Reformulating nonequilibrium steady states and generalized Hopfield discrimination

Uğur Çetiner and Jeremy Gunawardena*
Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02115, USA

(Received 8 August 2022; accepted 28 November 2022; published 20 December 2022)

Despite substantial progress in nonequilibrium physics, steady-state (s.s.) probabilities remain intractable to analysis. For a Markov process, s.s. probabilities can be expressed in terms of transition rates using the graph-theoretic matrix-tree theorem (MTT). The MTT reveals that, away from equilibrium, s.s. probabilities become globally dependent on all rates, with expressions growing exponentially in the system size. This overwhelming complexity and lack of thermodynamic interpretation have greatly impeded analysis. Here we show that s.s. probabilities are proportional to the average of $\exp(-S(P))$, where $S(P)$ is the entropy generated along minimal paths P in the graph, and the average is taken over a probability distribution on spanning trees. Assuming Arrhenius rates, this "arboreal" distribution becomes Boltzmann-like, with the energy of a tree being its total edge barrier energy. This reformulation offers a thermodynamic interpretation that smoothly generalizes equilibrium statistical mechanics and reorganizes the expression complexity: the number of distinct minimal-path entropies depends not on graph size but on the entropy production index, a measure of nonequilibrium complexity. We demonstrate the power of this reformulation by extending Hopfield's analysis of discrimination by kinetic proofreading to any graph with index 1. We derive a general formula for the error ratio and use it to show how local energy dissipation can yield optimal discrimination through global synergy.

DOI: 10.1103/PhysRevE.106.064128

I. INTRODUCTION

The foundations of nonequilibrium physics remain under vigorous development, following impressive breakthroughs in proving exact fluctuation theorems for systems arbitrarily far from thermodynamic equilibrium [1–4]. This continuing progress has fostered significant interactions with biology [5–8]. Life itself is quintessentially far from equilibrium. However, despite considerable knowledge of the molecular pathways of energy transduction, the functional significance of energy expenditure has been harder to unravel, especially for cellular information processing. In this setting, one of the barriers to progress has been the intractability to exact analysis of the probabilities of mesoscopic states, even at steady state.

The problem becomes particularly clear from a Markovian perspective, as usually adopted in stochastic thermodynamics [9]. Here, we exploit the "linear framework," a graph-theoretic approach to Markov processes developed in previous work [10,11]. Graph vertices represent mesostates, directed edges represent transitions, and edge labels represent transition rates [more details below, and see Fig. 1(a) for an example]. Such graphs correspond to continuous-time, finite-state Markov processes, for which the graphs are the infinitesimal generators. This framework extends the earlier approach of Hill [12,13] and Schnakenberg [14] by treating the graph as a mathematical object in its own right. It thereby provides a rigorous way to prove theorems that rise above the molecular complexity of biology [15–18].

The steady-state (s.s.) probabilities of a Markov process can be expressed in terms of the corresponding graph edge labels using certain subgraphs—spanning trees (Fig. 1)—as described by the matrix-tree theorem [MTT, Eq. (3)]. The MTT shows that as soon as the system is away from thermodynamic equilibrium, even if that occurs through energy expenditure at only a single edge, s.s. probabilities become globally dependent on all edge labels in the graph. The resulting expressions, which depend on enumerating all spanning trees, become extremely complex. Consider, for example, a graph whose mesostates correspond to the presence or absence at k sites of some feature such as a post-translational modification, so that there are 2^k mesostates. For $k = 2$, there are four spanning trees; for $k = 3$ there are 384 spanning trees; but for $k = 4$ there are 42 467 328 spanning trees [15]. All of these trees are required to exactly calculate s.s. probabilities. Moreover, while the mathematical details are clear, a thermodynamic interpretation has been lacking. We have not been able to see the wood for the trees! The combinatorial complexity and lack of thermodynamic meaning have greatly hindered nonequilibrium calculations.

We offer a solution to both these difficulties. Since the complexity cannot be avoided, it must be reorganized and reinterpreted. Consider any minimal path P from a given vertex i to a reference vertex 1. A minimal path has no repeated vertices and so there are only finitely many such paths. Detailed balance tells us that at thermodynamic equilibrium $S(P)$, the overall entropy production along P is independent of the choice of P. This is equilibrium statistical mechanics in the graph setting. We define a probability distribution, the "arboreal distribution," on spanning trees rooted at 1. We write edge labels in Arrhenius form, in terms of a vertex energy and an edge barrier energy; the constructions are independent of the choices involved. Surprisingly, in view of the nonequilibrium setting, the arboreal distribution is Boltzmann-like, with

*jeremy@hms.harvard.edu

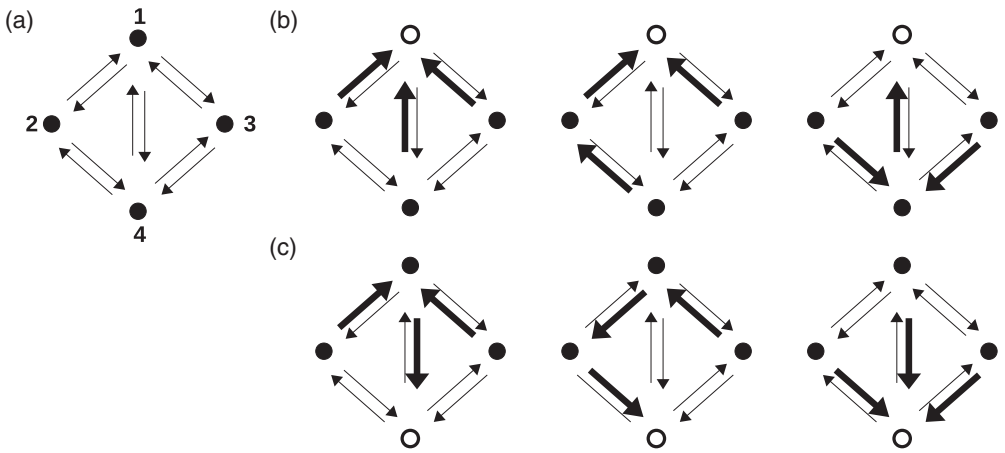

FIG. 1. Graphs and spanning trees. (a) Reversible linear framework graph, with named vertices and labels omitted for clarity. (b) Three spanning trees (thick edges) rooted at vertex 1 (open disk), chosen from eight possibilities. (c) Corresponding spanning trees rooted at vertex 4, obtained by applying the map $\Phi_{1,4}$, as described in the text, to the tree vertically above in panel (b), with the same format as (b).

$\Pr(T) \propto e^{-E(T)}$, where the energy $E(T)$ is the sum of the barrier energies on all edges of the tree T [Eq. (9)]. Each spanning tree rooted at 1 yields a unique path from i to 1 for each vertex i. Our main result is that, away from thermodynamic equilibrium, the s.s. probability of vertex i is proportional to the average of $\exp(-S(P))$, where P runs over the minimal paths from i to 1 and the average is taken over the arboreal distribution [Eq. (8)].

This reformulation provides a thermodynamic interpretation in place of a forest of trees and smoothly generalizes equilibrium statistical mechanics. It also finesses the combinatorial explosion: The number of distinct values for the entropy production on minimal paths does not depend on the size of the graph; rather, it depends on how many edges are experiencing energy expenditure or, more precisely, on the entropy production index $\iota(G)$ of a graph G, a measure of nonequilibrium complexity that we introduce here. This revised scaling dramatically simplifies the calculation of s.s. probabilities.

To illustrate this point, we substantially generalize Hopfield's classic study of kinetic proofreading [19]. Hopfield analyzed a simple graph with three vertices. We analyze any graph in which energy is expended at only one edge or, equivalently, for which $\iota(G) = 1$, and give a general formula for the error ratio [Eq. (18)]. We exploit this formula to show that optimal discrimination is possible even in complex graphs, despite energy being expended at only one edge. One of the messages of this paper is that systems for which $\iota(G) = 1$, although being away from equilibrium and suffering all the problems of global parameter dependence and combinatorial complexity, are nevertheless algebraically tractable through the reformulation of steady-state probabilities introduced here.

II. BACKGROUND ON THE LINEAR FRAMEWORK

A. Steady-state probabilities

We briefly describe the linear framework. More details and background can be found in [10,11,15,17,20–22]. Let G denote a finite, directed graph with labeled edges and no self-loops [Fig. 1(a)]. We denote the vertices of G by the indices, $1, 2, \ldots, n$, an edge from i to j by $i \to j$, and the label on this edge by $\ell(i \to j)$. The vertices are mesostates of the system under study, implying thereby that they are coarse-grained abstractions of the underlying physical microstates. The edges correspond to transitions between mesostates with the label being the transition rate with dimensions of $(\text{time})^{-1}$. Labels may be complex expressions which describe interactions between mesostates and environmental reservoirs, such as those for molecules (particles) or heat. Here, we make the customary thermodynamic assumption that exchanges between the graph and the reservoirs do not change the thermodynamic potentials of the reservoirs. Edge labels may then be treated as constants.

G describes the infinitesimal generator of a continuous-time Markov process $X(t)$, given by a conditional probability distribution on the same mesostates for times $s > t$, $\Pr(X(s) = j \,|\, X(t) = i)$. The edge labels are those infinitesimal transition rates,

$$\ell(i \to j) = \lim_{\Delta t \to 0} \frac{\Pr(X(t + \Delta t) = j \,|\, X(t) = i)}{\Delta t}, \qquad (1)$$

which are not zero. Provided the limits in Eq. (1) exist, there is an exact correspondence between Markov processes and graph representations [21]. In particular, the master equation of the Markov process can be recovered from the graph. Let $u_i(t)$ denote the probability of mesostate i at time t, $u_i(t) = \Pr(X(t) = i)$. The master equation is the linear matrix equation

$$\frac{du(t)}{dt} = \mathcal{L}(G)u(t), \qquad (2)$$

where $\mathcal{L}(G)$ is the $n \times n$ Laplacian matrix of G, and $u(t) = (u_1(t), \ldots, u_n(t))^T$ is the $n \times 1$ column vector of probabilities [10].

Since Eq. (2) is linear, there is no difficulty in solving it in terms of eigenvalues, but these are not known in terms of the edge labels, at least for $n \geqslant 5$. However, the s.s. probabilities of the mesostates, denoted $u^*(G)$, can be expressed in terms of the labels. If H is a subgraph of G, let $q(H)$ denote the product of the labels on the edges of H: $q(H) = \prod_{i \to j \in H} \ell(i \to j)$. Let $\Theta_i(G)$ denote the set of spanning trees of G rooted at i.

A spanning tree is a subgraph which includes each vertex of G (spanning) and has no cycles if edge directions are ignored (tree); it is rooted at i if the tree has no edges outgoing from i [Fig. 1(b)]. Provided G is strongly connected, so that any two vertices, i and j, are joined by a directed path, $i = i_1 \to i_2 \to \cdots \to i_k = j$, there exist rooted spanning trees at each vertex. Moreover, the kernel of $\mathcal{L}(G)$ is one-dimensional. A canonical basis element, $\rho(G) \in \ker \mathcal{L}(G)$, is given by the MTT,

$$\rho_i(G) = \sum_{T \in \Theta_i(G)} q(T). \tag{3}$$

Since $u^*(G) \in \ker \mathcal{L}(G)$, it follows that $u^*(G) \propto \rho(G)$. The proportionality constant comes from solving for total probability, $u_1^*(G) + \cdots + u_n^*(G) = 1$, which gives

$$u_i^*(G) = \frac{\rho_i(G)}{\rho_1(G) + \cdots + \rho_n(G)}. \tag{4}$$

B. Path entropies and thermodynamic equilibrium

We assume from now on that G is reversible: if $i \to j$, then also $j \to i$, and furthermore, the reverse edge represents the reverse process to the forward edge and not simply some alternative process for moving between the mesostates [23]. The log label ratio, $\ln[\ell(i \to j)/\ell(j \to i)]$, is then the total entropy change in taking the transition from i to j: The entropy change in the reservoirs together with the internal entropy difference between j and i. (Here, we assume $k_B = 1$ for convenience.) This form of "local detailed balance" goes back to Hill and Schnakenberg and has been broadly justified within stochastic thermodynamics [9,24,25]. Let $R(i, j)$ denote the set of reversible paths, $i = i_1 \leftrightarrows i_2 \leftrightarrows \cdots \leftrightarrows i_k = j$, from i to j. If $P \in R(i, j)$ is such a path, let $S(P)$ denote the total entropy change, as above, along the path. Evidently,

$$S(P) = \ln\left[\left(\frac{\ell(i_1 \to i_2)}{\ell(i_2 \to i_1)}\right) \cdots \left(\frac{\ell(i_{k-1} \to i_k)}{\ell(i_k \to i_{k-1})}\right)\right]. \tag{5}$$

If $P \in R(i, j)$, let $P^* \in R(j, i)$ denote the reverse path, so that $S(P^*) = -S(P)$.

If the graph can reach thermodynamic equilibrium, an alternative basis element, $\mu(G) \in \ker \mathcal{L}(G)$, exists. In this case, detailed balance holds: each pair of reversible edges, $i \leftrightarrows j$, is in s.s. flux balance, so that $u_i^*(G)\ell(i \to j) = u_j^*(G)\ell(j \to i)$. Equivalently, given any cycle of reversible edges, $Q \in R(i, i)$, $S(Q) = 0$. The latter equality is the *cycle condition*: at thermodynamic equilibrium, given any cycle of reversible edges, the product of the edge labels in one direction around the cycle is equal to the product of the edge labels in the reverse direction. Hence, if $P_1, P_2 \in R(i, j)$, then $S(P_1) = S(P_2)$. We can then define $\mu_i(G) = \exp(-S(P))$ for any $P \in R(i, 1)$. As before, $u^*(G) \propto \mu(G)$, which gives the following specification for equilibrium steady-state probabilities:

$$u_i^*(G) \propto \exp(-S(P)), \tag{6}$$

for any $P \in R(i, 1)$. A similar formula to Eq. (4) holds, with μ in place of ρ. This formula is the prescription of equilibrium statistical mechanics, with the denominator, $\mu_1(G) + \cdots + \mu_n(G)$, being the partition function for the grand canonical ensemble.

III. RESULTS

A. Reformulating steady-state probabilities

Path entropies enable the first step in reformulating Eq. (3). Following [17], let $\Phi_{i,j} : \Theta_i(G) \to \Theta_j(G)$ be defined as follows. Choose $T \in \Theta_i(G)$. By construction, there is a unique path in T from j to i. Since it has no repeated vertices, this path is minimal. Reversing the edges on this minimal path yields a spanning tree rooted at j, which is $\Phi_{i,j}(T) \in \Theta_j(G)$ [Fig. 1(c)]. $\Phi_{i,j}$ is a bijection—there are the same number of spanning trees at each vertex of a reversible graph—and $\Phi_{i,j}^{-1} = \Phi_{j,i}$; see [17] for a proof. Let $M(i, j) \subseteq R(i, j)$ be the set of minimal paths from i to j. While $R(i, j)$ is infinite, $M(i, j)$ is finite. If we focus on the reference vertex and consider any $T \in \Theta_1(G)$, let $T_i \in M(i, 1)$ be the unique minimal path, as in the definition of $\Phi_{1,i}$. It is easy to see that [17] $q(\Phi_{1,i}(T)) = \exp(S(T_i^*))q(T)$. Because $\Phi_{i,j}$ is a bijection, we can rewrite Eq. (3) for any vertex i in terms of only the spanning trees rooted at 1. Recalling that $S(T_i^*) = -S(T_i)$, we see that

$$\rho_i(G) = \sum_{T \in \Theta_1(G)} \exp(-S(T_i))q(T). \tag{7}$$

We define the arboreal probability distribution on $\Theta_1(G)$ by normalizing $q(T)$ to its total over all trees T, so that $\Pr_{\Theta_1(G)}(T) = q(T)/(\sum_{T \in \Theta_1(G)} q(T))$. This arboreal distribution has been previously studied (Discussion). It follows from Eq. (7) that $\rho_i(G) \propto \langle \exp(-S(T_i)) \rangle$, where the average is taken over the arboreal distribution. Since $u^*(G) \propto \rho(G)$ in $\ker \mathcal{L}(G)$, we see that

$$u_i^*(G) \propto \langle \exp(-S(T_i)) \rangle_{\Theta_1(G)}. \tag{8}$$

An easy consequence of Eq. (8) is that

$$\min(S(T_i)) \leqslant \ln\left(\frac{u_1^*(G)}{u_i^*(G)}\right) \leqslant \max(S(T_i)),$$

where the extrema are taken over $T \in \Theta_1(G)$. Maes *et al.* derive these bounds [26, Cor. 2.2] but without the probabilistic interpretation in Eq. (8).

To interpret the arboreal distribution thermodynamically, we express edge labels in Arrhenius form, $\ell(i \to j) = \exp(\epsilon_i - W_{i \to j})$. The quantity ϵ_i can be thought of as a vertex energy for mesostate i and $W_{i \to j}$ as the resulting barrier energy of the edge from i to j, expressed in units of $k_B T$, where T is the temperature of the reservoir with which the edge $i \to j$ is exchanging particles or heat. In general, $W_{i \to j} \neq W_{j \to i}$. Such a representation is always numerically possible but is not unique. Choose any Arrhenius representation and let $T \in \Theta_1(G)$. Let $E(T)$ be the total edge barrier energy,

$$E(T) = \sum_{i \to j \in T} W_{i \to j}. \tag{9}$$

Since $q(T) = \exp(\sum_{1 \leqslant i \leqslant n} \epsilon_i) \exp(-E(T))$, and the first term is independent of T, the arboreal distribution may be expressed in terms of $E(T)$ as

$$\Pr_{\Theta_1(G)}(T) = \frac{\exp(-E(T))}{\sum_{T \in \Theta_1(G)} \exp(-E(T))}. \tag{10}$$

Equation (10) is independent of the choice of Arrhenius rates. It reveals the arboreal distribution to be "Boltzmann-like,"

with the energy of a spanning tree being the total edge barrier energy over the tree.

Equations (8) and (10) constitute our reformulation of s.s. probabilities. In contrast to the expression coming from the MTT in Eq. (3), which lacks thermodynamic meaning, Eq. (8) smoothly generalizes the equilibrium formula in Eq. (6). At equilibrium, s.s. probabilities are given by path entropies: $u_i^* \propto \exp(-S(P))$. Away from equilibrium they are given by averages over path entropies: $u_i^* \propto \langle \exp(-S(T_i)) \rangle$, where the average is calculated over the arboreal distribution. At equilibrium, the entropies of all paths in $R(i,1)$ are identical; the arboreal distribution factors out and Eq. (8) reduces to Eq. (6).

B. Combinatorial scaling by energetic edges

Equation (8) has a further important advantage. Unlike spanning trees, minimal path entropies do not scale with the size of the graph. Suppose that G satisfies detailed balance and let $\ell_{eq}(i \to j)$ denote the edge labels under this condition. Suppose that edge labels are altered to break detailed balance and the new labels are given by $\ell(i \to j) = m(i \to j)\ell_{eq}(i \to j)$. We will say that $i \to j$ is an *energetic edge* if $m(i \to j) \neq 1$. Let $P : v = v_1 \leftrightarrows \cdots \leftrightarrows v_k = w \in R(v,w)$ and define $F(P)$ to be the set of energetic edges in the forward direction of P,

$$F(P) = \{v_l \to v_{l+1} \,|\, m(v_l \to v_{l+1}) \neq 1\}. \tag{11}$$

The set of energetic edges in the reverse direction is then $F(P^*)$. It follows from Eq. (5) that

$$S(P) = \ln\left[\frac{\prod_{i\to j \in F(P)} m(i \to j)}{\prod_{i\to j \in F(P^*)} m(i \to j)}\right] + S_{eq}(P), \tag{12}$$

where $S_{eq}(P)$ is the total entropy change along P at thermodynamic equilibrium. As noted previously, $S_{eq}(P)$ is independent of $P \in R(v,w)$. Given $P_1, P_2 \in R(v,w)$, we will say that P_1 is *energetically similar* to P_2, denoted $P_1 \sim P_2$, if $F(P_1) = F(P_2)$ and $F(P_1^*) = F(P_2^*)$. It follows from Eq. (12) that if $P_1 \sim P_2$, then $S(P_1) = S(P_2)$. Hence the number of distinct minimal path entropies in Eq. (8) is independent of the size of the graph and depends only on the number of energetic edges. This scaling still incurs a combinatorial increase, since minimal paths may have different subsets of energetic edges but the scaling is substantially less intimidating than that arising from all rooted spanning trees (above). We examine below the implications of this scaling for a graph with a single energetic edge.

C. The entropy production index

Identifying a single energetic edge relies on the prior choice of a state of thermodynamic equilibrium. Energy expenditure could also occur from coupling at multiple edges to reservoirs with different chemical potentials [27]. We therefore sought a more general way to exploit the reformulation in Eq. (8). Consider any basis B of oriented minimal cycles, $B = \{B_1, \cdots, B_q\}$ [14]. Such a basis arises by choosing a spanning tree and adjoining an edge outside the tree to create each minimal cycle for which an arbitrary orientation is chosen (see Appendixes). The number q of minimal cycles is the first Betti number of G, $\beta_1(G)$, or the rank of the first homology group of G considered as a topological space. Let $\iota(B)$ denote the number of cycles B_i which break the cycle condition, $\iota(B) = \#i,\ 1 \leqslant i \leqslant q, S(B_i) \neq 0$. Define the entropy production index of G to be $\iota(G) = \min_B \iota(B)$. Evidently, $0 \leqslant \iota(G) \leqslant \beta_1(G)$. $\iota(G) = 0$ corresponds to thermodynamic equilibrium, and it can be readily shown that $\iota(G) = 1$ corresponds to a single energetic edge (see Appendixes). It is $\iota(G)$ that determines the extent of departure from thermodynamic equilibrium and, as we will see, the algebraic complexity of nonequilibrium expressions. It thereby offers a general way to partition the nonequilibrium landscape without choosing beforehand a state of thermodynamic equilibrium.

D. Generalized Hopfield discrimination

Hopfield's seminal analysis of discrimination between a correct and incorrect substrate sought to explain the low error rates in RNA and DNA synthesis [19]. It has been widely influential both theoretically and experimentally [28–31]. Here we analyze a general mechanism of Hopfield discrimination using the linear framework approach of [16]. Figure 2(a) shows a butterfly graph, $G = \overline{C} \bowtie C$, consisting of two "wings," $\overline{C}$ and C, sharing a common reference vertex 1 (open disk). If C and $\overline{C}$ are strongly connected, so too is G, and [16]

$$\rho_i(G) = \begin{cases} \rho_i(\overline{C})\rho_1(C) & \text{if } i \in \overline{C} \\ \rho_1(\overline{C})\rho_i(C) & \text{if } i \in C. \end{cases} \tag{13}$$

For Hopfield discrimination, C represents the mesostates interacting with the correct substrate and $\overline{C}$ the same for the incorrect substrate. Structurally (i.e., ignoring labels), C and $\overline{C}$ are mirror images of each other. Using overlines to map graph entities in C to their mirror images in $\overline{C}$, $i \to j$ if, and only if, $\overline{i} \to \overline{j}$. C is assumed to be reversible and strongly connected but otherwise arbitrary. Ligand binding to vertex 1 leads to K *proximal vertices*, $p_1, \ldots, p_K$ and ligand is selected to be correct at a distinguished *exit vertex*, $e \neq 1$. The *error ratio* is

$$\varepsilon = \frac{u_{\overline{e}}^*(G)}{u_e^*(G)} = \frac{\rho_{\overline{e}}(\overline{C})\rho_1(C)}{\rho_1(\overline{C})\rho_e(C)}, \tag{14}$$

where the second equality comes from Eqs. (4) and (13). Assume to begin with that G is at thermodynamic equilibrium with labels $\ell_{eq}(i \to j)$. Following Hopfield, discrimination only takes place through unbinding from proximal vertices. Accordingly, $\ell_{eq}(i \to j) = \ell_{eq}(\overline{i} \to \overline{j})$ as long as $j \neq 1$, and $\ell_{eq}(\overline{p}_u \to 1) = \alpha \ell_{eq}(p_u \to 1)$, where $\alpha > 1$, so that the incorrect substrate has a higher off rate. Using Eq. (6), it follows that the equilibrium error ratio is $\varepsilon_{eq} = \alpha^{-1}$. Accordingly, $\varepsilon_{eq} < 1$.

Hopfield's insight was that ε_{eq} is independent of the number of discriminations K, and the only way to exceed this "Hopfield barrier" [15] is to expend energy. He analyzed the graph in Fig. 2(b), for which $K = 2$ and C has only three vertices, with energy expenditure on the thick edge, and identified a parametric region for kinetic proofreading in which $(\varepsilon_{eq})^2 < \varepsilon < \varepsilon_{eq}$ [16,19]. Here we ask what determines ε when $K > 2$ and C is a general graph [Fig. 2(a)] in which energy is expended at only a single energetic edge where $\ell(z_1 \to z_2) = m\ell_{eq}(z_1 \to z_2)$. As noted above, this is equivalent to assuming that $\iota(C) = 1$.

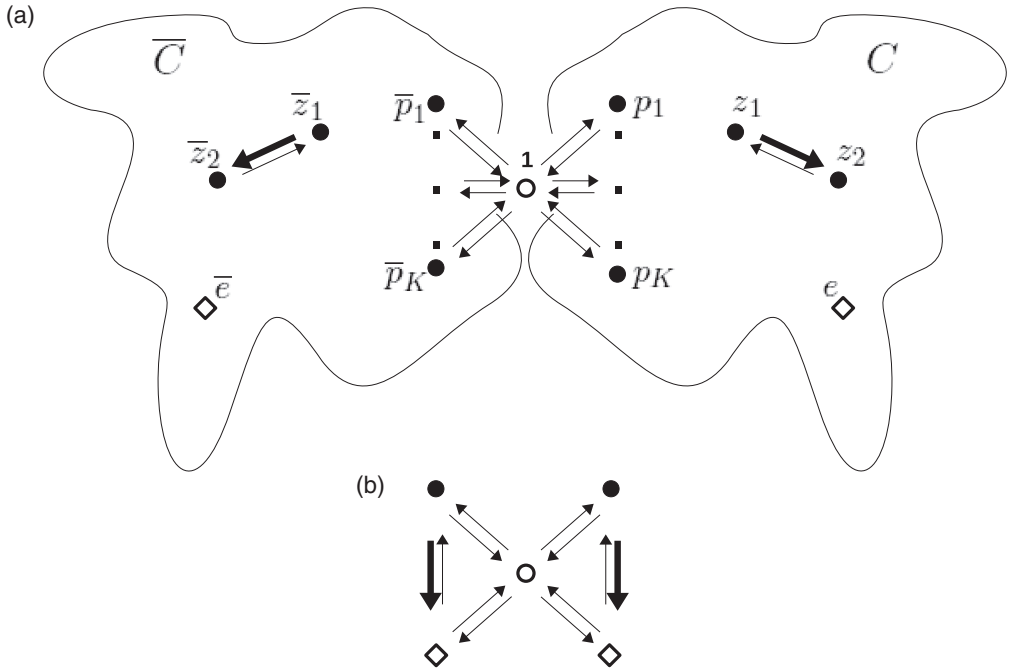

FIG. 2. Hopfield discrimination. (a) Schematic butterfly graph, $G = \overline{C} \bowtie C$, for generalized Hopfield discrimination [16]. The subgraphs C and $\overline{C}$ for discriminating the correct and incorrect substrate, respectively, are structurally mirror images with a shared reference vertex (open disk). C is essentially arbitrary (cloud outline—see the text), with K proximal vertices, a single energetic edge (thick edge), and an exit vertex (open diamond). Only the graph structure is shown, with labels omitted for clarity. (b) The butterfly graph structure for Hopfield's original analysis [19].

We calculate how ε depends on m and α by exploiting the reformulation above and partitioning trees according to proximal edges (i.e., edges $j \to 1$, where j is a proximal vertex). See the Appendixes for full details.

A polynomial $\mathcal{P}$ in the edge labels is m-free, respectively α-free, if m, respectively α, does not divide any monomial in $\mathcal{P}$. Let $\Theta = \Theta_1(G)$. We partition Θ according to the combinatorics of the energetic edge on minimal paths in $M(e, 1)$:

$$\begin{aligned} \Theta_0 &= \{T \in \Theta \,|\, z_1 \to z_2, z_2 \to z_1 \notin T_e\} \\ \Theta_+ &= \{T \in \Theta \,|\, z_2 \to z_1 \in T_e\} \\ \Theta_- &= \{T \in \Theta \,|\, z_1 \to z_2 \in T_e\}. \end{aligned} \tag{15}$$

Let $P_0, P_+, P_- \in M(e, 1)$ be any choice of minimal paths arising as T_e for $T \in \Theta_0, \Theta_+, \Theta_-$, respectively. By construction, different choices are energetically similar, so that $S(P_0), S(P_+), S(P_-)$ are well defined for any choices of minimal paths. Furthermore, by Eq. (12), $S(P_+) = S(P_0) - \ln(m)$, $S(P_-) = S(P_0) + \ln(m)$. Let us extend q to subsets $X \subseteq \Theta$ by defining $q(X) = \sum_{T \in X} q(T)$ and note that q_{eq} corresponds to $m = 1$. We see from Eq. (15) that $q(\Theta_+)$ is m-free and that $q(\Theta_-) = m\, q_{eq}(\Theta_-)$. Hence, we can rewrite Eq. (7) as

$$\rho_e(C) = \exp(-S(P_0))(q(\Theta_0) + m\, q(\Theta_+) + q_{eq}(\Theta_-)), \tag{16}$$

where $q(\Theta_0)$ and $q_{eq}(\Theta_-)$ are m-free but $q(\Theta_0)$ may not be. The limitation of Eq. (16) to just three parts reflects the fact that $\iota(C) = 1$, or, equivalently, the presence of just a single energetic edge.

We now introduce the partitioning scheme. Given $X \subseteq \Theta$, let $X^{(j,u)} \subseteq X$, for $1 \leqslant j \leqslant K$ and $0 \leqslant u \leqslant 1$, consisting of those trees in X with exactly j proximal edges to 1 and exactly u energetic edges. The $X^{(j,u)}$ form a partition of X into mutually disjoint subsets, so that $q(X) = \sum_{j,u} q(X^{(j,u)})$. By construction, $q(X^{(j,0)})$ is m-free and $q(X^{(j,1)}) = m\, q_{eq}(X^{(j,1)})$, where $q_{eq}(X^{(j,1)})$ is m-free. Most importantly, again by construction,

$$q(\overline{X}^{(j,u)}) = \alpha^j q(X^{(j,u)}), \tag{17}$$

where, evidently, $q(X^{(j,u)})$ is α-free.

Equations (16) and (17) make it straightforward to calculate the error ratio from Eq. (14),

$$\varepsilon = \varepsilon_{eq}\left(\frac{(\overline{P}m + \overline{Q})(Rm + S)}{(\overline{R}m + \overline{S})(Pm + Q)}\right), \tag{18}$$

where the eight coefficients in Eq. (18), which are all m-free, are expressed in terms of the constructions above in Eq. (A11). It is striking that Eq. (18) has the same algebraic form for the general graph in Fig. 2(a) as for Hopfield's simple graph in Fig. 2(b) [16, Eq. (4)], albeit with vastly more complicated coefficients. The overlined coefficients in Eq. (18) are each of degree K in $\alpha = \varepsilon_{eq}^{-1}$, so that ε is a rational function of ε_{eq} and m. Equation (18) enables us to analyze discrimination in complex graphs (below).

The rational dependence of ε on ε_{eq} was previously shown in [32, Eq. (10)]. This study was based on a network similar to the graph in Fig. 2(a), with certain structural restrictions—the same number of edges leave the reference vertex as enter the exit vertex—but allowed for discriminations at nonproximal edges and global energy expenditure [32, Fig. 3]. The latter assumption amounts to taking the entropy production index to be maximal: if G is the graph, then $\iota(G) = \beta_1(G)$. No

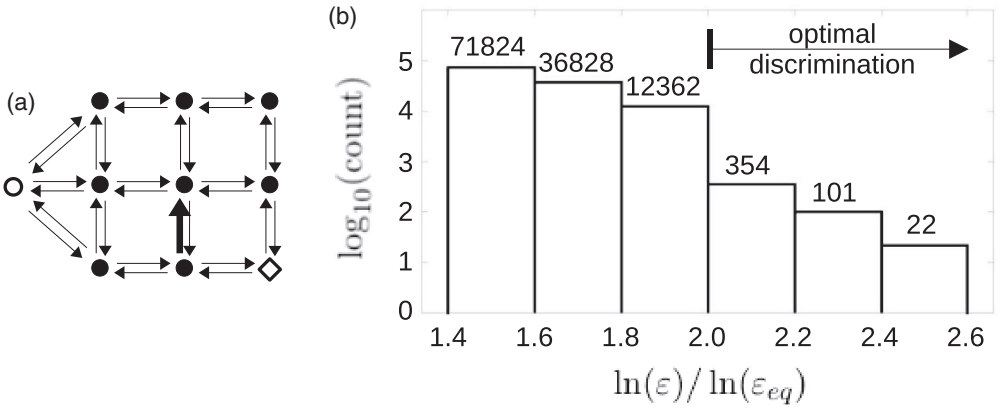

FIG. 3. Optimal Hopfield discrimination in a complex graph. (a) Right-hand (correct) wing of butterfly graph structure, with $K = 3$ proximal vertices, following the same conventions as Fig. 2, with a single energetic edge (thick edge) separated from the proximal vertices and the exit vertex (open diamond). (b) Histogram of numerical calculations for the graph in panel (a) giving logarithmic counts of randomly sampled parameter sets, with actual numbers over each bar, for specified ranges of normalized logarithmic error ratios. Those parameter sets achieving optimal discrimination, with $(\varepsilon_{eq})^3 < \varepsilon < (\varepsilon_{eq})^2$, are indicated. The calculations used Eq. (18).

formula like Eq. (18) was found but the following bounds were observed:

$$(\varepsilon_{eq})^K < \varepsilon < (\varepsilon_{eq})^{2-K}, \tag{19}$$

where the quantity corresponding to our K is the number of "discriminatory edges" [32, C(i)]. Equation (19) is easy to deduce from Eq. (18) [see Eq. (A14), below]. The left-hand inequality in Eq. (19) generalizes Hopfield's finding for Fig. 2(b) with $K = 2$. Equation (19) allows for "antiproofreading" regimes, where energy expenditure worsens the error ratio above the equilibrium value [32]. Indeed, this is already seen in the quadratic dependency of ε on m in Eq. (18) [16].

E. Other previous work and optimal discrimination

Early studies [33–36] rigorously analyzed Hopfield's remarks on multistage proofreading schemes [19]. They vertically extended the graph in Fig. 2(b) to have multiple "stages" and established conditions for minimizing energy expenditure for a given error ratio. More recently, Refs. [32,37] identified multiple proofreading regimes in general networks similar to Fig. 2(a) and found that optimal discrimination, where $(\varepsilon_{eq})^K < \varepsilon < (\varepsilon_{eq})^{K-1}$, can be achieved [37, Eq. (3)]. An important distinction here is that [32,37] allow energy expenditure throughout the network, corresponding to $\iota(G) = \beta_1(G)$, while the earlier studies allow it at only one edge, corresponding to $\iota(G) = 1$, as we have done above. We were therefore interested to find optimal discrimination in a complex graph [Fig. 3(a)] with $\iota(G) = 1$ (see Appendixes). Note that this graph does not meet the structural restrictions imposed in [32,37].

Optimal discrimination requires that each proximal edge contributes synergistically to the error reduction. This synergy presumably arises from the global dependence of steady-state probabilities on all edge labels when energy is being expended. What is surprising, however, is that $\iota(G) = 1$, and energy is being expended at only a single edge, which is distant from the proximal edges where discrimination occurs. For the graph in Fig. 3(a), it is not difficult to see that the first Betti number is $\beta_1(G) = 6$, so that, in an appropriate cycle basis, five out of six cycles still satisfy the cycle condition. The edge labels on these five cycles are therefore constrained, which reduces the parametric choices. This would not be the case if energy was being expended throughout the graph, so that $\iota(G) = \beta_1(G)$. Nevertheless, Fig. 3(b) shows that it is still possible to achieve optimal discrimination. We see that away from thermodynamic equilibrium, global functional capabilities can arise through only local energy expenditure. Further examination suggests that the position of the energetic edge within the graph has an impact on the extent of discrimination. The effect is subtle, and a complete analysis of this intriguing question remains a work in progress.

IV. DISCUSSION

The matrix-tree theorem gives an exact solution for the s.s. probabilities of a Markov process [Eqs. (3) and (4)], but its elegant statement belies its intractability. Away from thermodynamic equilibrium, the s.s. probability of a vertex becomes globally dependent on all graph edge labels and requires enumerating all spanning trees, leading to a combinatorial explosion. Studies have had to rely, in effect, on astute approximations to a few dominant trees. While this has provided important insights, it also suggests that those behaviors which depend on small contributions from many trees may have been overlooked. This could be a particularly serious omission in biology, where functionality can arise from many small contributions [38].

We have overcome the intractability of s.s. probabilities in two ways. First, by recasting the MTT as a smooth generalization of the equilibrium case, in place of minimal path entropies at equilibrium [Eq. (6), averages of these quantities must be taken away from equilibrium Eq. (8)]. Here the average is calculated over the arboreal distribution on spanning trees [Eq. (10)]. Second, energetically similar minimal paths have identical entropies so that the number of distinct minimal path entropies scales with the number of energetic edges, not the size of the graph. It is this scaling which has enabled exact calculation of the error ratio [Eq. (18)] for an arbitrary graph with only one energetic edge [Fig. 2(a)]. Equivalently, systems

away from equilibrium become amenable to calculation if their entropy production index is 1, despite suffering all the problems of global parameter dependence and combinatorial explosion. Such calculations would not have been feasible with the unreformulated MTT [32].

Similar expressions to Eq. (8) have been previously described in [39, Eq. (15)] and [40, Eq. (3.12)] but these lack the arboreal distribution on spanning trees. This distribution is important because it gives an exact description and is Boltzmann like [Eq. (10)]. It has also been studied mathematically, although not named, as the fixed point of algorithms for generating random spanning trees [41–43].

Landauer pointed out the necessity for kinetic, nonthermodynamic quantities to exactly describe the global nature of nonequilibrium steady states [44], despite attempts to characterize them thermodynamically [45,46]. Landauer's point has been reiterated in Maes' concept of "frenesy" [47]. Equation (8) cleanly separates the local, thermodynamic contribution to steady-state probabilities, $\exp(-S(T))$, from the kinetic, global contribution coming from the arboreal distribution, $\Pr_{\Theta_1(G)}$. An important task for subsequent work is to characterize the arboreal distribution, for which the mathematical connections mentioned above may be helpful.

The reformulation in Eq. (8) has a formal resemblance to other path ensemble formulations in physics [3,48,49]. Interestingly, it is nonequilibrium information which is recovered here from a Boltzmann-like distribution. Whether some more fundamental setting underlies these different formulations is beyond the scope of this paper, but the graph-theoretic framework used here is essential to describing the ensemble.

Our reformulation further clarifies Hopfield's analysis of kinetic proofreading [19]. At equilibrium, path independence [Eq. (6)] implies that ε_{eq} is independent of the number of discriminations. Away from equilibrium, global path dependency allows multiple discriminations to exploit energy expenditure at a single distant edge, to collectively achieve optimal discrimination (Fig. 3). Being "away from thermodynamic equilibrium" is not a unitary condition but rather a nuanced and complex landscape, in which how and where energy is expended can profoundly influence functional outcomes. The ideas introduced here suggest how we can begin to "follow the energy" through this nonequilibrium landscape to uncover the logic of biological information processing.

ACKNOWLEDGMENTS

We thank two anonymous reviewers for their suggestions, which improved the manuscript, and Christian Maes and Kee-Myoung Nam for helpful comments. U.C. was supported by the Giovanni Armenise Harvard Foundation. U.C. and J.G. were also supported by U.S. NSF Award No. 1462629 and by U.S. NIH Award No. R01GM105375.

APPENDIX

We provide further details here of the analysis of generalized Hopfield discrimination, following on from the main text and the notation introduced there.

1. Partitioning scheme

Recall the partitioning scheme introduced in the main text in which a subset of spanning trees rooted at the reference vertex, $X \subseteq \Theta$, is divided into parts, $X^{(j,u)}$, for $1 \leqslant j \leqslant K$ and $0 \leqslant u \leqslant 1$. $X^{(j,u)}$ consists of those trees in X with exactly j proximal edges to 1 and exactly u energetic edges. In other words, $T \in X^{(j,u)}$ if, and only if, there are distinct proximal vertices, $p_{i_1}, \ldots, p_{i_j}$ in C, which may depend on T, such that $p_{i_1} \to 1, \ldots, p_{i_j} \to 1 \in T$, but no other proximal vertices have this property; and, if $u = 0$, $z_1 \to z_2 \notin T$, while if $u = 1$, $z_1 \to z_2 \in T$. Note that any tree in Θ must have between 1 and K proximal edges to 1 and no more than 1 energetic edge [Fig. 2(a)]. It follows that the $X^{(j,u)}$ form a partition of X into disjoint subsets,

$$X = (X^{(1,0)} \sqcup X^{(1,1)}) \sqcup \cdots \sqcup (X^{(K,0)} \sqcup X^{(K,1)}). \tag{A1}$$

Here $A \sqcup B$ denotes $A \cup B$ with the added stipulation that $A \cap B = \emptyset$. Equation (A1) implies that we can decompose $q(X)$,

$$\begin{aligned} q(X) = {} & [q(X^{(1,0)}) + \cdots + q(X^{(K,0)})] \\ & + m[q_{eq}(X^{(1,1)}) + \cdots + q_{eq}(X^{(K,1)})], \end{aligned} \tag{A2}$$

where the terms $q(X^{(j,0)})$ and $q_{eq}(X^{(j,1)})$ are all m-free. The value of this decomposition becomes clear by applying Eq. (17) to see that

$$\begin{aligned} q(\overline{X}) = {} & [\alpha q(X^{(1,0)}) + \cdots + \alpha^K q(X^{(K,0)})] \\ & + m[\alpha q_{eq}(X^{(1,1)}) + \cdots + \alpha^K q_{eq}(X^{(K,1)})], \end{aligned} \tag{A3}$$

where the expressions $q(X^{(j,0)})$ and $q_{eq}(X^{(j,1)})$ are both m-free, as above, and, by construction, α-free.

2. Calculation of $\rho_1(C)$ and $\rho_1(\overline{C})$

Let us abbreviate $\sum_{1 \leqslant j \leqslant K}$ by $\sum_j$. Since $\rho_1(C) = q(\Theta)$, it follows from Eq. (A2) that

$$\rho_1(C) = \left(\sum_j q(\Theta^{(j,0)}) \right) + m \left(\sum_j q_{eq}(\Theta^{(j,1)}) \right). \tag{A4}$$

Similarly, applying Eq. (A3) to $\rho_1(\overline{C}) = q(\overline{\Theta})$, we see that

$$\rho_1(\overline{C}) = \left(\sum_j \alpha^j q(\Theta^{(j,0)}) \right) + m \left(\sum_j \alpha^j q_{eq}(\Theta^{(j,1)}) \right). \tag{A5}$$

3. Calculation of $\rho_e(C)$ and $\rho_{\overline{e}}(\overline{C})$

We begin with Eq. (16), which represented $\rho_e(C)$ in terms of a partition into disjoint subsets of Θ,

$$\Theta = \Theta_0 \sqcup \Theta_+ \sqcup \Theta_-. \tag{A6}$$

It is clear from Eq. (16) that, for all $1 \leqslant j \leqslant K$,

$$\Theta_+^{(j,1)} = \emptyset \text{ and } \Theta_-^{(j,0)} = \emptyset. \tag{A7}$$

If we now apply the decomposition in Eq. (A2) to each of the three subsets of Eq. (A6) and collect together the m-free terms

in Eq. (16), we find that

$$\rho_e(C) = \exp\left(-S(P_0)\right)\Bigg\{\sum_j \left(q\left(\Theta_0^{(j,0)}\right) + q_{eq}\left(\Theta_-^{(j,1)}\right)\right) + m\Bigg[\sum_j \left(q_{eq}\left(\Theta_0^{(j,1)}\right) + q\left(\Theta_+^{(j,0)}\right)\right)\Bigg]\Bigg\}. \tag{A8}$$

Similarly, using the decomposition in Eq. (A3), we find that

$$\rho_{\overline{e}}(\overline{C}) = \alpha^{-1}\exp\left(-S(P_0)\right)\Bigg\{\sum_j \alpha^j\left(q\left(\Theta_0^{(j,0)}\right) + q_{eq}\left(\Theta_-^{(j,1)}\right)\right) + m\Bigg[\sum_j \alpha^j\left(q_{eq}\left(\Theta_0^{(j,1)}\right) + q\left(\Theta_+^{(j,0)}\right)\right)\Bigg]\Bigg\}, \tag{A9}$$

where we have used the fact that $S(\overline{P_0}) = \alpha^{-1}S(P_0)$.

4. Proof of Eq. (18)

We have now calculated each of the four terms appearing in Eq. (18). Using overlines again to indicate coefficients from the mirror image subgraph $\overline{C}$ and recalling that $\alpha = (\varepsilon_{eq})^{-1}$, we find that

$$\varepsilon = \varepsilon_{eq}\left(\frac{(\overline{P}m + \overline{Q})(Rm + S)}{(\overline{R}m + \overline{S})(Pm + Q)}\right). \tag{A10}$$

The eight coefficients in Eq. (A10) are given by

$$\begin{aligned}
P &= \textstyle\sum_j q_{eq}\left(\Theta_0^{(j,1)}\right) + q\left(\Theta_+^{(j,0)}\right)\\
Q &= \textstyle\sum_j q\left(\Theta_0^{(j,0)}\right) + q_{eq}\left(\Theta_-^{(j,1)}\right)\\
R &= \textstyle\sum_j q_{eq}(\Theta^{(j,1)})\\
S &= \textstyle\sum_j q(\Theta^{(j,0)})\\
\overline{P} &= \textstyle\sum_j \alpha^j\left[q_{eq}\left(\Theta_0^{(j,1)}\right) + q\left(\Theta_+^{(j,0)}\right)\right]\\
\overline{Q} &= \textstyle\sum_j \alpha^j\left[q\left(\Theta_0^{(j,0)}\right) + q_{eq}\left(\Theta_-^{(j,1)}\right)\right]\\
\overline{R} &= \textstyle\sum_j \alpha^j q_{eq}(\Theta^{(j,1)})\\
\overline{S} &= \textstyle\sum_j \alpha^j q(\Theta^{(j,0)}).
\end{aligned} \tag{A11}$$

The various expressions of the form $q(-)$ or $q_{eq}(-)$ in Eq. (A11) are all both α-free and m-free. Equations (A10) and (A11) establish Eq. (18).

5. Proof of Eq. (19)

Using the fact that $\alpha > 1$, which encodes the distinction between the correct and incorrect substrate, we see by inspection of Eq. (A11) that

$$\begin{aligned}
\alpha(P+Q) &< \overline{P}+\overline{Q} < \alpha^K(P+Q)\\
\alpha(R+S) &< \overline{R}+\overline{S} < \alpha^K(R+S).
\end{aligned}$$

It follows that

$$\alpha < \frac{\overline{P}+\overline{Q}}{P+Q} < \alpha^K, \tag{A12}$$

and also that

$$\alpha^{-K} < \frac{R+S}{\overline{R}+\overline{S}} < \alpha^{-1}. \tag{A13}$$

Combining Eqs. (A12) and (A13), and using Eq. (A10) and the fact that $\alpha = (\varepsilon_{eq})^{-1}$, we see that

$$(\varepsilon_{eq})^K < \varepsilon < (\varepsilon_{eq})^{2-K}, \tag{A14}$$

which proves Eq. (19).

6. Figure 3(b)

Equation (18) expresses ε as a quadratic rational function of m. This graph of this function can exhibit many shapes [16], and we looked for the conditions under which it has a positive minimum, which implies proofreading. The derivative, $d\varepsilon/dm$, has a quadratic numerator, and it is readily seen that ε is decreasing at $m = 0$ if

$$\frac{S\overline{P}+R\overline{Q}}{P\overline{S}+Q\overline{R}} < \frac{S\overline{Q}}{Q\overline{S}}. \tag{A15}$$

When Eq. (A15) holds, ε has a single positive minimum if also

$$\frac{S\overline{P}+R\overline{Q}}{P\overline{S}+Q\overline{R}} < \frac{R\overline{P}}{P\overline{R}}. \tag{A16}$$

We sought this minimum error ratio for the graph in Fig. 3(a). First, an independent set of edge labels was chosen using a spanning tree, as previously described [15]. Numerical values for these rates on the correct wing of the graph, C, were chosen independently as 10^x, where x was drawn randomly from the uniform distribution on $[-3, +3]$. The remaining edge labels, not on the spanning tree, which determine independent cycles, were chosen to make $\iota(G) = 0$, so that G is at thermodynamic equilibrium. The labels on the incorrect wing $\overline{C}$ were then determined by the relationships described in the main text. The value of α, which gives the equilibrium error ratio, $\varepsilon_{eq} = \alpha^{-1}$, was arbitrarily set to 0.1. Departure from thermodynamic equilibrium was imposed through the multiplier m on the energetic edge $\ell(z_1 \to z_2) = m\ell_{eq}(z_1 \to z_2)$, while all other edge labels kept their previously assigned values, $\ell(i \to j) = \ell_{eq}(i \to j)$ when $i \neq z_1$ or $j \neq z_2$. Spanning trees were enumerated using MATLAB'S `generateSpanningTrees` function, and the eight coefficients, $P, Q, R, S, \overline{P}, \overline{Q}, \overline{R}, \overline{S}$, in Eq. (18) were numerically calculated from Eq. (A11). If they did not satisfy the inequalities in Eqs. (A15) and (A16), the parameter set was rejected. If they did, the value of m giving the positive minimum of ε was calculated from the quadratic numerator of $d\varepsilon/dm$ as

$$m_* = \frac{P\overline{Q}S\overline{R} - Q\overline{P}R\overline{S} + \sqrt{(P\overline{Q}-Q\overline{P})(SP-QR)(S\overline{P}-Q\overline{R})(R\overline{S}-S\overline{R})}}{R\overline{P}(P\overline{S}+Q\overline{R}) - P\overline{R}(S\overline{P}+R\overline{Q})},$$

and the corresponding minimum value of ε was determined by substituting m_* into Eq. (18). $\sim 10^7$ parameter sets were sampled, of which $\sim 1.2 \times 10^5$ had minimum error ratios satisfying $\ln(\varepsilon)/\ln(\varepsilon_{eq}) > 1.4$, as reported in Fig. 3(b).

7. Cycle basis and $\iota(G) = 1$

Assuming that G is reversible, it is simpler to work with the corresponding undirected and unlabeled graph G^u, in which $i \sim j$ in G^u if, and only if, $i \leftrightarrows j$ in G. An undirected spanning tree T in G^u is a connected, acyclic subgraph that includes each vertex. Choosing any edge which is not in T defines a minimal cycle in G^u. Such a cycle can be lifted to a minimal reversible cycle in G, and an arbitrary orientation around the cycle may be chosen. Doing this for each minimal cycle obtained from T defines a basis of oriented minimal cycles. If G has v vertices and e reversible pairs of edges, then T has $v - 1$ edges and the number of minimal cycles in a basis is the first Betti number of G, $\beta_1(G) = e - v + 1$. If $\iota(G) = 1$, then there is a basis of oriented cycles in which only one cycle breaks the cycle condition. Take the undirected edge on the cycle which is not on the corresponding spanning tree in G^u. By changing the label of either of the corresponding reversible edges in G, the oriented cycle can always be returned to equilibrium. Conversely, suppose G is moved from thermodynamic equilibrium by changing just $\ell(i \to j)$. Let u denote the undirected edge $i \sim j$ in G^u. Take any basis B of oriented minimal cycles in G and let T be the corresponding undirected spanning tree in G^u. If $\iota(B) > 1$, so that multiple cycles in B break the cycle condition, then u must be an edge in T, for otherwise it would occur on only one cycle. Choose any cycle on which u appears and suppose that v is the corresponding edge which is not in T. Removing u from T and adjoining v creates a new spanning tree T' in G^u. Because u is not in T' by construction, the basis B' of minimal cycles corresponding to T' has only the cycle defined by u away from equilibrium. Hence, $\iota(B') = 1$ and so $\iota(G) = 1$. It follows that having a single edge away from equilibrium corresponds exactly to $\iota(G) = 1$.

[1] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, Probability of Second Law Violations in Shearing Steady States, Phys. Rev. Lett. **71**, 2401 (1993).

[2] G. Gallavotti and E. G. D. Cohen, Dynamical Ensembles in Nonequilibrium Statistical Mechanics, Phys. Rev. Lett. **74**, 2694 (1995).

[3] C. Jarzynski, Equilibrium free-energy differences from nonequilibrium measurements: A master equation approach, Phys. Rev. E **56**, 5018 (1997).

[4] G. E. Crooks, Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences, Phys. Rev. E **60**, 2721 (1999).

[5] J. Liphardt Jr., S. Dumont, S. B. Smith, I. Tinoco Jr., and C. Bustamante, Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality, Science **296**, 1832 (2002).

[6] S. Raman, T. Utzig, T. Baimpos, B. R. Shrestha, and M. Valtiner, Deciphering the scaling of single-molecule interactions using Jarzynski's equality, Nat. Commun. **5**, 5539 (2014).

[7] J. Camunas-Soler, A. Alemany, and F. Ritort, Experimental measurement of binding energy, selectivity, and allostery using fluctuation theorems, Science **355**, 412 (2017).

[8] U. Çetiner, O. Raz, S. Sukharev, and C. Jarzynski, Recovery of Equilibrium Free Energy from Nonequilibrium Thermodynamics with Mechanosensitive Ion Channels in *E. coli*, Phys. Rev. Lett. **124**, 228101 (2020).

[9] U. Seifert, Stochastic thermodynamics: Principles and perspectives, Eur. Phys. J. B **64**, 423 (2008).

[10] J. Gunawardena, A linear framework for time-scale separation in nonlinear biochemical systems, PLoS One **7**, e36321 (2012).

[11] K.-M. Nam, R. Martinez-Corral, and J. Gunawardena, The linear framework: Using graph theory to reveal the algebra and thermodynamics of biomolecular systems, Interface Focus **12**, 20220013 (2022).

[12] T. L. Hill, Studies in irreversible thermodynamics, IV. Diagrammatic representation of steady state fluxes for unimolecular systems, J. Theor. Biol. **10**, 442 (1966).

[13] T. L. Hill, *Free Energy Transduction and Biochemical Cycle Kinetics* (Dover Publications, New York, 2004), originally published in 1977.

[14] J. Schnakenberg, Network theory of microscopic and macroscopic behaviour of master equation systems, Rev. Mod. Phys. **48**, 571 (1976).

[15] J. Estrada, F. Wong, A. DePace, and J. Gunawardena, Information integration and energy expenditure in gene regulation, Cell **166**, 234 (2016).

[16] F. Wong, A. Amir, and J. Gunawardena, Energy-speed-accuracy relation in complex networks for biological discrimination, Phys. Rev. E **98**, 012420 (2018).

[17] F. Wong, A. Dutta, D. Chowdhury, and J. Gunawardena, Structural conditions on complex networks for the Michaelis-Menten input-output response, Proc. Natl. Acad. Sci. USA **115**, 9738 (2018).

[18] J. W. Biddle, R. Martinez-Corral, F. Wong, and J. Gunawardena, Allosteric conformational ensembles have unlimited capacity for integrating information, eLife **10**, e65498 (2021).

[19] J. J. Hopfield, Kinetic proofreading: A new mechanism for reducing errors in biosynthetic processes requiring high specificity, Proc. Natl. Acad. Sci. USA **71**, 4135 (1974).

[20] J. W. Biddle, M. Nguyen, and J. Gunawardena, Negative reciprocity, not ordered assembly, underlies the interaction of Sox2 and Oct4 on DNA, eLife **8**, e41017 (2019).

[21] I. Mirzaev and J. Gunawardena, Laplacian dynamics on general graphs, Bull. Math. Biol. **75**, 2118 (2013).

[22] F. Wong and J. Gunawardena, Gene regulation in and out of equilibrium, Annu. Rev. Biophys. **49**, 199 (2020).

[23] R. D. Astumian, C. Pezzato, Y. Feng, Y. Qiu, P. R. McGonigal, C. Cheng, and J. F. Stoddart, Non-equilibrium kinetics and trajectory thermodynamics of synthetic molecular pumps, Mater. Chem. Front. **4**, 1304 (2020).

[24] U. Seifert, Entropy Production along a Stochastic Trajectory and an Integral Fluctuation Theorem, Phys. Rev. Lett. **95**, 040602 (2005).

[25] M. Bauer and F. Cornu, Local detailed balance: A microscopic derivation, J. Phys. A: Math. Theor. **48**, 015008 (2015).
[26] C. Maes and K. Netočný, Heat bounds and the blowtorch theorem, Ann. Henri Poincaré **14**, 1193 (2013).
[27] J. W. Biddle and J. Gunawardena, Reversal symmetries for cyclic paths away from thermodynamic equilibrium, Phys. Rev. E **101**, 062125 (2020).
[28] D. Hartich, A. C. Barato, and U. Seifert, Nonequilibrium sensing and its analogy to kinetic proofreading, New J. Phys. **17**, 055026 (2015).
[29] S. A. Cepeda-Humerez, G. Rieckh, and G. Tkačik, Stochastic Proofreading Mechanism Alleviates Crosstalk in Transcriptional Regulation, Phys. Rev. Lett. **115**, 248101 (2015).
[30] J. Pettmann, A. Huhn, M. A. Kutuzov E. A. Shah, D. B. Wilson, M. L. Dustin, S. J. Davis, P. A. van der Merwe, and O. Dushek, The discriminatory power of the T cell receptor, eLife **10**, e67092 (2021).
[31] H. Boeger, Kinetic proofreading, Annu. Rev. Biochem. **91**, 423 (2022).
[32] A. Murugan, D. A. Huse, and S. Leibler, Discriminatory Proofreading Regimes in Nonequilibrium Systems, Phys. Rev. X **4**, 021016 (2014).
[33] M. Ehrenberg and C. Blomberg, Thermodynamic constraints on kinetic proofreading in biosynthetic pathways, Biophys. J. **31**, 333 (1980).
[34] C. Blomberg and M. Ehrenberg, Energy considerations for kinetic proofreading in biosynthesis, J. Theor. Biol. **88**, 631 (1981).
[35] R. R. Freter and M. A. Savageau, Proofreading systems of multiple stages for improved accuracy of biological discrimination, J. Theor. Biol. **85**, 99 (1980).
[36] M. A. Savageau and D. S. Lapointe, Optimization of kinetic proofreading: A general method for derivation of the constraint relations and an exploration of a specific case, J. Theor. Biol. **93**, 157 (1981).
[37] A. Murugan, D. A. Huse, and S. Leibler, Speed, dissipation, and error in kinetic proofreading, Proc. Natl. Acad. Sci. USA **109**, 12034 (2012).
[38] J. Crocker, N. Abe, L. Rinaldi, A. P. McGregor, N. Frankel, S. Wang, A. Alsawadi, P. Valenti, S. Plaza, F. Payre, R. S. Mann, and D. L. Stern, Low affinity binding site clusters confer Hox specificity and regulatory robustness, Cell **160**, 191 (2015).
[39] T. S. Komatsu and N. Nakagawa, Expression for the Stationary Distribution in Nonequilibrium Steady States, Phys. Rev. Lett. **100**, 030601 (2008).
[40] C. Maes, K. Netočný, and B. Wynants, On and beyond entropy production: The case of Markov jump processes, Markov Processes Relat. Fields **14**, 445 (2008).
[41] A. Broder, Generating random spanning trees, in *Proceedings of the 30th IEEE Symposium on Foundations of Computer Science* (IEEE, New York, 1989), pp. 442–447.
[42] D. Aldous, The random walk construction of uniform spanning trees and uniform labelled trees, SIAM J. Discrete Math. **3**, 450 (1990).
[43] D. B. Wilson, Generating random spanning trees more quickly than the cover time, In *Proceedings 28th Annual ACM Symposium on Theory of Computing* (ACM, New York, 1996), pp. 296–303.
[44] R. Landauer, Inadequacy of entropy and entropy derivatives in characterizing the steady state, Phys. Rev. A **12**, 636 (1975).
[45] E. T. Jaynes, Information theory and statistical mechanics, Phys. Rev. **106**, 620 (1957).
[46] I. Prigogine, *Introduction to Thermodynamics of Irreversible Processes* (Interscience, New York, 1967).
[47] C. Maes, Frenesy: Time-symmetric dynamical activity in nonequilibria, Phys. Rep. **850**, 1 (2020).
[48] R. P. Feynman, Space-time approach to non-relativistic quantum mechanics, Rev. Mod. Phys. **20**, 367 (1948).
[49] G. E. Crooks, Path-ensemble averages in systems driven far from equilibrium, Phys. Rev. E **61**, 2361 (2000).