ER_{m+1}}$. Therefore, the ratio β_m/α_m is actually only related to the difference between the energy level of the two transition states:

$$\frac{\beta_m}{\alpha_m} = \frac{k_{2m+1}}{k_{2m+2}} = \frac{1 - (\eta f)^{1/n}}{f(\eta f)^{1/n} - 1} \propto \exp\left(\epsilon_{m,m+1}^\dagger - \epsilon_{m,p}^\dagger \right). \tag{S62}$$

As discussed in the main text, this is a manifestation of how the error-cost relation is kinetically controlled.

E. Effect of the thermodynamic constraints

Similar to the case of the original Hopfield scheme, the thermodynamic constraints prevent any reaction to be completely irreversible and introduces a correction term to the minimum energy cost (Eq. S55) in the n-stage DBD scheme. Here we calculate the leading order contribution of this correction term.

In the derivation of the error-cost bound, many of the fluxes were set to zero since they only increase the overall cost. These terms must be recovered as we study the effect of the thermodynamic constraints. Fortunately, due to the linearity of the stationary conditions, they contribute to the cost through a linear relation:

$$C = C(j_{-2}, \{\alpha\}, \{\beta\}) = C_0 + a_0 j_{-2} + \sum_{m=1}^{n-1} a_m \alpha_{-m} + \sum_{m=1}^{n} b_m \beta_{-m}, \tag{S63}$$

where C_0 is the minimum cost in Eq. S55. The coefficients a_i $(i = 0, 1, 2, \ldots, n-1)$ and b_i $(i = 1, 2, 3, \ldots, n)$ are positive functions of f and η_m $(m = 1, 2, \ldots, n)$. Following the inductive method used to derive the bound, we find the following coefficients:

$$a_0 = \frac{(\eta_0 - \eta_1)(1 + \eta_1 f)}{(1 + \eta)(\eta_1 f - \eta_0)} \tag{S64}$$

$$a_i = \frac{1 + \eta_0}{1 + \eta} \frac{\pi_1}{\pi_{i+2}} \frac{f(\eta_i - \eta_{i+1})^2}{\eta_i \eta_{i+1}(f-1)^2}, \quad i = 1, 2, \ldots, n-1 \tag{S65}$$

$$b_i = \frac{1 + \eta_0}{1 + \eta} \frac{\pi_1}{\pi_{i+1}} \frac{1 - f\eta_i}{\eta_i(f-1)}, \quad i = 1, 2, \ldots, n, \tag{S66}$$

where $\eta_0 = f^{-1}$, $\pi_m = \prod_{k=m}^{n} \frac{\eta_k(f-1)}{\eta_k f - \eta_{k-1}}$, and $\pi_{n+1} = 1$.

The thermodynamic constraints are:

$$\gamma = \frac{j_2}{j_{-2}} \cdot \frac{\beta_1}{\beta_{-1}} = \frac{j_2}{j_{-2}} \cdot \frac{\alpha_1 \beta_2}{\alpha_{-1}\beta_{-2}} = \cdots = \frac{j_2}{j_{-2}} \cdot \prod_{k=1}^{m-1} \frac{\alpha_k}{\alpha_{-k}} \frac{\beta_m}{\beta_{-m}}, \quad (m = 1, 2, \ldots n) \tag{S67}$$

For any futile cycle, the thermodynamic correction to the energy cost is of the order $\gamma^{-1/L}$, where L is the number of reactions needed to be driven strongly forward in this cycle. This is because the cost always depends on the reverse reaction fluxes, which should vanish without the thermodynamic constraint, in a linear fashion. Thus, in the presence of the thermodynamic constraint, the cost is minimized when those reverse fluxes are of the same order of magnitude, i.e. of order $\gamma^{-1/L}$. Hence, the first order contribution in γ comes from the largest futile cycle, which has $(n+1)$ reaction steps that need to be driven forward. The thermodynamic constraint for this cycle can be reorganized to:

$$j_{-2}\beta_{-n} \prod_{m=1}^{n-1} \alpha_{-m} = \frac{j_2 \beta_n}{\gamma} \prod_{m=1}^{n-1} \alpha_m. \tag{S68}$$

Therefore, the first correction to the cost is calculated as follows:

$$\begin{aligned}
C &= C_0 + \left(a_0 j_{-2} + \sum_{m=1}^{n-1} a_m \alpha_{-m} + b_n \beta_{-n} \right) + \sum_{m=1}^{n-1} b_m \beta_{-m} \\
&\geq C_0 + (n+1) \left(a_0 j_{-2} \cdot \prod_{m=1}^{n-1} a_m \alpha_{-m} \cdot b_n \beta_{-n} \right)^{1/(n+1)} + \sum_{m=1}^{n-1} b_m \beta_{-m} \\
&= C_0 + (n+1) \left(a_0 j_2 \cdot b_n \beta_n \prod_{m=1}^{n-1} a_m \alpha_m \right)^{1/(n+1)} \gamma^{-1/(n+1)} + O(\gamma^{-2/(n+1)}) \\
&= C_0 + C_1 \gamma^{-1/(n+1)} + O(\gamma^{-1/n}).
\end{aligned} \tag{S69}$$

The $O(\gamma^{-1/n})$ term is due to the second largest futile cycle which has length n. The coefficient C_1 is given by

$$C_1 = (n+1) \left(a_0 j_2 \cdot \prod_{m=1}^{n-1} a_m \alpha_m \cdot b_n \beta_n \right)^{1/(n+1)}, \tag{S70}$$

where the coefficients (a_0, a_m, b_n) and fluxes (j_2, α_m, β_n) are evaluated in the optimal scheme, i.e. as if the thermodynamic constraints are not present. Thus, the correction is of the order $\gamma^{-1/(n+1)}$. Although the correction term becomes increasingly significant as n is increased, the number of proofreading pathways in real biological systems is usually limited, so the correction term remains small. Moreover, note that the correction term due to thermodynamic constraints is always positive, so the original error-cost bound could never be violated.

III. A SIMPLE KINETIC MODEL FOR n-STAGE PROOFREADING

In this section, we study the n-stage proofreading scheme shown in Fig. 4A by directly solving the Chemical Master Equation (CME). We introduce $P_m(t)$ to denote the probability for state ER_m at time t and $P_{-m}(t)$ to denote the

probability for state EW_m. The probability for the free enzyme state E is denoted by $P_0(t)$. The probabilities are normalized by the condition

$$\sum_{m=-n}^{n} P_m(t) = 1, \quad \forall t \in (-\infty, +\infty). \tag{S71}$$

The CME reads

$$\frac{dP_0(t)}{dt} = (1+a)\kappa_n P_n(t) + (1+fa)\kappa_n P_{-n}(t) + fa\sum_{m=1}^{n-1}\kappa_m P_{-m}(t) + a\sum_{m=1}^{n-1}\kappa_m P_m(t) - \left(1+f^{-1}\right)\kappa_0 P_0(t), \tag{S72}$$

$$\frac{dP_m(t)}{dt} = \kappa_{m-1}P_{m-1}(t) - (1+a)\kappa_m P_m(t), \quad m = 1, 2, \ldots n \tag{S73}$$

$$\frac{dP_{-1}(t)}{dt} = f^{-1}\kappa_0 P_0(t) - (1+fa)\kappa_1 P_{-1}(t), \tag{S74}$$

$$\frac{dP_{-m}(t)}{dt} = \kappa_{m-1}P_{-(m-1)}(t) - (1+fa)\kappa_m P_{-m}(t), \quad m = 2, 3, \ldots n. \tag{S75}$$

We are interested in the steady-state solution, which satisfies $\frac{dP_m}{dt} = 0$ ($m = -n, -(n-1), \ldots, n-1, n$). The stationary condition for state m ($m = 1, 2, 3, \ldots, n$) leads to:

$$\frac{dP_m(t)}{dt} = \kappa_{m-1}P_{m-1}(t) - (1+a)\kappa_m P_m(t) = 0 \Rightarrow P_m = \frac{1}{(1+a)^m}\frac{\kappa_0}{\kappa_m}P_0. \tag{S76}$$

Similarly, the stationary condition for state $(-m)$ leads to

$$P_{-m} = \frac{1}{f(1+fa)^m}\frac{\kappa_0}{\kappa_m}P_0. \tag{S77}$$

The error rate η is given by

$$\eta = \frac{J_W}{J_R} = \frac{P_{-n}}{P_n} = f^{-1}\left(\frac{1+a}{1+fa}\right)^n. \tag{S78}$$

Note that the error η is always bound between $\eta_{\min} = f^{-n-1}$ (in the limit $a \to \infty$) and $\eta_{eq} = f^{-1}$ (in the limit $a \to 0$). From this relation, we can solve for a as a function of η:

$$a = \frac{1 - (\eta f)^{1/n}}{f(\eta f)^{1/n} - 1}. \tag{S79}$$

On the other hand, the energy cost C is given by

$$\begin{aligned}
C &= \frac{1}{J_R + J_W}\left(a\sum_{m=1}^{n}\kappa_m P_m + fa\sum_{m=1}^{n}\kappa_m P_{-m}\right) \\
&= \frac{a}{(1+\eta)\kappa_n P_n}\sum_{m=1}^{n}\kappa_m(P_m + fP_{-m}) \\
&= \frac{a(1+a)^n}{(1+\eta)\kappa_0 P_0}\sum_{m=1}^{n}\kappa_0 P_0\left(\frac{1}{(1+a)^m} + \frac{1}{(1+fa)^m}\right) \\
&= \frac{a(1+a)^n}{1+\eta}\left[\frac{1-(1+a)^{-n}}{a} + \frac{1-(1+fa)^{-n}}{fa}\right] \\
&= \frac{(1+a)^n\left(1+f^{-1}\right)}{1+\eta} - 1.
\end{aligned} \tag{S80}$$

Substituting a with $a = \frac{1-(\eta f)^{1/n}}{f(\eta f)^{1/n}-1}$, we obtain the full expression for the minimum cost

$$C = (1+f)\left(\frac{f-1}{f(\eta f)^{1/n}-1}\right)^n\frac{\eta}{1+\eta} - 1 = \frac{(1+f)(f-1)^n}{\left(f(\eta f)^{1/n}-1\right)^n}\frac{\eta}{1+\eta} - 1 \tag{S81}$$

which is exactly the dissipation bound for n-stage DBD scheme reported in the main text (Eq. 13).

In this simplified model, both error and energy dissipation are modulated by the partition ratio a, which is equivalent to the flux-splitting ratio β_m/α_m in the flux-based formalism calculated above. When $a \to 0$, the system approaches the non-dissipative, equilibrium discrimination regime with $\eta \to \eta_{\mathrm{eq}} = f^{-1}$ and $C \to 0$. When $a \to \infty$, the system approaches the limit to which error can be reduced by dissipative proofreading, namely $\eta \to \eta_{\min} = f^{-n-1}$ and $C \to \infty$. Moreover, the system is optimized as long as the partition ratio a is uniform for all proofreading pathways. The continuous tuning of $a \in (0, \infty)$ therefore represents a trade-off between error and dissipation, where error can be reduced by increasing a at the cost of more dissipation.

IV. MICHAELIS-MENTEN SCHEME WITH DISSIPATIVE RESETTING

This section provides detailed derivation of the error-cost relation in the MM-with-proofreading scheme reported in Fig. 5 in the main text. More complex reaction networks can be considered as combination or generalization of this type of reaction network.

The reaction scheme is presented in Fig. S1A with notations introduced in the main text. Due to the kinetic control of both error and energy cost, we introduce a set of variables ξ_i to quantify the difference between energy barriers:

$$\xi_1 = f_1, \quad \xi_2 = \frac{f_1 f_2}{f_{-1}} = f_{-2}, \quad \xi_p = \frac{f_1 f_p}{f_{-1}} = \frac{f_{-2} f_p}{f_2}. \tag{S82}$$

Their relation with energy barrier differences are reported in Eq. 17 in the main text.

In the absence of proofreading ($k_{\pm 2} = 0$), the minimum error is determined by the maximum difference in energy barriers

$$\eta_{\mathrm{eq}} = \min(\xi_1, \xi_p) = f_1 \min\left(1, \frac{f_p}{f_{-1}}\right). \tag{S83}$$

The minimum error is achieved by making the step with the largest barrier difference rate-limiting. Namely, k_p is rate-limiting if $f_p < f_{-1}$, and k_1 is rate-limiting if $f_p > f_{-1}$.

We study the relation between error and energy cost in the parameter regime where the dissipative proofreading mechanism is relevant, i.e. it could achieve some error rate $\eta < \eta_{\mathrm{eq}}$ which is otherwise inaccessible. The condition for the proofreading mechanism to reduce error below η_{eq} is

$$f_2 > \max(f_{-1}, f_p). \tag{S84}$$

In the main text, this condition is justified with the heuristic argument that proofreading only improves the accuracy if it creates more bias in the dissociation of incorrect complexes compared to the bias in non-dissipative dissociation (unbinding) or product formation. In the following, this condition is justified *a posteriori* after the minimum error is derived.

In Fig. S1B, we present the flux-based formalism for the MM-with-proofreading scheme, where the noncognate fluxes (i.e. $j'_{\pm 1, \pm 2}$) has already been derived and labeled on the reactions. η stands for the error. The fluxes are constrained by the stationary conditions for ER and EW:

$$j_1 - j_{-1} = 1 + j_2 - j_{-2}, \quad f_1 j_1 - \eta \frac{f_{-1}}{f_p} j_{-1} = \eta + \eta \frac{f_2}{f_p} j_2 - f_{-2} j_{-2}. \tag{S85}$$

The energy cost C is given by

$$C = \frac{1}{1+\eta}\left(j_2 - j_{-2} + \eta \frac{f_2}{f_p} j_2 - f_{-2} j_{-2}\right) = \frac{1}{1+\eta}\left[(1 + f_1) j_1 - \left(1 + \eta \frac{f_{-1}}{f_p}\right) j_{-1}\right] - 1. \tag{S86}$$

Eliminating j_2 from the stationary conditions yields:

$$\left(\eta \frac{f_2}{f_p} - f_1\right) j_1 = \eta \frac{f_2 - f_{-1}}{f_p} j_{-1} + \eta \frac{f_2 - f_p}{f_p} + f_2\left(\frac{f_1}{f_{-1}} - \frac{1}{f_p}\eta\right) j_{-2}, \tag{S87}$$

where f_{-2} has been substituted by $\frac{f_1 f_2}{f_{-1}}$ due to thermodynamic constraints. We note that the right hand side is positive due to conditions $f_2 > \max(f_{-1}, f_p)$ and $\eta < \eta_{\mathrm{eq}} \le \frac{f_1 f_p}{f_{-1}}$. Thus, the left hand side must also be positive, leading to the minimum error

$$\eta > \eta_{\min} = \frac{f_1 f_p}{f_2}. \tag{S88}$$

The condition for the minimum error in the presence of proofreading to be smaller than the minimum error without proofreading is

$$\eta_{\min} < \eta_{\text{eq}} \Leftrightarrow \frac{f_1 f_p}{f_2} < f_1 \min\left(1, \frac{f_p}{f_{-1}}\right) \Leftrightarrow f_2 > \frac{f_p}{\min\left(1, \frac{f_p}{f_{-1}}\right)} = \max(f_p, f_{-1}), \tag{S89}$$

which recovers the condition Eq. S84. This is the condition for the nonequilibrium proofreading mechanism to be relevant. It can be verified that if f_2 is smaller than either f_{-1} or f_p, the minimum error can always be achieved without proofreading.

Finally, we consider the energy cost for $\eta \in (\eta_{\min}, \eta_{\text{eq}})$:

$$\begin{aligned}
C &= \frac{1}{1+\eta}\left[(1+f_1)j_1 - \left(1+\eta\frac{f_{-1}}{f_p}\right)j_{-1}\right] - 1 \\
&= \frac{1}{1+\eta}\left[(1+f_1)\frac{\eta\frac{f_2-f_{-1}}{f_p}j_{-1} + \eta\frac{f_2-f_p}{f_p} + f_2\left(\frac{f_1}{f_{-1}} - \frac{1}{f_p}\eta\right)j_{-2}}{\eta\frac{f_2}{f_p} - f_1} - \left(1+\eta\frac{f_{-1}}{f_p}\right)j_{-1}\right] - 1 \\
&= C_0 + a_1 j_{-1} + a_2 j_{-2}.
\end{aligned} \tag{S90}$$

The coefficients are given by

$$C_0 = \frac{(f_1 - \eta)\left(1 + \eta\frac{f_2}{f_p}\right)}{(1+\eta)(\eta\frac{f_2}{f_p} - f_1)}, \tag{S91}$$

$$a_1 = \frac{\left(f_1 - \eta\frac{f_{-1}}{f_p}\right)\left(1 + \eta\frac{f_2}{f_p}\right)}{(1+\eta)\left(\eta\frac{f_2}{f_p} - f_1\right)}, \tag{S92}$$

$$a_2 = \frac{(1+f_1)f_2\left(\frac{f_1}{f_{-1}} - \frac{\eta}{f_p}\right)}{(1+\eta)\left(\eta\frac{f_2}{f_p} - f_1\right)}. \tag{S93}$$

$$\tag{S94}$$

Since $a_{1,2} > 0$, the energy cost is minimized when the reverse fluxes $j_{-1,-2} \to 0$. The minimum energy cost is given by

$$C_{\min} = C_0 = \frac{(f_1 - \eta)\left(1 + \eta\frac{f_2}{f_p}\right)}{(1+\eta)(\eta\frac{f_2}{f_p} - f_1)}, \tag{S95}$$

which recovered Eq. 16 of the main text.

The effect of the thermodynamic constraint $\gamma = \frac{k_1 k_2}{k_{-1} k_{-2}}$ can be analysed following the method used in the original Hopfield scheme (section I C). The correction is of the order $\gamma^{-1/2}$ due to having two reactions driven irreversibly forward in the futile cycle.

V. PARAMETERS AND ADDITIONAL SIMULATION RESULTS FOR THE REAL BIOLOGICAL SYSTEMS

In this section, we provide additional details for the three biological examples analysed in the main text (Fig. 6).

A. T7 DNA polymerase

The reaction network and parameters for the DNA replication network are obtained from previous works [1, 2]. For reference purposes, the reaction network has been reproduced in Fig. S1C.
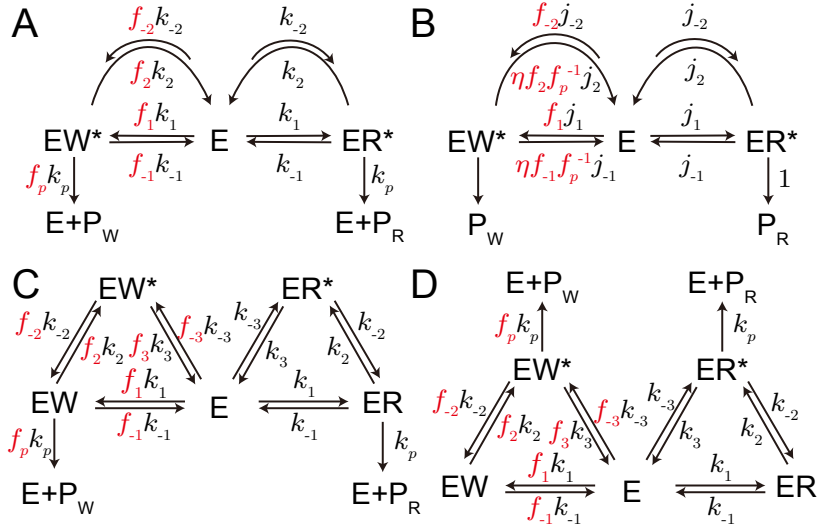
FIG. S1. Reaction schemes used in the main text and SI. (A) Michaelis-Menten scheme with dissipative resetting (reproduced from Fig. 5A in main text for comparison with the flux-based formalism). (B) Flux-based formalism for the MM-with-proofreading scheme. (C) Reaction scheme for T7 DNA polymerase, reproduced from ref. [1, 2]. (D) Reaction scheme for *E. coli* ribosome, reproduced from ref. [1, 2].

    a. Relation between various error rates. We first verify that the native system operates in the regime where dissipative proofreading is necessary. The error of the native system is $\eta_{\text{wt}} = 7.39 \times 10^{-8}$; the minimum error for discrimination without proofreading is $\eta_{\text{eq}} = \min(\xi_1, \xi_p) = \xi_1 = 8.00 \times 10^{-6}$; the minimum error for the first step is $\eta_0 = \xi_1 = f_1 = 8.00 \times 10^{-6}$; the overall minimum error is $\eta_{\min} = \frac{f_p}{f_2} f_1 = 3.34 \times 10^{-11}$. Therefore, the relation between these error rates is

$$\eta_0 = \eta_{\text{eq}} > \eta_{\text{wt}} > \eta_{\min}. \tag{S96}$$

The native system η_{wt} falls within the non-equilibrium discrimination regime.

    b. Optimal and native proofreading systems The only difference between the DNA replication network and the MM-with-proofreading scheme is the addition of intermediate states EW^* and ER^*. The additional states will not change the error-cost bound since proofreading reaction is driven irreversibly forward in the optimal scheme, as indicated by the derivation in the last section. Hence, the error-cost bound is the same as that derived in the MM-with-proofreading scheme:

$$C_{\min} = \frac{(f_1 - \eta)\left(1 + \eta \frac{f_2}{f_p}\right)}{(1 + \eta)(\eta \frac{f_2}{f_p} - f_1)}, \tag{S97}$$

which is the red line in Fig. 6A of the main text. This bound indeed encapsulates all the systems sampled.

    At the native error rate, the optimal partition ratio is given by:

$$a_{\text{optimal}} = (j_2)_{\text{optimal}} = \frac{f_1 - \eta_{\text{wt}}}{\eta_{\text{wt}} \frac{f_2}{f_p} - f_1} = 4.5 \times 10^{-4}. \tag{S98}$$

The native partition ratio is

$$a_{\text{wt}} = \frac{k_2}{k_p} = 8.0 \times 10^{-4}. \tag{S99}$$

B. *E. coli* ribosome

    The reaction network and parameters for the protein replication network are also obtained from previous works [1, 2]. For reference purposes, the reaction network has been reproduced in Fig. S1D.
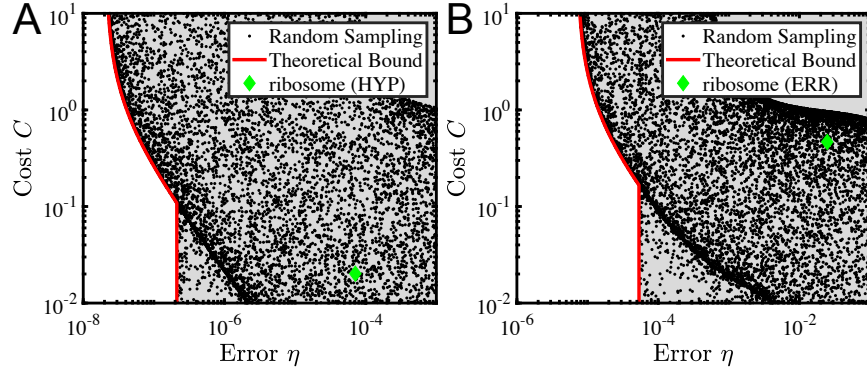
FIG. S2. The error-dissipation relations in two mutants of the *E. coli* ribosome. Left: mutant *rpsL141*, which is hyperaccurate (HYP). Right: mutant *rpsD12*, which is more error-prone than WT (ERR).

    a.   *Relation between various error rates.*   The error of the native system is $\eta_{\rm wt} = 8.65 \times 10^{-4}$; the minimum error for discrimination without proofreading is $\eta_{\rm eq} = \min(\xi_1, \xi_2, \xi_p) = \xi_p = 1.45 \times 10^{-6}$; the minimum error for the first two steps is $\eta_0 = \min(\xi_1, \xi_2) = \xi_2 = 3.45 \times 10^{-4}$; the overall minimum error is $\eta_{\min} = \frac{f_p}{f_3} \eta_0 = 1.83 \times 10^{-7}$. Therefore, the relation between these error rates is

$$\eta_{\rm wt} > \eta_0 > \eta_{\rm eq} > \eta_{\min}. \tag{S100}$$

The native error rate $\eta_{\rm wt}$ falls within the equilibrium discrimination regime, which can in principle be achieved without the proofreading step. As discussed in the main text, achieving $\eta_{\rm eq}$ requires the product formation step k_p to be much smaller than the preceding reactions $k_{\pm 1, \pm 2}$, which could not be realized due to speed requirements. Similarly, achieving η_0 without proofreading requires GTP hydrolysis (k_2) to be rate-limiting, which is also prevented by speed requirements.

    b.   *Energy-cost bound in the translation network.*   The network has only one proofreading pathway, and the error-cost bound takes the same form as the bound in the MM-with-proofreading scheme (Eq. S95) with f_1 replaced by η_0 (the minimum error in the first two steps) and f_2 replaced by f_3 (the discrimination factor for the proofreading step). Therefore, the error-cost bound in the ribosome network is

$$C_{\min} = \frac{(\eta_0 - \eta)\left(1 + \eta \frac{f_3}{f_p}\right)}{(1 + \eta)\left(\eta \frac{f_3}{f_p} - \eta_0\right)} = \frac{\left(\frac{f_1 f_2}{f_{-1}} - \eta\right)\left(1 + \eta \frac{f_3}{f_p}\right)}{(1 + \eta)\left(\eta \frac{f_3}{f_p} - \frac{f_1 f_2}{f_{-1}}\right)}. \tag{S101}$$

This bound correspond to the red line in Fig. 6B in main text.

    c.   *Results in mutants.*   The simulation results for the ERR (error-prone) and HYP (hyperaccurate) mutants are qualitatively similar to the results in WT. The parameters for these two mutants are obtained from ref. [1]. The numeric results are presented in Fig. S2.

C.   *E. coli* isoleucyl-tRNA synthetase (IleRS)

The reaction network and parameters for the IleRS network are obtained from ref. [3]. The reaction network is presented in Fig. S3. The error and cost of the IleRS network are bounded by the following piecewise function:

$$C_{\min} = \begin{cases} \frac{1 + \eta_0}{1 + \eta} \frac{(b_1 - 1)(b_2 - 1)(b_3 - 1)\eta}{\left((b_1 b_2 b_3 \eta)^{1/3} - \eta_0^{1/3}\right)^3} - 1 & \frac{\eta_0}{b_1 b_2 b_3} < \eta < \frac{\eta_0 b_1^2}{b_2 b_3} \\[2mm] \frac{1 + \eta_0}{1 + \eta} \frac{(b_2 - 1)(b_3 - 1)\eta}{\left((b_2 b_3 \eta)^{1/2} - \eta_0^{1/2}\right)^2} - 1 & \frac{\eta_0 b_1^2}{b_2 b_3} < \eta < \frac{\eta_0 b_2}{b_3} \\[2mm] \frac{1 + \eta_0}{1 + \eta} \frac{(b_3 - 1)\eta}{b_3 \eta - \eta_0} - 1 & \frac{\eta_0 b_2}{b_3} < \eta < \eta_0 \end{cases} \tag{S102}$$

where $b_1 = \frac{f_{h1}}{f_3}$, $b_2 = \frac{f_{h2}}{f_4}$, and $b_3 = \frac{f_{h3}}{f_p}$. $\eta_0 = \min\left(f_+, \frac{f_+}{f_-} f_a\right)$ is the minimum error of the equilibrium discrimination by the first two steps (binding and activation). Eq. S102 corresponds to the red line in Fig. 6C in the main text. The three error intervals correspond to the three phases of proofreading in Fig. 6D in the main text. In this section, we provide detailed derivation for the error-cost bound and the optimal partition ratios.
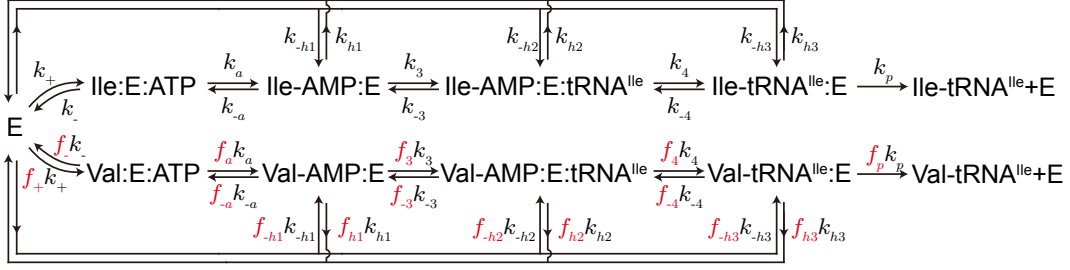
FIG. S3. Reaction scheme for *E. coli* tRNA$^{\text{Ile}}$ aminoacylation, reproduced from ref. [3].

a. Derivation of the error-cost bound. Here, we show that due to the different discrimination factors in the three proofreading pathways, the optimal partition ratios are no longer uniform, and there will be three proofreading regimes due to the sequential "shutdown" of proofreading pathways.

First, we investigate the case where all three proofreading pathways are utilized, which can be considered as three MM-with-proofreading schemes applied in tandem. The minimum cost in the MM-with-proofreading scheme (Eq. S95) gives the ratio of the total input flux (product formation plus proofreading) to the output (product-forming) flux:

$$\frac{J_{\text{in}}}{J_{\text{out}}} = \frac{J_{\text{hydrolysis}}}{J_{\text{product}}} = 1 + C = 1 + \frac{(f_1 - \eta)\left(1 + \eta\frac{f_2}{f_p}\right)}{(1+\eta)\left(\eta\frac{f_2}{f_p} - f_1\right)} = 1 + \frac{(\eta_0 - \eta)(1 + \eta b)}{(1+\eta)(\eta b - \eta_0)} = \frac{\eta(\eta_0 + 1)(b - 1)}{(1+\eta)(\eta b - \eta_0)}, \tag{S103}$$

where $b = \frac{f_2}{f_p}$ is the discrimination of the partition ratio, and $\eta_0 = f_1$ can be considered as the error of the last proofreading stage. The optimal partition ratio corresponding to this minimum cost is

$$a = j_2 = \frac{\eta_0 - \eta}{\eta b - \eta_0}. \tag{S104}$$

The minimum proofreading cost can thus be calculated by taking the product of the ratios $J_{\text{in}}/J_{\text{out}}$ in all three proofreading pathways, assuming optimal partition ratios. We denote the error at the three proofreading stages as $\eta_{1,2,3}$, respectively. $\eta_3 = \eta$ is the final error rate. $\eta_0 = \min\left(f_+, \frac{f_+}{f_-}f_a\right)$ is the minimum error before proofreading. The minimum cost is therefore

$$C = \frac{J_{\text{in}}}{J_{\text{out}}} - 1 \tag{S105}$$

$$= \frac{\eta_1(\eta_0 + 1)(b_1 - 1)}{(1+\eta_1)(\eta_1 b_1 - \eta_0)} \frac{\eta_2(\eta_1 + 1)(b_2 - 1)}{(1+\eta_2)(\eta_2 b_2 - \eta_1)} \frac{\eta_3(\eta_2 + 1)(b_3 - 1)}{(1+\eta_3)(\eta_3 b_3 - \eta_2)} - 1 \tag{S106}$$

$$= \frac{\eta_1\eta_2\eta_3(1 + \eta_0)(b_1 - 1)(b_2 - 1)(b_3 - 1)}{(1+\eta_3)(\eta_1 b_1 - \eta_0)(\eta_2 b_2 - \eta_1)(\eta_3 b_3 - \eta_2)} - 1 \tag{S107}$$

$$= \frac{1 + \eta_0}{1 + \eta} \frac{\left(1 - b_1^{-1}\right)\left(1 - b_2^{-1}\right)\left(1 - b_3^{-1}\right)}{\left(1 - \frac{\eta_0}{\eta_1 b_1}\right)\left(1 - \frac{\eta_1}{\eta_2 b_2}\right)\left(1 - \frac{\eta_2}{\eta_3 b_3}\right)} - 1 \tag{S108}$$

The denominator can be maximized with Jensen's inequality. Since $f(x) = \ln(1 - e^x)$ $(x \in (0,1))$ is a concave function $(f''(x) < 0)$, we have

$$f\left(\ln\frac{\eta_0}{\eta_1 b_1}\right) + f\left(\ln\frac{\eta_1}{\eta_2 b_2}\right) + f\left(\ln\frac{\eta_2}{\eta_3 b_3}\right) \leq 3f\left(\frac{1}{3}\ln\frac{\eta_0}{b_1 b_2 b_3 \eta_3}\right) \tag{S109}$$

$$\Rightarrow \left(1 - \frac{\eta_0}{\eta_1 b_1}\right)\left(1 - \frac{\eta_1}{\eta_2 b_2}\right)\left(1 - \frac{\eta_2}{\eta_3 b_3}\right) \leq \left[1 - \left(\frac{\eta_0}{b_1 b_2 b_3 \eta_3}\right)^{1/3}\right]^3. \tag{S110}$$

Hence, we obtain the minimum cost in this regime:

$$C_{\min} = \frac{1 + \eta_0}{1 + \eta} \frac{\left(1 - b_1^{-1}\right)\left(1 - b_2^{-1}\right)\left(1 - b_3^{-1}\right)}{\left[1 - \left(\frac{\eta_0}{b_1 b_2 b_3 \eta_3}\right)^{1/3}\right]^3} - 1 = \frac{1 + \eta_0}{1 + \eta} \frac{(b_1 - 1)(b_2 - 1)(b_3 - 1)\eta}{\left((b_1 b_2 b_3 \eta)^{1/3} - \eta_0^{1/3}\right)^3} - 1, \tag{S111}$$

where $\eta_3 = \eta$. The condition for minimizing the cost is

$$\frac{\eta_0}{\eta_1 b_1} = \frac{\eta_1}{\eta_2 b_2} = \frac{\eta_2}{\eta b_3} = \left(\frac{\eta_0}{b_1 b_2 b_3 \eta}\right)^{1/3} \tag{S112}$$

$$\Rightarrow \eta_1 = \eta_0^{2/3} \eta^{1/3} \left(\frac{b_2 b_3}{b_1^2}\right)^{1/3}, \quad \eta_2 = \eta_0^{1/3} \eta^{2/3} \left(\frac{b_3^2}{b_1 b_2}\right)^{1/3}. \tag{S113}$$

Note that the error rates no longer form a geometric series. Instead, their ratios are modulated by factors $b_{1,2,3}$. The optimal partition ratios are:

$$a_1 = \frac{\eta_0 - \eta_1}{\eta_1 b_1 - \eta_0}, \quad a_2 = \frac{\eta_1 - \eta_2}{\eta_2 b_2 - \eta_1}, \quad a_3 = \frac{\eta_2 - \eta}{\eta b_3 - \eta_2}. \tag{S114}$$

where $\eta_{1,2}$ take the optimal values indicated in Eq. S113. All three partition ratios decrease as the error η is increased. In the n-stage DBD scheme, the partition ratios are equal, and they go to zero simultaneously at $\eta_{\mathrm{eq}} = f^{-1}$. For the IleRS network, however, the three partition ratios are not equal, and one of them vanishes first. This takes place in the proofreading pathway with the least b, which is b_1 in the IleRS network:

$$a_1 = 0 \Leftrightarrow \eta_0 = \eta_1 \Leftrightarrow \eta = \eta_{\mathrm{th1}} = \eta_0 \frac{b_1^2}{b_2 b_3}. \tag{S115}$$

For error rates greater than the threshold η_{th1}, the above calculation leads to a negative partition ratio ($a_1 < 0$), which must be regularized to zero.

Hence, the three-stage proofreading analysis only applies to $\eta \in \left(\frac{\eta_0}{b_1 b_2 b_3}, \frac{\eta_0 b_1^2}{b_2 b_3}\right)$. For larger error, the first proofreading pathway does not function ($a_1 = 0$), and we treat the system as two MM-with-proofreading schemes operating in tandem. Similarly, an error-cost bound can be obtained:

$$C_{\min} = \frac{1 + \eta_0}{1 + \eta} \frac{(b_2 - 1)(b_3 - 1)\eta}{\left((b_2 b_3 \eta)^{1/2} - \eta_0^{1/2}\right)^2} - 1, \quad \frac{\eta_0 b_1^2}{b_2 b_3} < \eta < \frac{\eta_0 b_2}{b_3}. \tag{S116}$$

The maximum error for this two-pathway regime is determined by $a_2 = 0$, which leads to $\eta = \eta_{\mathrm{th2}} = \frac{\eta_0 b_2}{b_3}$. For error larger than this value, we have $a_1 = a_2 = 0$, and the optimal system operates as if there is only one proofreading pathway:

$$C_{\min} = \frac{1 + \eta_0}{1 + \eta} \frac{(b_3 - 1)\eta}{b_3 \eta - \eta_0} - 1, \quad \frac{\eta_0 b_2}{b_3} < \eta < \eta_0. \tag{S117}$$

Therefore, we have derived the piecewise error-cost bound for the IleRS network, which is in agreement with the numeric sampling (Fig. 6C, main text).

b. Analysis of the native system. In the IleRS network, the error rate thresholds which separate the three proofreading regimes are

$$\eta_0 = 9.2 \times 10^{-3}, \quad \eta_{\mathrm{th2}} = \eta_0 \frac{b_2}{b_3} = 1.3 \times 10^{-4}, \quad \eta_{\mathrm{th1}} = \eta_0 \frac{b_1^2}{b_2 b_3} = 7.7 \times 10^{-5}, \quad \eta_{\min} = \frac{\eta_0}{b_1 b_2 b_3} = 4.5 \times 10^{-8}. \tag{S118}$$

The native system operates in the one-stage proofreading phase, where the optimal system only utilizes the last (post-transfer) proofreading pathway:

$$\eta_{\mathrm{wt}} = 2.2 \times 10^{-4} \in (\eta_{\mathrm{th2}}, \, \eta_0). \tag{S119}$$

The native partition ratios are

$$a_1 = \frac{f_{h1}}{f_3} = 3.4 \times 10^{-6}, \quad a_2 = \frac{f_{h2}}{f_4} = 3.6 \times 10^{-3}, \quad a_3 = \frac{f_{h3}}{f_p} = 8.9 \times 10^{-2}. \tag{S120}$$

The optimal partition ratios are

$$a_1 = a_2 = 0, \quad a_3 = \frac{\eta_0 - \eta_{\mathrm{wt}}}{\eta_{\mathrm{wt}} b_3 - \eta_0} = 3.9 \times 10^{-2}. \tag{S121}$$
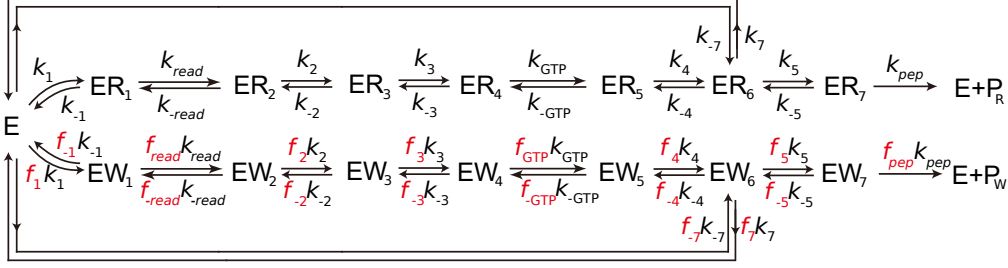
FIG. S4. Schematics of the detailed ribosome model [4].

Hence, it would seem that the first two proofreading pathways are not utilized, consistent with the theory prediction ($a_1 = a_2 = 0$). The last proofreading pathway is responsible for most of the proofreading, but the third-stage partition ratio in the native system (8.9×10^{-2}) is more than twice of its optimal value (3.9×10^{-2}). The reason for the extra proofreading is that η_0, which is the minimum error before proofreading, is never realized in the real system. It is only achieved if the amino acid activation step k_a is much slower than binding $k_\pm$, but such time scale separation is not realized in the native system. The error rate before proofreading, which in theory could be as low as $\eta_0 = f_a \frac{f_+}{f_-} = 9.2 \times 10^{-3}$, is actually $\eta_{\text{activation}} = 2.1 \times 10^{-2}$ in the native system (calculated by taking the ratio of net fluxes in the activation step). If we calculate the partition ratio with η_0 replaced by $\eta_{\text{activation}}$, the optimal partition ratio becomes $a_3' = 9.3 \times 10^{-2}$, which is indeed closer to the native system. The reason why η_0 could not be realized is similar to what was discussed in the main text about the ribosome network. η_0 could be approached by either speeding up binding/unbinding reactions or by slowing down the amino acid activation step. Reducing the activation rate, however, will slow down the speed of product formation. One possible interpretation is that while the binding and unbinding reactions are already as fast as possible, the native system chooses not to further decrease the activation rate so as to produce isoleucyl-tRNA$^{\text{Ile}}$ sufficiently fast, which necessitates additional proofreading in the post-transfer proofreading pathway.

Hence, the main conclusion here is that the deviation of the native IleRS from the optimal error-cost bound is due to prioritizing speed in the activation step. This is consistent with the trade-off analysis in the previous work [3], where k_a prefers to optimize speed rather than error or dissipation. If k_a (and the reverse reaction k_{-a}) becomes much slower than the binding and unbinding rates $k_\pm$, the accuracy before proofreading will be improved, which will lead to a smaller partition ratio a_3 and lower cost C.

D. Detailed model of the ribosome

The ribosome model presented in the main text was based on previous theoretical work [1] and experimental work [5]. Here, we apply our theoretical framework to study another model of the ribosome, which is based ref. [4]. The reaction scheme is shown in Fig. S4. Compared to the ribosome model discussed in the main text, this model now includes multiple intermediate states. However, there is still only one proofreading pathway, namely the futile cycle containing k_7. For the sake of generality, we allow for discrimination in all reaction steps, subject to the thermodynamic constraint:

$$\frac{f_1 f_{\text{read}} f_2 f_3 f_{\text{GTP}} f_4 f_7}{f_{-1} f_{-\text{read}} f_{-2} f_{-3} f_{-\text{GTP}} f_{-4} f_{-7}} = 1. \tag{S122}$$

In the following, we derive the error-cost bound with the flux-based formalism detailed above.

For step i, the rate constant is denoted by k_i in the cognate network and $k_i' = k_i f_i$ in the noncognate network. The normalized flux is denoted by j_i in the cognate network and j_i' in the noncognate network. The fluxes forming products are $j_{\text{pep}} = 1$ and $j_{\text{pep}}' = \eta$, where η is the final error rate. The cost is defined by

$$C = \frac{j_7 + j_7' - j_{-7} - j_{-7}'}{j_{\text{pep}} + j_{\text{pep}}'} = \frac{j_7 + j_7' - j_{-7} - j_{-7}'}{1 + \eta}. \tag{S123}$$

Similar to the steps taken to derive the error-cost bound in previous sections of the SI, we establish the relation between fluxes inductively. We define the intermediate error rates:

$$\eta_{\text{read}} = \frac{j_{\text{read}}'}{j_{\text{read}}}, \quad \eta_2 = \frac{j_2'}{j_2}, \quad \eta_3 = \frac{j_3'}{j_3}, \quad \eta_{\text{GTP}} = \frac{j_{\text{GTP}}'}{j_{\text{GTP}}}, \quad \eta_4 = \frac{j_4'}{j_4}, \quad \eta_5 = \frac{j_5'}{j_5}. \tag{S124}$$

All the noncognate fluxes $\{j'\}$ can now be expressed in terms of the cognate fluxes $\{j\}$, the discrimination factors $\{f\}$, and the error rates $\{\eta\}$. The stationary conditions for states ER_7 and EW_7 read

$$j_5 - j_{-5} = 1, \quad j'_5 - j'_{-5} = \eta, \tag{S125}$$

where $j'_5 = \eta_5 j_5$ and $j'_{-5} = \eta \frac{f_{-5}}{f_{\text{pep}}} j_{-5}$. These equations lead to

$$\left(\eta - \eta_5 \frac{f_{\text{pep}}}{f_{-5}} \right) j_5 = \left(1 - \frac{f_{\text{pep}}}{f_{-5}} \right) \eta. \tag{S126}$$

Since $j_5 = 1 + j_{-5} > 1$, the intermediate error rate η_5 satisfies:

$$\eta_5 < \eta_{5,\text{max}} = \eta \cdot \max\left(1, \frac{f_{-5}}{f_{\text{pep}}} \right). \tag{S127}$$

The stationary conditions for states ER_6 and EW_6 read

$$j_4 - j_{-4} = j_7 - j_{-7} + j_5 - j_{-5}, \tag{S128}$$
$$j'_4 - j'_{-4} = j'_7 - j'_{-7} + j'_5 - j'_{-5}, \tag{S129}$$

where $j'_4 = \eta_4 j_4$, $j'_{-4} = \eta_5 f_{-4} f_5^{-1} j_{-4}$, $j'_7 = \eta_5 f_5^{-1} f_7 j_7$, $j'_{-7} = f_{-7} j_{-7}$. Elimination of the forward proofreading flux j_7 yields

$$\left(\eta_5 \frac{f_7}{f_5} - \eta_4 \right) j_4 = \eta_5 \frac{f_7 - f_{-4}}{f_5} j_{-4} + \left(f_{-7} - \eta_5 \frac{f_7}{f_5} \right) j_{-7} + \left(\frac{f_7}{f_5} \eta_5 - \eta \right). \tag{S130}$$

Based on analysis employed in previous models, error rates $\eta_5 > f_5 \frac{f_{-7}}{f_7}$ could be achieved without any proofreading (by making the k_5 step rate-limiting). Hence, we study the cost for error rates $\eta_5 < f_5 \frac{f_{-7}}{f_7}$. The coefficient $\left(f_{-7} - \eta_5 \frac{f_7}{f_5} \right)$ is positive. Proofreading preferentially dissociates noncognate complexes, indicating $f_7 > f_{-4}$ and $f_7 > f_5$ (which is the case for experimental data). Thus, LHS must also be positive, leading to

$$\eta_5 > \eta_4 \frac{f_5}{f_7}. \tag{S131}$$

The minimum j_4 is

$$j_4 \geq \frac{\left(\frac{f_7}{f_5} \eta_5 - \eta \right) + \eta_5 \frac{f_7 - f_{-4}}{f_5} j_{-4}}{\eta_5 \frac{f_7}{f_5} - \eta_4}, \tag{S132}$$

with equality condition $j_{-7} \to 0$. The cost is

$$C = \frac{j_7 + j'_7 - j_{-7} - j'_{-7}}{1 + \eta} = \frac{j_4 + j'_4 - j_{-4} - j'_{-4}}{1 + \eta} - 1 \tag{S133}$$

$$= \frac{(1 + \eta_4) j_4 - \left(1 + \frac{\eta_5 f_{-4}}{f_5} \right) j_{-4}}{1 + \eta} - 1 \tag{S134}$$

$$\geq \frac{(\eta_4 - \eta) \left(1 + \frac{f_7}{f_5} \eta_5 \right)}{(1 + \eta) \left(\frac{f_7}{f_5} \eta_5 - \eta_4 \right)} + \frac{\left(\eta_4 - \eta_5 \frac{f_{-4}}{f_5} \right) \left(1 + \frac{f_7}{f_5} \eta_5 \right)}{(1 + \eta) \left(\frac{f_7}{f_5} \eta_5 - \eta_4 \right)} j_{-4} \tag{S135}$$

For error rates satisfying $\eta_4 > \eta_5 f_5^{-1} f_{-4}$, the proofreading pathway is unnecessary, and the minimum cost is zero. For error rates satisfying $\eta_4 > \eta_5 f_5^{-1} f_{-4}$, the cost is minimized in the limit $j_{-4} \to 0$:

$$C \geq \frac{(\eta_4 - \eta) \left(1 + \frac{f_7}{f_5} \eta_5 \right)}{(1 + \eta) \left(\frac{f_7}{f_5} \eta_5 - \eta_4 \right)}. \tag{S136}$$

The above minimum cost increases with η_4 but decreases with η_5. The maximum value of η_5 is given by:

$$\eta_{5,\mathrm{max}} = \eta \cdot \max\left(1, \frac{f_{-5}}{f_{\mathrm{pep}}}\right). \tag{S137}$$

The minimum value of η_4 is determined by the maximal difference in the energy barriers along the chain of reversible reactions from state E to state $\mathrm{ER}_6/\mathrm{EW}_6$:

$$\eta_4 > \eta_{4,\mathrm{min}} = e^{-\Delta\mu_{\mathrm{max}}} \tag{S138}$$

$$= \min\left(f_1, \frac{f_1 f_{\mathrm{read}}}{f_{-1}}, \frac{f_1 f_{\mathrm{read}} f_2}{f_{-1} f_{-\mathrm{read}}}, \frac{f_1 f_{\mathrm{read}} f_2 f_3}{f_{-1} f_{-\mathrm{read}} f_{-2}}, \frac{f_1 f_{\mathrm{read}} f_2 f_3 f_{\mathrm{GTP}}}{f_{-1} f_{-\mathrm{read}} f_{-2} f_{-3}}, \frac{f_1 f_{\mathrm{read}} f_2 f_3 f_{\mathrm{GTP}} f_4}{f_{-1} f_{-\mathrm{read}} f_{-2} f_{-3} f_{-\mathrm{GTP}}}\right). \tag{S139}$$

Therefore, we have derived the error-cost bound:

$$C_{\mathrm{min}} = \frac{(\eta_{4,\mathrm{min}} - \eta)\left(1 + \frac{f_7}{f_5}\eta_{5,\mathrm{max}}\right)}{(1 + \eta)\left(\frac{f_7}{f_5}\eta_{5,\mathrm{max}} - \eta_{4,\mathrm{min}}\right)}, \quad \eta \in (\eta_{\mathrm{min}}, \eta_{\mathrm{eq}}), \tag{S140}$$

where $\eta_{4,\mathrm{min}}$ and $\eta_{5,\mathrm{max}}$ are given by Eq. S139 and Eq. S137, respectively. The minimum error is

$$\eta_{\mathrm{min}} = \frac{\eta_{4,\mathrm{min}}}{f_7} \cdot \min\left(1, \frac{f_{\mathrm{pep}}}{f_{-5}}\right), \tag{S141}$$

and the minimum error without proofreading is

$$\eta_{\mathrm{eq}} = \min\left(\eta_{4,\mathrm{min}}, \frac{f_1 f_{\mathrm{read}} f_2 f_3 f_{\mathrm{GTP}} f_4 f_5}{f_{-1} f_{-\mathrm{read}} f_{-2} f_{-3} f_{-\mathrm{GTP}} f_{-4}}, \frac{f_1 f_{\mathrm{read}} f_2 f_3 f_{\mathrm{GTP}} f_4 f_5 f_{\mathrm{pep}}}{f_{-1} f_{-\mathrm{read}} f_{-2} f_{-3} f_{-\mathrm{GTP}} f_{-4}}\right). \tag{S142}$$

Thus, the flux-based formalism could be used to fully determine the fundamental error-cost bound in this detailed kinetic model of the ribosome. The methodology is completely same as that used in for the other models, and the cost-error bound exhibits similar quantitative behavior. In fact, the mathematical form of this bound (Eq. S140) is similar to that of the ribosome model studied in the main text (Eq. S101).

REFERENCES

[1] K. Banerjee, A. B. Kolomeisky, and O. A. Igoshin, Elucidating interplay of speed and accuracy in biological error correction, Proc. Natl. Acad. Sci. U. S. A. **114**, 5183 (2017).

[2] J. D. Mallory, A. B. Kolomeisky, and O. A. Igoshin, Trade-offs between error, speed, noise, and energy dissipation in biological processes with proofreading, J. Phys. Chem. B **123**, 4718 (2019).

[3] Q. Yu, J. D. Mallory, A. B. Kolomeisky, J. Ling, and O. A. Igoshin, Trade-offs between speed, accuracy, and dissipation in tRNA$^{\mathrm{Ile}}$ aminoacylation, J. Phys. Chem. Lett. **11**, 4001 (2020).

[4] I. Wohlgemuth, C. Pohl, J. Mittelstaet, A. L. Konevega, and M. V. Rodnina, Evolutionary optimization of speed and accuracy of decoding on the ribosome, Philos. Trans. R. Soc. Lond., B, Biol. Sci. **366**, 2979 (2011).

[5] H. S. Zaher and R. Green, Hyperaccurate and error-prone ribosomes exploit distinct mechanisms during tRNA selection, Mol. Cell **39**, 110 (2010).