A. DEFINITION OF MODEL AND PARAMETER CHOICES

A schematic of the model we are considering is shown in Fig. 1. As described in the main text, we denote a receptor-ligand complex that has been phosphorlyated \( n \) times by \( X_n \) with \( X = C \) for foreign ligands and \( X = D \) for self ligands. Furthermore, we denote the maximum number of phosphorylations as \( N \). With this notation, using the law of mass action, we have

\[
\begin{align*}
\dot{C}_0 &= \kappa R L_f - \left( \tau_f^{-1} + \phi \right) C_0 + b C_1 \\
\dot{C}_n &= \gamma^n R L_f / \tau_f + \phi C_{n-1} - \left( \phi + \tau_f^{-1} + b \right) C_n + b C_{n+1} \\
\dot{C}_N &= \gamma^N R L_f / \tau_f + \alpha K C_{N-1} - \left( b + \tau_f^{-1} \right) C_N \\
\dot{K} &= -\epsilon K (C_m + D_m) + \sigma (K_T - K) \\
\dot{D}_0 &= \kappa R L_s - \left( \tau_s^{-1} + \phi \right) D_0 + b D_1 \\
\dot{D}_n &= \gamma^n R L_s / \tau_s + \phi D_{n-1} - \left( \phi + \tau_s^{-1} + b \right) D_n + b D_{n+1} \\
\dot{D}_N &= \gamma^N R L_s / \tau_s + \alpha K D_{N-1} - \left( b + \tau_s^{-1} \right) D_N
\end{align*}
\]

where \( N > n > 0 \), \( R = R_T - \sum_{j=0}^{N} (C_j + D_j) \approx R_T \), \( L_s = L_s^T - \sum_{j=0}^{N} D_j \) and \( L_f = L_f^T - \sum_{j=0}^{N} C_j \). Typically, we set: \( R_T = 10^4 \), \( L_s^T = L_f^T = 10^4 \), \( \kappa = 300 \text{s}^{-1} \), \( \sigma = 1 \text{s}^{-1} \), \( \epsilon = 2 \text{s}^{-1} \), \( K_T = 10^4 \), \( \alpha = 3 \times 10^{-4} \), \( \gamma = 10^{-3} \), \( \tau_s = 1 \text{s} \) and \( \tau_f = 10 \text{s} \). Any deviations from this choice of parameter is explicitly noted.

Accuracy

At steady state, the error rate can be written as

\[
\eta = \frac{D_N}{C_N}
\]

In the presence of the kinase feedback \( K = \frac{\sigma K_T}{\sigma + \epsilon (C_2 + D_2)} \), the set of eqs. (1) are no longer linear and but the steady-state solution can still be found easily using an iterative method.
Energy Consumption

The power dissipation is calculated based on the net flux and the chemical potential difference\[1, 2\]. We define the net flux \(J_{n,i}, i \in [s, f]\) at \(X_n = X_{n+1}\) in the main pathway.

\[
J_{i,n} = \begin{cases} 
\phi X_n - bX_{n+1}, & \text{for } 0 \leq n < N - 1 \\
\alpha K X_{N-1} - bX_N, & \text{for } n = N 
\end{cases}
\]

Considering the flux conservation, the power dissipation \(P_i\) can be written as

\[
P_i = k_B T J_{i,0} \frac{\kappa R_i^L_{free}}{\tau_i^{-1} X_0} + k_B T \sum_{n=0}^{N-2} J_{i,n} \ln \frac{\phi X_i}{bX_{i+1}} + k_B T J_{i,N-1} \ln \frac{\alpha K X_{N-1}}{bX_N} \\
+ k_B T \sum_{n=0}^{N-2} (J_{i,n} - J_{i,n+1}) \ln \frac{X_{n+1}}{\gamma X_i R_i^L_{free}} + k_B T J_{i,N-1} \ln \frac{X_N}{\gamma N R_i^L_{free}}
\]

\[
= k_B J_{i,0} \ln \frac{\kappa}{\tau_i^{-1}} + k_B T \sum_{n=0}^{N-2} J_{i,n} \ln \frac{\phi}{b\gamma} + k_B T J_{i,N-1} \ln \frac{\alpha K}{b\gamma}
\]

The total power dissipation is from the contribution of both foreign and self ligands: \(P = P_s + P_f\).

Role of \(\gamma\)

In KPR, the reversible decay rate is ignored as it has extremely small value. Let \(\gamma_{n,i}\) denote the rate at which a self \((i = s)\) or foreign ligand \((i = f)\) can directly form a complex at \(n - th\) step of the KPR cascade (see Fig. 2 of main text). In such a reaction, the first \(n - 1\) steps of the KPR cascade are bypassed resulting in lower accuracy. There are several natural choices for how to choose \(\gamma_{n,i}\). One common choice in the literature is to assume that \(\gamma_{n,i}\) is independent of \(n\) and given by \(\gamma_{n,i} = \gamma/\tau_i\). However, with this choice never saturates the KPR accuracy bound for an N-step cascade, \(\eta_{min} = \tau_s^N/\tau_f^N\), especially when \(N\) is large (see Fig 2).

For this reason, in this work we choose a step-dependent rate, \(\gamma_{n,i} = \gamma^n/\tau_i\) \((i = s, f)\), for directly forming a complex \(C_n\) and \(D_n\). This functional form is a direct consequence of assuming that there is a constant free energy difference \(k_B T \log \phi/\gamma b\) per phosphorylation. Having a large \(\gamma\) will result in a bypassing of the proofreading steps and a high error threshold for any KPR-based circuit.

One choice \(\gamma^n/\tau_i\). There are two reasons for this form: 1. the production rate from ligands and receptors to \(X_{n+1}\) should be smaller then the one to \(X_n\) as one more phosphorylation step is involved. If not, it is hard for the KPR circuit to achieve the theoretical limit, \(\tau_s^N/\tau_s^N\). 2. it is also natural to assume the energy consumption is the same for each phosphorylation step.

The free energy difference between \(n\)th and \(n+1\)th phosphorylation round can be calculated as:

\[
\Delta G_n = k_B T \log \left[ \frac{\kappa}{\tau_i} \frac{\phi^{n+1}}{b^{n+1}} \frac{\tau_i^{-1}}{\gamma^{n+1}/\tau_i} \right] - k_B T \log \left[ \frac{\kappa}{\tau_i} \frac{\phi^n}{b^n} \frac{\tau_i^{-1}}{\gamma^n/\tau_i} \right]
\]

\[
= k_B T \log \frac{\phi}{\gamma b}
\]

Speed

The speed is defined by the mean first passage time (MFPT) for the foreign ligand. Here we mainly follow the procedures in Ref. [3]. The concentration vector is defined as \(c = [L_f, C_0, C_1, \ldots, C_N, p_f]\). An final ‘dark’ state is added because the response is only activated at the end and it can be treated as absorbing markov chain. Added this absorb state, it becomes an irreversible process, which is impossible to calculate the energy consumption. The transfer probability from \(C_N\) to the ‘dark’ state is \(W(\text{irreversible})\). We set \(W = 100s^{-1}\), a large value, which means the final step has little effect on MFPT. Without loss of generality, we begin with \(N = 4\) and \(m = 2\), which can be
generalized other cases easily. The master equations i.e. eqs. (1) can be rewritten as \( \dot{c} = A \cdot c \) and

\[
A = \begin{bmatrix}
-\kappa R - \sum_{j=1}^{4} \gamma_j / \tau_f & \frac{1}{\tau_f} & \frac{1}{\tau_f} & \frac{1}{\tau_f} & \frac{1}{\tau_f} & 0 \\
\kappa R & -\frac{1}{\tau_f} - \phi & b & 0 & 0 & 0 \\
\gamma_j / \tau_f & \phi & -\phi - \frac{1}{\tau_f} - b & b & 0 & 0 \\
\gamma_j^2 / \tau_f & 0 & \phi & -\phi - \frac{1}{\tau_f} - b & b & 0 \\
\gamma_j^3 / \tau_f & 0 & 0 & \phi & -\alpha K - \frac{1}{\tau_f} - b & b \\
\gamma_j^4 / \tau_f & 0 & 0 & 0 & \alpha K - b & -\frac{1}{\tau_f} - W & 0 \\
0 & 0 & 0 & 0 & \alpha K - b & -\frac{1}{\tau_f} - W & 0 \\
\end{bmatrix}
\]

But eqs. (1) are not linear. The first order perturbation approximation is adapted and we can linearize (with bar denoting average) to get \( c = \bar{c} + \delta c \).

\[
\delta \dot{c} = A' \delta c, \quad \delta c = [\delta L_f, \delta C_0, \delta C_1, \ldots, \delta C_N, p_f]
\]

where \( A' \) is

\[
A' = \begin{bmatrix}
-\kappa R - \sum_{j=1}^{4} \gamma_j / \tau_f & \frac{1}{\tau_f} & \frac{1}{\tau_f} & \frac{1}{\tau_f} & \frac{1}{\tau_f} & 0 \\
\kappa R & -\frac{1}{\tau_f} - \phi & b & 0 & 0 & 0 \\
\gamma_j / \tau_f & \phi & -\phi - \frac{1}{\tau_f} - b & b & 0 & 0 \\
\gamma_j^2 / \tau_f & 0 & \phi & -\phi - \frac{1}{\tau_f} - b & b & 0 \\
\gamma_j^3 / \tau_f & 0 & 0 & \phi + \frac{\alpha K \sigma C_3}{D_2} & -\alpha K - \frac{1}{\tau_f} - b & b \\
\gamma_j^4 / \tau_f & 0 & 0 & 0 & \alpha K & -b - \frac{1}{\tau_f} - W & 0 \\
0 & 0 & 0 & 0 & \alpha K & -b - \frac{1}{\tau_f} - W & 0 \\
\end{bmatrix}
\]

Applying the Laplace transform, \( \delta C(s) = \int_0^\infty \delta c e^{-st} dt \), the master equations can be rewritten as:

\[(s - A) \delta C(s) = \delta c(t = 0) = [1, \ldots, 0]^T \]

The MFPT can be written:

\[
T = \int_0^\infty t p_f(t) dt = -\frac{d}{ds} \int_0^\infty p_f(t) e^{-st} dt \bigg|_{s=0} = -W \frac{d\delta C_N(s)}{ds} \bigg|_{s=0}
\]
which can be calculated numerically. It should be notified that the concentration and probability have the same master equations but a different pre-factor. When choosing the initial condition \( [1, \ldots, 0]^T \), the pre-factor is set to be 1 and \( \delta C_N(s) \) solved from eq. (7) is exactly a probability distribution.

**B. SIMULATION DETAILS FOR PHASE DIAGRAM**

In this figure, we run \( 10^6 \) samples with random sets \( \gamma \in [10^{-4}, 10] \), \( \phi \in [10^{-10}, 10^{10}]s^{-1} \), \( b \in [10^{-15}, 10^{15}]s^{-1} \).

It can be observed that a large amount of red points distributes over regimes C and D with \( \eta \sim 100 \). This is because of \( \gamma \sim 10 \) and the inverse flux at the final step dominates. In the extreme case: \( b/\phi \) is very large, \( C_0 \simeq D_0 \) will occupy most of products and free ligands \( L_s, L_f \) have little concentration.

\[
\frac{L_s}{L_f} = \frac{D_0 \tau_s^{-1}}{C_0 \tau_f^{-1}} = \frac{\tau_s^{-1}}{\tau_f^{-1}}
\]

As \( \gamma^N/\tau_i \) dominates,

\[
\eta = \frac{D_N}{C_N} = \frac{L_s \gamma^N/\tau_s}{L_f \gamma^N/\tau_f} = \frac{\tau_f^2}{\tau_s^2} = 100
\]

---

[1] T. L. Hill, *Free energy transduction and biochemical cycle kinetics* (Springer Science & Business Media, 2012).
[2] H. Qian, *Annu. Rev. Phys. Chem.* **58**, 113 (2007).
[3] G. Bel, B. Munsky, and I. Nemenman, *Physical biology* **7**, 016003 (2009).