Geometric Requirements for Nonequilibrium Discrimination

Gaurav G. Venkataraman,[1,2,*] Eric A. Miska,[1,3,4] and David J. Jordan[1,3,*]

[1] *Wellcome/CRUK Gurdon Institute, University of Cambridge. Tennis Court Rd, Cambridge, CB2 1QN, UK.*

[2] *Division of Medicine, University College London. London WC1E 6BT, UK.*

[3] *Department of Genetics, University of Cambridge. Downing Street, Cambridge CB2 3EH, UK.*

[4] *Wellcome Sanger Institute, Wellcome Genome Campus, Cambridge CB10 1SA, UK.*

We study: necessary conditions for the steady-state occupancies of multistep networks to achieve high sensitivity via thermodynamic drive; the ability of these networks to toggle between outputs by adjusting the amount of drive; the possible implementation of this dissipation-driven output toggling in the biochemical context of RNA granules. In general, sensitivity allows the ratios of non-equilibrium steady states to depart far from their equilibrium limit, known as discrimination. Discrimination is crucial for high fidelity information processing at the molecular scale, where steady-state occupancies correspond to (competing) products of biochemical reactions. We define an analytically tractable measure on network discrimination schemes, termed orthogonality, which measures the extent to which discrimination is local. The central proposition of our paper is that discrimination is fundamentally constrained by orthogonality. We demonstrate that discrimination which amplifies binding energy differences requires low orthogonality, whereas discrimination which amplifies activation energy differences requires high orthogonality. Subject to orthogonality requirements, *both* types of discrimination are maximized by maximizing dissipation. Dissipation itself can drive orthogonality up or down. When increasing thermodynamic drive conflicts with orthogonality requirements, discrimination is non-monotonic. We find that, due to its effect on orthogonality, modulating thermodynamic drive alone can sharply select between products which are favored by different energy types, without network fine-tuning. Biologically, this corresponds to the ability to select between products by driving a single reaction type. We consider this possibility in the context of liquid-liquid phase separated collections of RNA and protein known as granules, which appear to have precisely the structure required to tune orthogonality via adjusting the rate of ATP hydrolysis.

I. INTRODUCTION

Systems driven out of thermodynamic equilibrium are capable of demonstrating extraordinarily varied behaviors. One such behavior is enhanced discrimination: the ratio of nonequilibrium steady-state occupancies can deviate far from what is allowed by their energetic differences alone. The cost of this flexibility is parametric complexity. Nonequilibrium steady-state occupancies will generally depend on the system's full parametric details.

Enhanced discrimination is particularly important in a biological context, where biochemical reactions achieve extraordinary fidelity by discriminating between competing substrates having only marginally different equilibrium energies. For example, equilibrium energy differences between nucleotides competing in a DNA replication reaction predict replication error rates of $\sim 10^{-4}$; but observed rates are $\sim 10^{-9}$. To resolve this discrepancy, Hopfield [13] and Ninio [20] proposed a network which uses non-equilibrium drive to enhance discrimination by amplifying equilibrium energy differences. Such networks have been analyzed extensively, extended [10, 16] and generalized [17].

The Hopfield-Ninio scheme achieves its minimum error rate in the slow, quasi-adiabatic regime. Some years later, Bennett introduced a discrimination scheme which achieves its error minimum in the fast, high-dissipation limit [4, 5]. An essential difference between these schemes was only recently noted: quasi-adiabatic discrimination amplifies *binding energy* differences, whereas high-dissipation discrimination amplifies *activation energy* differences. The two regimes were termed *energetic* and *kinetic* discrimination (respectively), and argued to be alternate in any single reaction step [26].

We study the general requirements for nonequilibrium discrimination from a geometric perspective. To this end, we introduce a measure on non-equilibrium systems, called *orthogonality*. The measure tightly bounds the degree to which a network's discrimination can be accurately represented by the transition rates near (0-2 links away from) the discriminatory nodes. We find that local (high orthogonality) discrimination relies on the existence

* Correspondence: gauravvman@gmail.com or dj333@cam.ac.uk

of many effective pathways directed towards the discriminatory nodes. In contrast, discrimination schemes that depend on global network parameters (low orthogonality) are characterized by having a single dominant path towards the discriminatory nodes.

We show that high orthogonality networks are necessary for kinetic discrimination, whereas low orthogonality networks are necessary for energetic discrimination. Subject to these orthogonality requirements, *both* kinetic and energetic discrimination increase with respect to dissipation. When the orthogonality requirement conflicts with the dissipation requirement - which can happen in either discriminatory regime - the level of discrimination is non-monotonic with respect to dissipation. This phenomena is responsible for the non-monotonicity observed in the original Hopfield-Ninio scheme.

We further show that orthogonality can be modulated by changing the drive of a single reaction type in an otherwise fixed network. Whether dissipation increases or decreases orthogonality depends on whether it is used to drive reactions which increase or decrease the number of effective pathways towards the discriminatory nodes. By shifting orthogonality, nonequilibrium drive can achieve simple, highly tunable selection between products which are favored by different energy types.

Hopfield noted that the reaction topologies necessary for energetic discrimination appear ubiquitously in biology. Our results demonstrate how these molecular discrimination mechanisms can explore alternative product spaces by spending energy to transition to the kinetic regime, without cost to their original fidelity. The degree to which a network can be repurposed in this manner is given by the degree to which its orthogonality can be modulated. We consider this repurposing proposition in the particular case of localized collections of RNA and protein known as *granules* [6], whose features bear all of the hallmarks of dissipation-driven orthogonality tuning.

II. DEFINITION OF ORTHOGONALITY

We consider systems whose dynamics are described by continuous time Markov chains. System states and transitions can be represented as a fully connected directed graph (Figure 1(a)), and have dynamics represented by a matrix differential equation known as the Master equation

$$\frac{dp}{dt} = \mathcal{L}p$$

where $\mathcal{L}$ is a *Laplacian matrix* encoding the transition rates $k_{ij} = (j \to i)$ of the network in its off-diagonal elements, and having diagonal elements set such that all columns sum to zero:

$$\mathcal{L}_{ij} = \begin{cases} k_{ij} & i \neq j,\ k_{ij} \geq 0 \\ -\sum_j k_{ij} & i = j, \end{cases}$$

and p is an n dimensional vector representing the dynamic occupancy of the network states.

We require our systems to be fully-connected[21], which implies that they have a single unique steady state (Appendix A). Let ρ be the vector of *steady states* of the Master equation determined by solving $\mathcal{L}\rho = 0$. The degree to which a network *discriminates* between state i and j is given by the ratio of corresponding steady state elements: ρ_i/ρ_j.

In general, for systems both in and out of equilibrium, ρ_i/ρ_j will be depend on transition rate parameters arbitrarily far from the discriminatory nodes i, j. We aim to determine in which cases the ratio ρ_i/ρ_j is more 'local'.

By considering the geometry of the steady-state solution space, we can compute that (Appendix A):

$$\frac{\rho_i}{\rho_j} = \frac{\|v_i - \text{proj}_{\mathcal{L}^{i,j}}(v_i)\|}{\|v_j - \text{proj}_{\mathcal{L}^{i,j}}(v_j)\|}. \tag{1}$$

where: v_i represents the ith column of $\mathcal{L}$; $\text{proj}_{\mathcal{L}^{i,j}}(v_i)$ denotes the projection of vector v_i onto the $n-2$ dimensional subspace spanned by the columns of $\mathcal{L}$ remove i, j.

In general, computing the subspace projections $\text{proj}_{\mathcal{L}^{i,j}}(v_i)$ will require orthogonalizing the columns of $\mathcal{L}^{i,j}$. From the perspective of Equation 1, this (recursive) orthogonalization procedure is the origin of global parametric complexity in ρ_i/ρ_j.

In the special case where the vectors that span $\mathcal{L}^{i,j}$ are orthogonal, ${\mathcal{L}^{i,j}}_{\text{orth}}$, the Equation 1 projections can be simply computed:

$$\text{proj}_{{\mathcal{L}^{i,j}}_{\text{orth}}}(v_i) = \sum_{l \in \mathcal{L}^{i,j}} \langle v_i, v_l \rangle v_i. \tag{2}$$

Graphically, $\langle v_i, v_l \rangle \neq 0$ only if the ith and lth nodes on the digraph associated with $\mathcal{L}$ are directly linked or direct arrows at one or more mutual nodes. Therefore, ${\mathcal{L}^{i,j}}_{\text{orth}}$ corresponds to the case in which ρ_i/ρ_j depends on only transition rates 0-2 connections away from the nodes i, j. The matrix $\mathcal{L}^{i,j}$ having orthogonal columns implies *local* discrimination between states i and j.

In practice, $\mathcal{L}^{i,j}$ having purely orthogonal columns conflicts with our requirement that $\mathcal{L}$ describe dynamics on a fully connected graph. Distance of an $\mathcal{L}^{i,j}$ matrix from the perfectly orthogonal case is the relevant quantity to consider. To that end, we compute the difference between the true projection of a vector v onto $\mathcal{L}^{i,j}$, $\text{proj}_{\mathcal{L}^{i,j}}(v)$, and the projection assuming that $\mathcal{L}^{i,j}$ is orthogonal, $\text{proj}_{\mathcal{L}^{i,j}{}_{\text{orth}}}(v)$,

$$\begin{aligned} &\left\|\text{proj}_{\mathcal{L}^{i,j}}(v) - \text{proj}_{\mathcal{L}^{i,j}{}_{\text{orth}}}(v)\right\| \\ &= \left\|\mathcal{L}^{i,j}(\mathcal{L}^{i,j^T}\mathcal{L}^{i,j})^{-1}\mathcal{L}^{i,j^T}v - \mathcal{L}^{i,j}\mathcal{L}^{i,j^T}v\right\| \\ &\leq \left\|\mathcal{L}^{i,j}(\mathcal{L}^{i,j^T}\mathcal{L}^{i,j})^{-1}\mathcal{L}^{i,j^T} - \mathcal{L}^{i,j}\mathcal{L}^{i,j^T}\right\| \|v\|, \\ &= \left\|\mathbf{I} - \mathcal{L}^{i,j^T}\mathcal{L}^{i,j}\right\| \|v\| \end{aligned} \tag{3}$$

where the inequality is due to Cauchy-Schwartz and the final line is arrived at by simplifying the term in the first norm (Appendix B).

We define *orthogonality* to capture this distance.

Definition II.1 (Orthogonality)**.** *We define the orthogonality of the discrimination between states i and j of network having Laplacian $\mathcal{L}$ to be*

$$\Theta(\mathcal{L}^{i,j}) = 1 - \left\|\mathbf{I} - \widehat{\mathcal{L}^{i,j}}^T \widehat{\mathcal{L}^{i,j}}\right\|_F$$

where $\|\ \|_F$ denotes the Frobenius norm and $\mathcal{L}^{i,j}$ is a $(n \times n-2)$ matrix constructed by removing the columns corresponding to the i,j discriminatory nodes from $\mathcal{L}$; and $\widehat{A}$ denotes the matrix A having every column normalized to be of unit length. Note that orthogonality is defined with respect to a network and two nodes of discriminatory interest. We will use the term 'discriminatory scheme,' to refer to this (network, nodes) pair and therefore refer to the 'orthogonality of a scheme' leaving the specific discriminatory nodes i,j implicit.

Orthogonality can be understood to capture the number of effective pathways directed towards the discriminatory nodes. To illustrate this, we consider discriminating between the end nodes of a 4 node toy model (Figure 1(a)) having bidirectional connections equal to either k or l (black, red, arrows in 1(a), respectively). As the ratio $r = k/l$ grows, the black path becomes singly dominant over the red paths. Correspondingly, the orthogonality decreases (Figure 1(b)) as $\Theta \sim O(r^{-2})$ (Appendix D).

Alternatively, we can ask: when does deleting the black links between the middle two nodes in Figure 1(a) affect orthogonality? We find that when these arrows do not constitute part of a dominant path ($r \approx 1$), deleting them has no effect on orthogonality. But when these links *do* constitute part of a dominant path ($r >> 1$), deleting them significantly increases orthogonality (Appendix D).

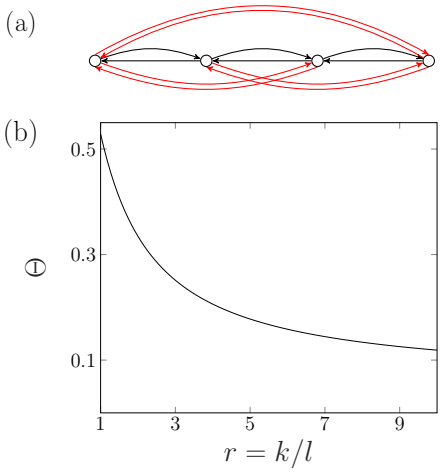


FIG. 1: Orthogonality captures the number of effective pathways directed at discriminatory nodes. (a) A toy four-node discriminatory scheme. (b) When each black and red arrow in panel (a) has weight k, l (respectively) we find that the orthogonality (Θ) decreases as the black arrows dominate, creating a singly dominant path. Note that the x-axis begins at $r = 1$, corresponding to an all-to-all connected graph with equal weights.

Our example demonstrates that an all-to-all connected discrimination scheme ($r = 1$) with equal rates has greater orthogonality than a line topology ($r >> 1$) having equal rates (k), for $N = 4$ nodes. We can compute that this holds for all N (Appendix C). The increased orthogonality of the all-to-all relative to line topology captures a more general fact: orthogonality tends to decrease as connections are removed from a discrimination scheme, so long as these connections are of equal order magnitude to remaining connections. We demonstrate this computationally (Figure S1).

III. ORTHOGONALITY IN THE HOPFIELD-NINIO SCHEME

We first demonstrate the relationship between orthogonality and discrimination using the classical Hopfield-Ninio scheme, shown graphically in Figure 2(a). Here, substrates $S = \{W,\ R\}$ compete

to form complexes with enzyme E. 'Wrong' and 'Right' products are formed from substrates R and W (respectively), at rates proportional to the steady state occupancy of the final complex ρ_{ES}. We thus define the *error fraction* achieved by the discrimination scheme to be

$$\xi = \frac{\rho_{EW}}{\rho_{ER}}.$$

Ninio and Hopfield designed this scheme to amplify differences in the binding energies of EW and ER complex formation. It is instructive to write the rate constants in Kramer's form.

We have for the EW reactions:

$$k'_W = \omega e^{\epsilon} \;\; l'_W = \omega_p$$

$$k_W = \omega e^{\gamma} \;\; l_W = \omega_p e^{\epsilon_p + \gamma}$$

where: ω, ω_p set overall rates; ϵ, ϵ_p represent the enthalpy differences between free enzyme and complexes ES^*, ES; and γ represents the binding energy difference between right and wrong complexes.

The ER reactions are given by:

$$k'_R = \omega e^{\epsilon+\delta} \;\; l'_R = \omega_p e^{-\delta_p}$$

$$k_R = \omega e^{\delta} \quad\; l_R = \omega_p e^{\epsilon_p - \delta_p}$$

where δ sets the activation energy differences between right and wrong complexes.

There is no discrimination along the transitions between the intermediary and final product:

$$m = \omega_i, \;\; m' = \omega_i e^{\epsilon_i}.$$

We begin by considering the relationship between error and orthogonality in the regime which is governed only by binding energy differences ($\gamma > 0,\ \delta = 0$), termed the 'energetic regime.' The Hopfield-Ninio scheme was originally designed for discrimination in this regime: the intermediary complex (ES^*) and dissipative drive (m') introduce a delay which amplifies the fact that EW complex formation has a faster off-rate (by factor of γ) relative to ER formation.

Simulations reveal that low orthogonality is necessary, but not sufficient, for low error rates in the energetic regime (Figure 2(b)). Analytically, we recall the long appreciated fact that the (energetic) Hopfield-Ninio scheme requires the parametric limit

$$\frac{\omega_p}{\omega e^{\epsilon}} \to 0$$

in order to minimize its error, and demonstrate that orthogonality is monotonically decreasing as this limit is approached (Appendix E).

A less well-appreciated requirement for energetic discrimination concerns the nonequilibrium drive, represented by m'. Some amount of drive is crucial for the discrimination scheme to work at all, but too much drive destroys discrimination [28]. We can understand this nonlinearity in terms of orthogonality (Figure 2(c)). Energy dissipation [22] is helpful for discrimination up *until* the point at which it begins to drive up orthogonality.

We next turn to the regime governed by only activation energy differences ($\gamma = 0,\ \delta > 0$), termed the 'kinetic regime.' Simulations reveal a bound opposite to that of the energetic regime: high orthogonality is necessary (but not sufficient) for low error (Supplemental Figure S2). Analytically, we can derive the error in this regime to be

$$\xi_{\text{kinetic}} = \frac{1 + e^{-\delta}\eta b + e^{-2\delta}\eta c}{1 + \eta b + \eta c}$$

where

$$a = \omega\omega_i, \quad b = \omega\omega_p, \quad c = \omega_p\omega_i e^{\epsilon_i}, \quad \eta = e^{\epsilon_p}/a.$$

The ξ_{kinetic} is minimized when $\eta \gg 1$ and $c \gg b$. That is, when there exists high drive ($\omega_i e^{\epsilon_i} \gg \omega$) and free enthalpy product differences ($\epsilon_p \gg 0$). We demonstrate that orthogonality is monotonically *increasing* as these limits are approached (Appendix E).

Differences between the two discriminatory regimes are summarized in Figure 2(d). Increasing the dissipative drive (ϵ_i) increases orthogonality, which allows for kinetic discrimination but precludes energetic discrimination.

The ability to modulate orthogonality via driving the single reaction m' suggests a simple strategy for dissipation-driven product switching. If products EW, ER are favored by different energy types, they can be selected between only by driving m' such that the network moves from low to high orthogonality. We achieve a four order of magnitude selection effect via this scheme (Figure 3).

Because the Hopfield-Ninio scheme only has one intermediary product, it is difficult to interpret in terms of the number of effective pathways towards the discriminatory products. In order to illustrate the connection between discrimination, effective pathways, and orthogonality, we turn to a more general setting.

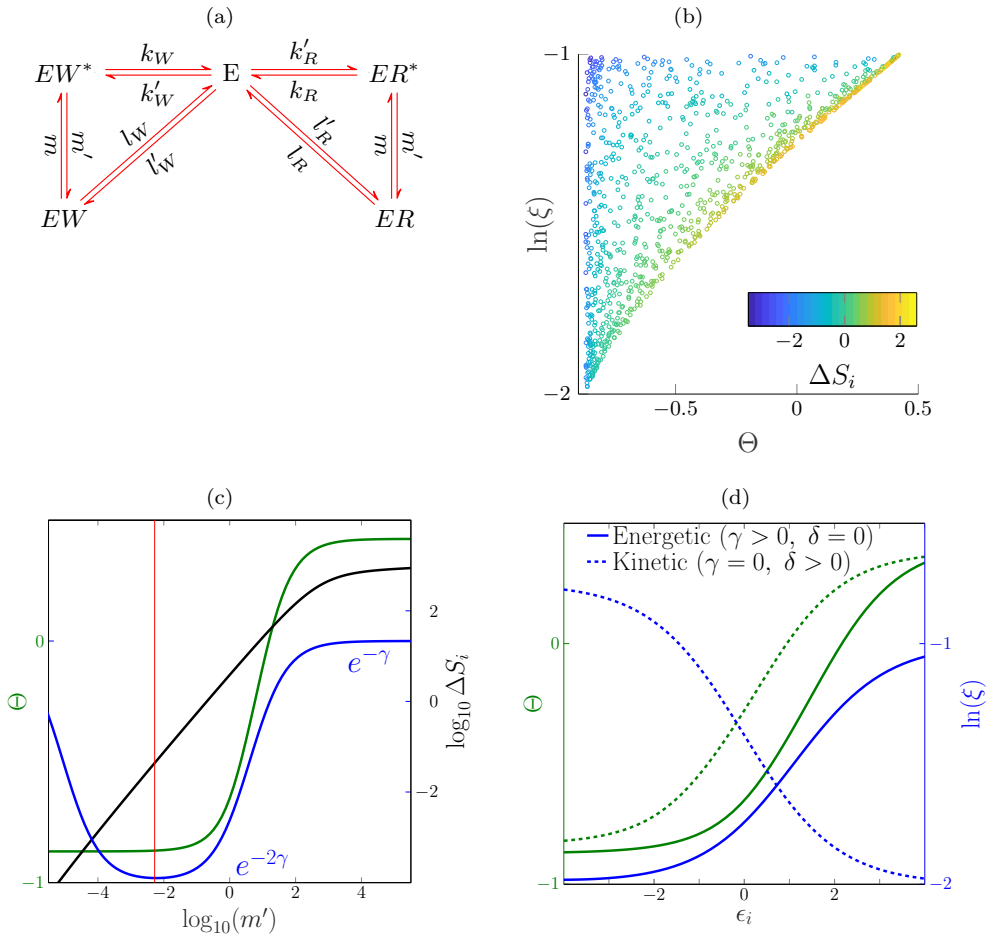


FIG. 2: Orthogonality in the Hopfield-Ninio scheme. (a) Reaction diagram of the scheme. Note that in the energetic regime, k_W and l_W will differ from k_R and l_R by a factor e^{γ}. In the kinetic regime, k_R and k'_R increase by a factor e^{δ} and l_W and l'_W increase by a factor e^{δ_p} [24]. (b) Orthogonality bounds minimum error rate in the energetic regime ($\gamma = 1, \delta = 0$). The log of the error rate ($\log(\xi)$) as a function of the orthogonality (Θ) is plotted for simulated data (parameter selection in Methods). Heatmap coloration represents relative dissipation ΔS_i [27] (hotter is higher dissipation); note that at a given orthogonality level, the error rate decreases as dissipation goes up (more red). (c) In the energetic regime, minimum error (red line, $\xi = e^{-2\gamma}$) is achieved by simultaneously minimizing orthogonality (Θ, green) and maximizing dissipation (black). Excess dissipation drives orthogonality upwards, asymptoting with error equal to the binding energy difference ($\xi = e^{-\gamma}$). (d) Orthogonality as a function of drive. In the energetic regime (solid curves), error rate (ξ) is minimized in the limit of low orthogonality (Θ). In the kinetic regime, (dashed curves), error rate is minimized in the limit of high orthogonality.

IV. ORTHOGONALITY IN A GENERAL SETTING

Murugan, Huse, and Leibler recently discovered that energetic discrimination in a general network requires a *discriminatory fence* [17], which can be idealized as a ladder graph having two sides, each with N loops (Figure 4(a)). The sides of the ladder are symmetric about the 0 node; the network aims to discriminate between states represented by its upper corners (i.e., x_{s2} in Figure 4(a)). Rate constants u^S, d^S, $S = \{W, R\}$ will differ for the 'Wrong' (W) and 'Right' (R) sides of the network.

The ladder idealization captures the fact that

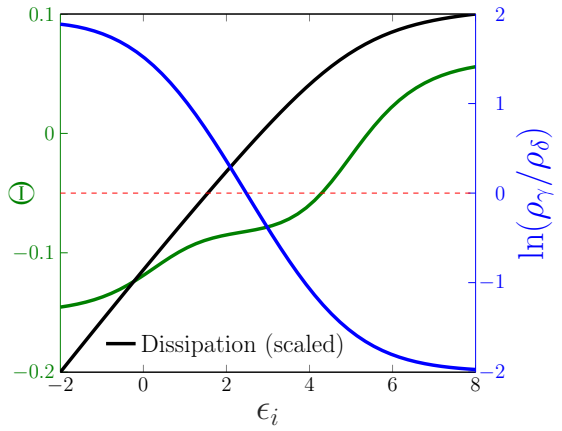


FIG. 3: A Hopfield-Ninio style network designed to tune product selectivity by modulating dissipation (black). One product ρ_γ has a lower binding energy and is favored in the energetic regime, while the other ρ_δ is has a lower activation energy and is favored in the kinetic regime. The log of the ratio between the products (ρ_γ/ρ_δ, blue), can be shifted from 2 (ρ_γ favored) to -2 (ρ_δ favored) by driving across a single reaction. This is due to orthogonality (green line) increasing, which shifts the network from the energetic to the kinetic regime.

a general energetic discrimination network must have dominant 'forward' (f) and 'reverse' (b) paths which are parallel to each other and effectively one-directional. On the pathway towards the product state, there is the constant threat of 'discard' (d), after which the reaction is exposed to a one-directional pathway away from the product state (b). There is also the possibility of 'rescue' (u) from discard.

The Kramer's form rate constants for this network are

$$u^R = \omega_d e^{\epsilon_u+\delta} \quad d^R = \omega_d e^{\delta}$$

$$u^W = \omega_d e^{\epsilon_u} \quad d^W = \omega_d e^{\gamma}.$$

And there is no discrimination ($f^R = f^W = f$) along the forward or reverse pathways:

$$f = \omega_f \quad b = \omega_b,$$

which we approximate to be one-directional for analytical convenience, but treat as bidirectional with small reverse rates when necessary for computing dissipation.

It is clear from the Kramer's form constants that to discriminate in the energetic regime (i.e., via γ), a high discard rate (d) is required. Indeed, the error rate for an N-loop network [23] in this regime is

$$\xi_{\text{energetic}} = \frac{1}{e^\gamma}\left(\frac{\omega_d + \omega_f}{\omega_d e^\gamma + \omega_f}\right)^N \tag{4}$$

which achieves its minimum when discards are high relative to steps toward reaction completion:

$$\omega_d/\omega_f \to \infty. \tag{5}$$

Discrimination in this regime is fundamentally processive, and global: accuracy relies on sequential exposure to frequently realized discard pathways, and *each* reaction step contributes to discrimination via the potential for discard. Correspondingly, orthogonality monotonically decreases in the Equation 5 limit (Appendix F), and is minimized in the high discard regime (Figure 4(b), solid lines).

In contrast, we find that the kinetic regime has error fraction given by (Appendix F):

$$\xi_{\text{kinetic}} = \frac{(\phi+1)^\alpha(1+\eta e^\delta)^\alpha}{(\phi e^\delta+1)^\alpha(\eta+1)^\alpha}. \tag{6}$$

where

$$\phi = \omega_d e^{\epsilon_u}/\omega_b, \quad \textit{and} \quad \eta = \omega_d/\omega_f.$$

The error ξ_{kinetic} is minimized when $\eta \to 0$ and $\phi \to \infty$, which is to say that:

$$\omega_d/\omega_f \to 0, \quad \omega_d e^{\epsilon_u}/\omega_b \to \infty. \tag{7}$$

These limits imply that network dynamics are being pushed quickly towards the final product nodes (ω_f, ϵ_u large, ω_b small). This makes local discrimination possible; and indeed orthogonality is monotonically increasing in the Equation 7 limit (Appendix G).

Quick movement towards final product nodes is in opposition to high discard rates; we can thus summarize the difference between the energetic and kinetic regimes by observing their difference with respect to the discard rate (d, Figure 4(b) x-axis), which reveal the expected orthogonality-error relationships in the two regimes. Note that these limits correspond to the dynamical phase localization limits described in [18].

We are now in a position to understand the orthogonality of this model in terms of its effective pathways towards the final product nodes. The energetic discrimination requirement that $f << d$ means that the network effectively contains only a single pathway to the product. Intuitively, the single

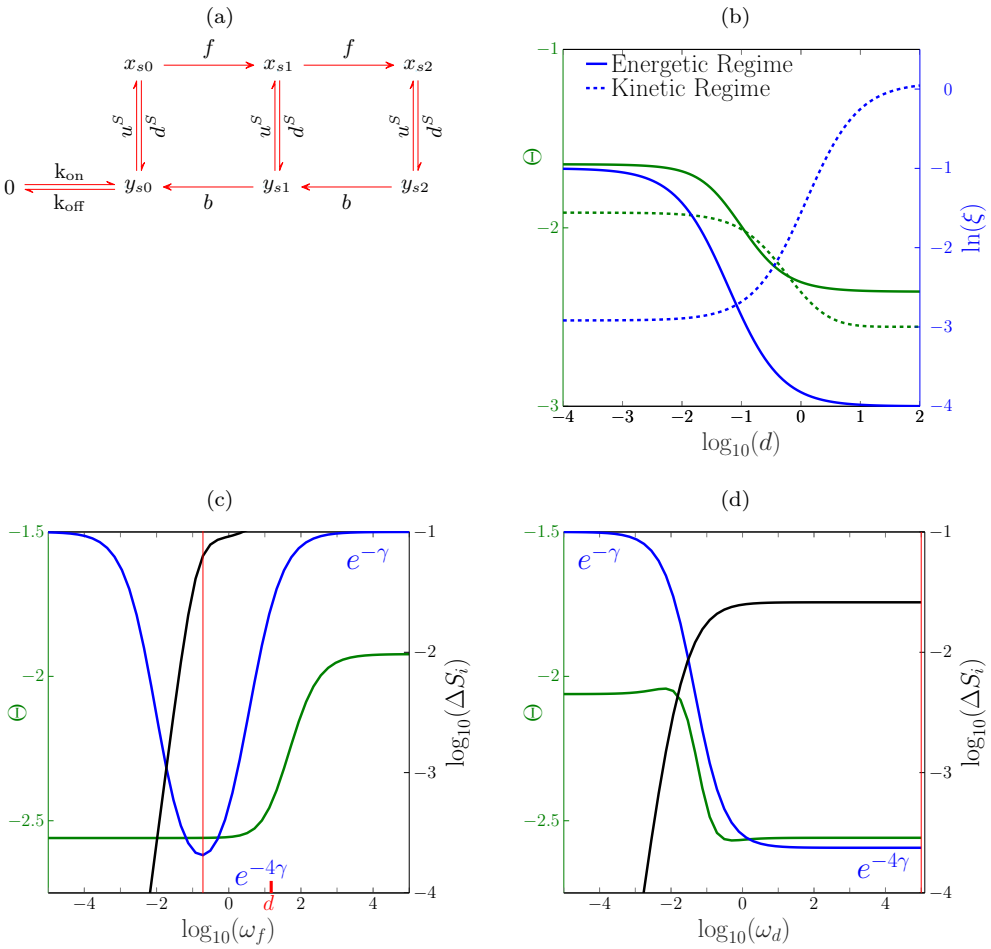


FIG. 4: (a) One side of the generalized ladder network [16]. The full ladder contains a second side, symmetric about the 0 node. The two sides of the ladder have different u^S, d^S constants ($S = \{R, W\}$ for 'right' and 'wrong' sides of the ladder, respectively). (b) Orthogonality and error for the two-loop ladder. In the energetic regime (δ=0, solid curves), minimum error (blue) is achieved in the low orthogonality (green) limit. In the kinetic regime (γ=0, dashed curves), minimum error is achieved in the high orthogonality limit. (c) Non-monotonicity in the energetic regime. The error rate (ξ, blue) is minimized (red line, $\xi = e^{-4\gamma}$ corresponding to $e^{-2\gamma}$ proofreading per loop) where dissipation (black) is maximized and orthogonality (Θ, green) is minimized. Red tick indicates value of rate $d \approx 15$. (d) Orthogonality is not always an increasing function of dissipation. Dissipation (black), error (blue), and orthogonality (green) for a two-loop ladder network in the energetic regime for which dissipation is maximized as orthogonality is minimized. Note that the error rate is minimized (red line, $\xi = e^{-4\gamma}$) at lower dissipation than in the energetic-regime network at left (black line in (c) vs black line in (d)).

pathway results from the slowness of one-directional progress towards the final product; rescue pathways cannot add additional paths to the final product because they are effectively equilibrated relative to the slow forward progress. Corresponding to this intuition, we find analytically that u, b, have essentially no effect on orthogonality in the $f << d$ regime (Appendix G). This argument is consistent with the fact that the discrimination error in the energetic regime (Equation 4) is independent of u, b, but in the kinetic regime, which requires $d << f$, we find that u, b are important factors in the error expression (Equation 6) and orthogonality requirements (Equation 7).

In the energetic regime, we observe that as f becomes close to d (red tick, Figure 4(c)), orthogonality rises sharply Figure 4(c). We understand this to result from many more effective pathways now leading to the final product. Again, the rise in orthogonality as we increase f leads to the non-monotonic behavior of the error rate.

Our understanding of orthogonality in terms of effective pathways allows us to apply thermodynamic drive in the energetic regime such that drive does *not* increase orthogonality. Our arguments above state that $f << d$, enforces the single pathway and hence orthogonal regime. Therefore, if we dissipate energy to drive d, we should find that the orthogonality decreases, and indeed we do (Figure 4(d)). Note that Figure 4(c) was generated with the same parameters as Figure 4(d); all that's changed is the reaction we choose to drive. In this parametric limit, the orthogonality and dissipation requirements are not contravening.

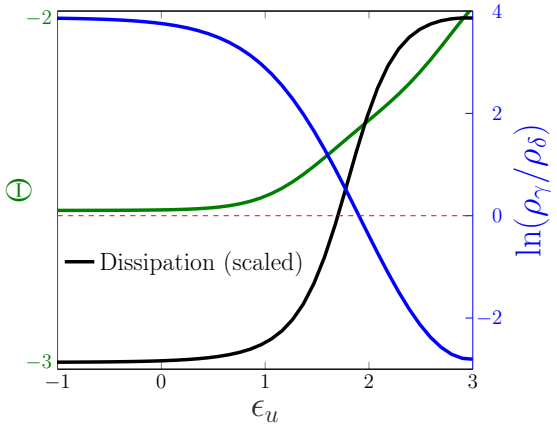


FIG. 5: The general ladder network can also achieve sensitive product switching. In this network, binding energies favor the product (ρ_γ) on one side of the ladder while activation energies favor the other product (ρ_δ). Dissipation is used to drive ϵ_u, increasing the ratio of rescues to discards u^S/d^S, thereby shifting the network from low orthogonality (ρ_γ favored) to high orthogonality (ρ_δ favored).

Finally, we note that (as in the Hopfield-Ninio regime) highly selective - seven orders of magnitude - dissipation driven product switching is possible between states which are favored by different energy types (Figure 5).

V. DISCUSSION

We have introduced orthogonality, which measures the degree to which a ratio of non-equilibrium steady states can be represented by rates local to the discriminatory nodes. We found that orthogonality tends to increase with the number of effectively realizable pathways directed towards the discriminatory nodes.

This connection between orthogonality and realizable pathways underlies its role in non-equilibrium discrimination. In order to discriminate via binding energies, networks require having a single dominant path along which discrimination occurs via frequently discarding intermediary products. These processes are inherently processive: discrimination is a global function of discards at sequential steps throughout the graph. Final product formation is rare, thus slow. In contrast, discrimination via kinetic barriers is fast. In the kinetic regime, discrimination relies on creating final products quickly, enabled by distributive networks which have many paths towards the final products. We thus find that orthogonal networks are necessary for kinetic discrimination, whereas non-orthogonal networks are necessary for energetic discrimination.

It is interesting to consider this result in the context of protein complex assembly [19]. Sartori and Leibler [25] have recently proposed that a significant proportion of the discrimination necessary for accurate protein complex assembly can be achieved by equilibrium energy differences in protein-protein interactions. Our results predict that non-equilibrium mechanisms which amplify these energetic differences should result in complexes being assembled sequentially, and slowly. If non-equilibrium mechanisms instead amplify kinetic differences to achieve accurate assembly, we expect a complex's component subunits to assemble in many different orders, quickly.

Our results clarify the role of thermodynamic drive in nonequilibrium discrimination. We find that both kinetic and energetic discrimination are enhanced by increasing dissipation, but are subject to necessary requirements on orthogonality, which itself can be modulated upwards or downwards by free energy expenditure.

By modulating orthogonality with energy expenditure, discriminatory networks can achieve sensitive product switching. In particular, driving a *single* reaction type is sufficient for sharp selection between products, if the products are favored by different energy types and driving shifts the orthogonality of the

network.

Biologically, this possibility may be realized in cytoplasmic ribonucleoprotein (RNP) granules [6]. These granules are composed of RNAs and proteins coloclazied in liquid-liquid phase separated droplets. Their liquid-liquid like components interact promiscuously, and are known to be enriched for multivalent components [2]. RNA contributes to promiscuous granule interactions via both RNA-RNA interactions and serving as a protein scaffold. Both RNA-RNA interactions and the number of RNA-protein contacts are dependent on RNA secondary structure [8]. It thus appears that RNA secondary structure can modulate the orthogonality of the granule interaction network.

RNA structure is appealing as a modulator of orthogonality because it can be modified by driving a single reaction type. It has been recently reported that ATP within granules is hydrolyzed by DEAD-box proteins, which remodel RNA by unwinding duplexes [12]. The (ATP-driven) DEAD-box unwinding of RNA has been reported responsible for the dynamic makeup of RNA inside of granules, and for granule dissolution. It is possible that driving this reaction type can tune the orthogonality of granule interaction networks.

Whether, and in which direction, ATP-driven RNA unwinding tunes orthogonality will depend on the molecular components of the granule. These components are not fixed; granules constantly exchange material with the local environment and are capable of exchanging components with each other. This combination of dynamic components and orthogonality driven selection may allow the cell to use existing components to explore new areas of biochemical reaction space. Such an ability is consistent with the apparent importance of granules in a wide variety of cellular responses to environmental cues, including stress response [7], transcriptional regulation [1], and local, activity dependent translation of mRNA at neuronal synapses [3, 14].

ACKNOWLEDGEMENTS

The authors would like to thank Tom Shimizu for useful discussions helping us to clarify the meaning of orthogonality and Gergo Bohner and Greg Wayne for useful discussions and Gergo Bohner and Pablo Sartori for critical reading of the manuscript.

The authors declare no competing or conflicting interests.

This work is partially supported by grants from the Wellcome Trust (104640/Z/14/Z, 092096/Z/10/Z) to E.A.M. G.V. is supported by a grant from Emergent Ventures. D.J. is funded by the Herchel Smith Post-doctoral Fellowship.

[1] Paul Anderson and Nancy Kedersha. RNA granules: post-transcriptional and epigenetic modulators of gene expression. *Nature Reviews Molecular Cell Biology*, 10(6):430–436, June 2009.

[2] Salman F. Banani, Hyun O. Lee, Anthony A. Hyman, and Michael K. Rosen. Biomolecular condensates: organizers of cellular biochemistry. *Nature Reviews Molecular Cell Biology*, 18(5):285–298, February 2017.

[3] Scott A. Barbee, Patricia S. Estes, Anne-Marie Cziko, Jens Hillebrand, Rene A. Luedeman, Jeff M. Coller, Nick Johnson, Iris C. Howlett, Cuiyun Geng, Ryu Ueda, Andrea H. Brand, Sarah F. Newbury, James E. Wilhelm, Richard B. Levine, Akira Nakamura, Roy Parker, and Mani Ramaswami. Staufen- and FMRP-containing neuronal RNPs are structurally and functionally related to somatic p bodies. *Neuron*, 52(6):997–1009, December 2006.

[4] Charles H. Bennett. Dissipation-error tradeoff in proofreading. *Biosystems*, 11(2-3):85–91, August 1979.

[5] Charles H. Bennett. The thermodynamics of computation. *International Journal of Theoretical Physics*, 21(12):905–940, December 1982.

[6] C. P. Brangwynne, C. R. Eckmann, D. S. Courson, A. Rybarska, C. Hoege, J. Gharakhani, F. Julicher, and A. A. Hyman. Germline p granules are liquid droplets that localize by controlled dissolution/condensation. *Science*, 324(5935):1729–1732, May 2009.

[7] J. Ross Buchan and Roy Parker. Eukaryotic stress granules: The ins and outs of translation. *Molecular Cell*, 36(6):932–941, December 2009.

[8] Natalia Sanchez de Groot, Alexandros Armaos, Ricardo Graña-Montes, Marion Alriquet, Giulia Calloni, R. Martin Vabulas, and Gian Gaetano Tartaglia. RNA structure drives interaction with proteins. *Nature Communications*, 10(1), July 2019.

[9] Daniel Drucker. A comprehensive pythagorean theorem for all dimensions. *The American Mathematical Monthly*, 122(2):164, 2015.

[10] M. Ehrenberg and C. Blomberg. Thermodynamic constraints on kinetic proofreading in biosynthetic pathways. *Biophysical Journal*, 31(3):333–358, September 1980.

[11] Eugene Gover and Nishan Krikorian. Determinants

and the volumes of parallelotopes and zonotopes. *Linear Algebra and its Applications*, 433(1):28–40, July 2010.

[12] Maria Hondele, Ruchika Sachdev, Stephanie Heinrich, Juan Wang, Pascal Vallotton, Beatriz M. A. Fontoura, and Karsten Weis. DEAD-box ATPases are global regulators of phase-separated organelles. *Nature*, 573(7772):144–148, August 2019.

[13] J. J. Hopfield. Kinetic proofreading: A new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proceedings of the National Academy of Sciences*, 71(10):4135–4139, October 1974.

[14] C. McCann, E. E. Holohan, S. Das, A. Dervan, A. Larkin, J. A. Lee, V. Rodrigues, R. Parker, and M. Ramaswami. The ataxin-2 protein is required for microRNA function and synapse-specific long-term olfactory habituation. *Proceedings of the National Academy of Sciences*, 108(36):E655–E662, July 2011.

[15] Inomzhon Mirzaev and Jeremy Gunawardena. Laplacian dynamics on general graphs. *Bulletin of Mathematical Biology*, 75(11):2118–2149, September 2013.

[16] Arvind Murugan, David A Huse, and Stanislas Leibler. Speed, dissipation, and error in kinetic proofreading. *Proceedings of the National Academy of Sciences*, 109(30):12034–12039, 2012.

[17] Arvind Murugan, David A Huse, and Stanislas Leibler. Discriminatory proofreading regimes in nonequilibrium systems. *Physical Review X*, 4(2):021016, 2014.

[18] Arvind Murugan and Suriyanarayanan Vaikuntanathan. Biological implications of dynamical phases in non-equilibrium networks. *Journal of Statistical Physics*, 162(5):1183–1202, February 2016.

[19] Arvind Murugan, Zorana Zeravcic, Michael P. Brenner, and Stanislas Leibler. Multifarious assembly mixtures: Systems allowing retrieval of diverse stored structures. *Proceedings of the National Academy of Sciences*, 112(1):54–59, December 2014.

[20] Jacques Ninio. Kinetic amplification of enzyme discrimination. *Biochimie*, 57(5):587–595, July 1975.

[21] Fully-connected graphs are ones in which it is possible to (not necessary directly) get from every state to every other state.

[22] Throughout this text dissipation is calculated as the entorpy production rate, $\Delta S_i = \frac{1}{2}\sum_{i,j}(k_{ij}\rho_j - k_{ji}\rho_i)\ln\frac{k_{ij}\rho_j}{k_{ji}\rho_i}$.

[23] An N-loop network will strictly speaking be composed of $2N+1$ loops, N on each side of the ladder and a single reactant node.

[24] Riccardo Rao and Luca Peliti. Thermodynamics of accuracy in kinetic proofreading: dissipation and efficiency trade-offs. *Journal of Statistical Mechanics: Theory and Experiment*, (6):P06001, 20, 2015.

[25] Pablo Sartori and Stanislas Leibler. Lessons from equilibrium statistical physics regarding the assembly of protein complexes. *Proceedings of the National Academy of Sciences*, 117(1):114–120, December 2019.

[26] Pablo Sartori and Simone Pigolotti. Kinetic versus energetic discrimination in biological copying. *Physical Review Letters*, 110(18), May 2013.

[27] Jürgen Schnakenberg. Network theory of microscopic and macroscopic behavior of master equation systems. *Reviews of Modern physics*, 48(4):571, 1976.

[28] Felix Wong, Ariel Amir, and Jeremy Gunawardena. Energy-speed-accuracy relation in complex networks for biological discrimination. *Physical Review E*, 98(1), July 2018.

Appendices

Appendix A: An expression for the discrimination ratio

In this section, we prove the following proposition.

Proposition A.1 (An expression for the ratio of kernel elements)**.** *Let $\mathcal{L}$ be a Laplacian matrix*

$$\mathcal{L}_{ij} = \begin{cases} k_{ij} & i \neq j,\ k_{ij} \geq 0 \\ -\sum_j k_{ij} & i = j \end{cases} \tag{A1}$$

representing a fully connected continuous time Markov chain. Note that the nullspace of $\mathcal{L}$ is one dimensional; let it have basis ρ. The ratio of any two elements of ρ is given by

$$\frac{\rho_i}{\rho_j} = \frac{\|v_i - \mathrm{proj}_{\mathcal{L}^{i,j}}(v_i)\|}{\|v_j - \mathrm{proj}_{\mathcal{L}^{i,j}}(v_j)\|}.$$

Proof. The result follows from three equalities:

$$\begin{aligned} \frac{\rho_i}{\rho_j} &= \frac{\det(\mathcal{L}^{kj})}{\det(\mathcal{L}^{ki})} \quad \forall\ k \in 1 \ldots N \\ &= \frac{\mathrm{vol}(P(\mathcal{L}^{0j}))}{\mathrm{vol}(P(\mathcal{L}^{0i}))} \\ &= \frac{\|v_i - \mathrm{proj}_{\mathcal{L}^{i,j}}(v_i)\|}{\|v_j - \mathrm{proj}_{\mathcal{L}^{i,j}}(v_j)\|} \end{aligned} \tag{A2}$$

where $\mathcal{L}^{ki}$ represents the matrix formed by removing row k and column i from matrix $\mathcal{L}$, and $\mathcal{L}^{0k}$ is formed from $\mathcal{L}$ by removing column k only. For matrix A, $\mathrm{vol}(P(A))$ represents the volume the parallelotope formed by the columns of A; vector v_i represents the ith column of $\mathcal{L}$; and $\mathrm{proj}_{\mathcal{L}^{i,j}}(v_i)$ denotes the projection of vector v_i onto the $n-2$ dimensional subspace S spanned by the columns of $\mathcal{L}$ remove i, j.

Please note that in our notation, $A^{xy} \neq A^{x,y}$. The former (no superscript commas) denotes the matrix A remove row x and column y. The latter (commas in superscript) denotes the matrix A remove *column* x and column y. □

We proceed by proving each of the three equalities in Equation A2. To prove the first equality, it will be useful to have the definition of the adjugate matrix at hand.

Definition A.2 (Adjugate matrix)**.** *The components of the adjugate of a matrix A,* $\text{adj}(A)$*, are given by taking the transpose of the cofactor matrix, C, of A:*

$$\begin{aligned} \text{adj}(A)_{ij} &= C_{ji} \\ &= \det(A^{ji}) \end{aligned} \tag{A3}$$

where A^{ji} is denotes the $(n-1) \times (n-1)$ matrix resulting form removing row j and column i from A.

For the second equality in Equation A3, we have recalled that the elements of the cofactor matrix C_{ij} of A are given by (up to sign):

$$C_{ij} = \det(A^{ij})$$

where A^{ij} denotes the $(n-1) \times (n-1)$ matrix resulting form removing row i and column j from A.

We can now prove the first equality in A2.

Proposition A.3 (Discrimination ratio in terms of determinants with column and row cuts)**.** *We aim to demonstrate that*

$$\frac{\rho_i}{\rho_j} = \frac{\det(\mathcal{L}^{kj})}{\det(\mathcal{L}^{ki})} \quad \forall\ k \in 1 \dots N.$$

Proof. The proposition was proved in [15]. We include the argument here for completeness.

By the Matrix-Tree theorem, the rank of a fully-connected N dimensional Laplacian matrix is $N-1$. The nullspace is therefore one-dimensional, and can be represented by a single basis vector ρ.

It will suffice to prove that $\rho_i = \det(\mathcal{L}^{ki})$. Recall the Laplace expansion for the determinant:

$$\begin{aligned} \text{adj}(\mathcal{L}) \cdot \mathcal{L} = \mathcal{L} \cdot \text{adj}(\mathcal{L}) &= \det(\mathcal{L}) \cdot I \\ &= 0_{n \times n}, \end{aligned} \tag{A4}$$

where $0_{n \times n}$ denotes the n by n zero matrix and the final equality follows from $\mathcal{L}$ not being full rank, hence $\det(\mathcal{L}) = 0$.

Consider that $\mathcal{L} \cdot \text{adj}(\mathcal{L}) = 0$ implies that $\mathcal{L}v = 0_{n \times 1}$ for all v which are columns of $\text{adj}(\mathcal{L})$. That is: the columns of $\text{adj}(\mathcal{L})$ are equal to ρ. This gives the result. □

We now prove the second equality in Equation A2.

Proposition A.4 (Discrimination ratio in terms of column cuts only)**.** *We now wish to demonstrate that the equality presented in the previous proposition does not require the removal of some row k [9]:*

$$\begin{aligned} \frac{\det(\mathcal{L}^{ki})}{\det(\mathcal{L}^{kj})} &= \frac{\text{vol}(P(\mathcal{L}^{ki}))}{\text{vol}(P(\mathcal{L}^{kj}))} \\ &= \frac{\text{vol}(P(\mathcal{L}^{0i}))}{\text{vol}(P(\mathcal{L}^{0j}))} \end{aligned}$$

Proof. The first equality is a common characterization of the determinant. The second result follows from a series of equalities

$$\begin{aligned} \frac{\text{vol}(P(\mathcal{L}^{0i}))}{\text{vol}(P(\mathcal{L}^{0j}))} &= \frac{\sqrt{\det[(\mathcal{L}^{0i})^T(\mathcal{L}^{0i})]}}{\sqrt{\det[(\mathcal{L}^{0j})^T(\mathcal{L}^{0j})]}} \\ &= \sqrt{\frac{\sum_k (\det[\mathcal{L}^{ki}])^2}{\sum_k (\det[\mathcal{L}^{kj}])^2}} \\ &= \sqrt{\frac{N(\det(\mathcal{L}^{ki}))^2}{N(\det(\mathcal{L}^{kj}))^2}} = \frac{\det(\mathcal{L}^{ki})}{\det(\mathcal{L}^{kj})} \end{aligned}$$

where: the first equality is by definition of a polytope volume generated by a non-square matrix; the second equality results from applying the Cauchy-Binet formula; the third equality follows from noting that $\det(\mathcal{L}^{ki}) = \det(\mathcal{L}^{k'i}),\ \forall\ k, k' \in 1 \dots N.$ □

We now prove the final equality in Equation A2. First, it is useful to recall the base-height formula for determinants.

Fact A.5 (The base-height formula)**.** The determinant of a matrix A can be written as

$$\det(A) = \prod_i \|a_i\|$$

where a_i is a vector representing the component of v_i that is perpendicular to the subspace spanned by the $N-i$ vectors $\{v_{i+1}, \cdots, v_n\}$. Crucially, this procedure can be done by selecting the v_i in any order [11].

Proof. Geometrically, the determinant of a matrix A having columns v_i can be thought of as the volume of the parallelotope generated by the v_i. Consider a parallelotope $P(A)$ generated by vectors $\{v_1, \cdots, v_n\}$. $P(A)$ can also be thought of as a prism with base generated by the vectors $\{v_2, \cdots, v_n\}$ and height equal to the magnitude of the component of

v_1 perpendicular to the span of $\{v_2, \cdots, v_n\}$. It follows that

$$\begin{aligned}\mathrm{vol}_n(P(A)) =& \mathrm{vol}_{n-1}(P(\{v_2, \cdots, v_n\})) \cdot \\ & \|v_1 - \mathrm{proj}(v_1; v_2, \cdots, v_n)\|\end{aligned}$$

And of course we can carry out this procedure successively for $\mathrm{vol}_{n-1}, \mathrm{vol}_{n-2}, \ldots$. This gives the desired result. □

Proposition A.6 (Discriminatory ratio in terms of normalized projections)**.** *Finally, we demonstrate that*

$$\frac{\mathrm{vol}(P(\mathcal{L}^{0i}))}{\mathrm{vol}(P(\mathcal{L}^{0j}))} = \frac{\|v_j - \mathrm{proj}_S(v_j)\|}{\|v_i - \mathrm{proj}_S(v_i)\|}$$

Proof. The result follows directly from the base-height formula for determinants.

$$\begin{aligned}\frac{\det(\mathcal{L}^{0i})}{\det(\mathcal{L}^{0j})} &= \frac{\|v_j - \mathrm{proj}_{\mathcal{L}^{i,j}}(v_j)\| \cdot \mathrm{vol}_{n-2} P(\{v_l\}_{l \neq i,j})}{\|v_i - \mathrm{proj}_{\mathcal{L}^{i,j}}(v_i)\| \cdot \mathrm{vol}_{n-2} P(\{v_l\}_{l \neq i,j})} \\ &= \frac{\|v_j - \mathrm{proj}_{\mathcal{L}^{i,j}}(v_j)\|}{\|v_i - \mathrm{proj}_{\mathcal{L}^{i,j}}(v_i)\|}\end{aligned}$$

where $\mathrm{proj}_{\mathcal{L}^{i,j}}(v_j)$ denotes the projection of vector v_j onto the subspace spanned by the vectors of matrix $\mathcal{L}^{i,j}$, formed by deleting columns $i,\ j$ from $\mathcal{L}$. Notice that in the numerator, we have chosen to begin the base-height iteration with vector v_j. Because L^{0i} already has column i removed, this procedure yields - in the numerator - a polytope base generated by the non-i, j columns in $\mathcal{L}$. In the denominator, beginning the base-height iteration v_i also yields a polytope base generated by the non-i, j columns. These bases cancel to give the desired result. □

Appendix B: An expression for the error of a projection approximation

In this section, we aim to prove the following proposition.

Proposition B.1 (Normed projection approximation)**.** *Let S be a matrix having full rank (note that our $\mathcal{L}^{i,j}$ are indeed of full rank). We have that*

$$\left\|(S(S^TS)^{-1}S^T) - SS^T\right\| = \left\|I - S^TS\right\|.$$

Proof. Let S have singular value decomposition $S = U\Sigma W^T$.

$$I - S^TS = I - W\Sigma^T\Sigma W^T = W[I - \Sigma^T\Sigma]W^T.$$

And similarly (noting that S^TS is invertible because S is full rank):

$$\begin{aligned}SS^T - S(S^TS)^{-1}S^T &= U\Sigma\Sigma^TU^T - U\Sigma W^T(W\Sigma^T\Sigma W^T)^{-1}W\Sigma^TU^T \\ &= U\Sigma\Sigma^TU^T - U\Sigma W^TW(\Sigma^T\Sigma)^{-1}W^TW\Sigma^TU^T \\ &= U\Sigma\Sigma^TU^T - U\Sigma(\Sigma^T\Sigma)^{-1}\Sigma^TU^T \\ &= U(\Sigma\Sigma^T - \Sigma(\Sigma^T\Sigma)^{-1}\Sigma^T)U^T\end{aligned}$$

It follows by direct calculation (Σ is diagonal) that

$$\left\|\Sigma\Sigma^T - \Sigma(\Sigma^T\Sigma)^{-1}\Sigma^T\right\| = \left\|I - \Sigma^T\Sigma\right\|.$$

This gives the result. □

Appendix C: Orthogonality of the line versus all-to-all graph

In this Appendix we demonstrate that the orthogonality of an N node line graph is strictly less than an N node all-to-all connected graph, in the toy case where all rate constants are the same. The result follows from directly calculating the orthogonality for each topology, which we do in turn.

Proposition C.1 (Θ for a line graph)**.** *For a a line graph with bidirectional connections of equal weight (set to 1 without loss of generality), the orthogonality is given by:* $\Theta = 1 - \sqrt{(N-1)\frac{8}{9} + \frac{1}{36}(N-4)}$.

Proof. The result follows from direct computation of $\langle i,\ j\rangle,\ \forall\ i \neq 1, N$. There are only two types of nonzero $\langle i, j\rangle$. The first type is $\langle i, i+1\rangle$; there exist $2(N-1)$ terms of this type. The second type is $\langle i, i+2\rangle$; there exist $N-4$ entries of this type.

The first type of nonzero term represents 'neighbors.' The second represents nodes separated by one node, which point at a mutual node. The two types of inner product have (squared, normalized) values:

$$\langle i, i+1\rangle^2 = \frac{(-\alpha \cdot 2\alpha - \alpha \cdot 2\alpha)^2}{(2\alpha^2 + 4\alpha^2)^2} = \frac{4}{9}$$

and

$$\langle i, i+2\rangle^2 = \frac{(\alpha^2)^2}{(6\alpha^2)^2} = \frac{1}{36}.$$

The result follows. □

The all-to-all calculation is slightly more complicated.

Proposition C.2 (Θ for an all-to-all graph)**.** *For an all-to-all connected graph with bidirectional connections of equal weight (set to 1 without loss of generality), the orthogonality is given by:* $\Theta = 1 - \sqrt{\frac{(N-2)(N-3)}{(N-1)^2}}$.

Proof. Let S be the n by $n-2$ matrix formed by removing two of the columns of the Laplacian for this graph.

Because the diagonal elements $(S^TS)_{ii} = 1$, we need only compute the off-diagonal elements of S^TS. A generic such element resulting from taking the (not normalized) inner product of columns j, k is given by

$$\begin{aligned}\langle j, k\rangle &= \sum_{\substack{i \\ i\neq j \\ i\neq k}} \theta_{ij}\theta_{ik} - \theta_{jk} \cdot \sum_{\substack{i \\ i\neq j}} \theta_{ij} - \theta_{kj} \sum_{\substack{i \\ i\neq k}} \theta_{ik} \\ &= (N-2)\alpha^2 - \alpha^2(N-1) - \alpha^2(N-1) \\ &= -\alpha^2 N.\end{aligned}$$

where the first line is a generic expression for the inner product of columns corresponding to connected nodes for matrix elements θ_{ij} of S, and the resulting lines follow from bidirectional all-to-all connectivity with equal rate constants.

We now need to compute the normalization factor:

$$\begin{aligned}(\|j\|\,\|k\|)^2 &= \left(\sum_{\substack{i \\ i\neq j}} \theta_{ij}^2 + \left(\sum_{\substack{i \\ i\neq j}} \theta_{ij}\right)^2\right) \\ &\quad \cdot \left(\sum_{\substack{i \\ i\neq k}} \theta_{ik}^2 + \left(\sum_{\substack{i \\ i\neq k}} \theta_{ik}\right)^2\right) \\ &= \left(\alpha^2(N-1) + (N-1)^2\alpha^2\right)^2 \\ &= \left(\alpha^2(N^2-N)\right)^2 \\ &= \alpha^4(N^2-N)^2\end{aligned}$$

where again we have begun with generic terms for the normalization of the inner product of columns of the Laplacian matrix, with θ_{ij} representing the elements of S.

Putting these together yields the expression for a generic element of S^TS:

$$\begin{aligned}\frac{\langle j, k\rangle^2}{(\|i\|\,\|j\|)^2} &= \frac{\alpha^4 N^2}{\alpha^4(N^2-N)^2} \\ &= \frac{1}{(N-1)^2}.\end{aligned}$$

How many such elements exist? We know that S^TS is a square $n-2$ length matrix, and we know that the diagonal terms are zero. We therefore have $(n-2)(n-3)$ entries each equal to $\frac{1}{(N-1)^2}$. The result follows. □

From the two propositions we can calculate that

$$\begin{aligned}\Theta_{\text{all-to-all}} - \Theta_{\text{line}} = &-\sqrt{\frac{(N-2)(N-3)}{(N-1)^2}} \\ &+ \sqrt{(N-1)\frac{8}{9} + \frac{1}{36}(N-4)}\end{aligned}$$

The former (negative) term quickly asymtotes to 1, whereas the latter (positive) term grows as $O(\sqrt{N})$. We conclude that the orthogonality of the all-to-all graph is greater than the line graph, and this difference is increasing for increasing N.

Appendix D: Orthogonality in the 4-Node toy model

We will show how orthogonality changes as the graph in Figure 1(a) is modified, in support of the

claims made in the main text. Because we are discriminating between the end nodes, the orthogonality of the scheme in Figure 1(a) is a function of a single (normalized) inner product:

$$\begin{aligned}\langle v_2, v_3\rangle^2 &= \tfrac{(2k(2k+l)-2kl)^2}{(2k^2+l^2+(2k+l)^2)^2} \\ &= \tfrac{4k^4}{(3k^2+2kl+l^2)^2}\end{aligned}$$

with k, l corresponding to black, red arrows in Figure 1(a), as defined in the main text.

We will first demonstrate how orthogonality changes as $r = k/l$ grows. We then demonstrate how orthogonality changes upon removing the black (bidirectional) connection between the middle nodes.

a. Adjusting rates to favor a single path reduces orthogonality We can rewrite Equation D1 in terms of $r = k/l$:

$$\langle v_2, v_3\rangle^2 = \frac{4r^4}{(3r^2+2r+1)^2}.$$

Two such terms contribute to the orthogonality giving

$$\begin{aligned}\Theta &= 1 - \sqrt{2\langle v_2, v_3\rangle^2} \\ &= 1 - \sqrt{\tfrac{8r^4}{(3r^2+2r+1)^2}}\end{aligned}$$

which is decreasing with r as $O(r^{-2})$, as claimed in the main text.

b. Removing a link What happens to the orthogonality when we remove the black bidirectional links between the middle nodes?

The expression for $\langle v_2, v_3\rangle^2_{\text{removed}}$ is given by

$$\begin{aligned}\langle v_2, v_3\rangle^2_{removed} &= \frac{k^2l^2}{\left(k^2+kl+l^2\right)^2} \\ &= \frac{r^2}{\left(r^2+r+1\right)^2}\end{aligned}$$

When $r \approx 1$ this expression is equal to Equation D1; there is no affect on orthogonality. However, as r increases, $\langle v_2, v_3\rangle^2_{\text{removed}}$ becomes smaller than Equation D1; deleting the connections increases orthogonality. We conclude that when $r > 1$, the black bidirectional links form part of the dominant path, removing them will therefore increase the orthogonality.

Appendix E: Error and Orthogonality in Ninio-Hopfield Model

We first consider the Hopfield model in the energetic regime. The Laplacian for this scheme with the columns corresponding to final products removed is given by

$$A = \begin{pmatrix} -\sum_1 & \omega e^{\gamma} & \omega \\ \omega e^{\epsilon} & -\sum_2 & 0 \\ \omega e^{\epsilon} & 0 & -\sum_3 \\ \omega_p & m' & 0 \\ \omega_p & 0 & m' \end{pmatrix}.$$

Orthogonality in this model will be a function of three inner products:

$$\Theta = 1 - \sqrt{2 * (s_{1,2}^2 + s_{1,3}^2 + s_{2,3}^2)}$$

where we have denoted the (normalized) inner product of the ith and jth elements of A as $s_{i,j}$. It will be useful to define and reason about

$$\sum s_{i,j}^2 = (s_{1,2}^2 + s_{1,3}^2 + s_{2,3}^2).$$

The relevant inner products are

$$\begin{aligned}s_{1,2}^2 &= \frac{\langle 1,2\rangle^2}{\left(\|1\|\,\|2\|\right)^2} = \frac{(3e^{\epsilon}\omega^2 + 2\omega\omega_p + e^{\epsilon}\omega m' - \omega_p m')^2}{4(3e^{2\epsilon}\omega^2 + 4e^{\epsilon}\omega_p\omega + 3\omega_p)(\omega^2 + \omega m' + m'^2)} \\ s_{1,3}^2 &= \frac{\langle 1,3\rangle^2}{\left(\|1\|\,\|3\|\right)^2} = \frac{(3e^{\epsilon+\gamma}\omega^2 + 2\omega\omega_p e^{\gamma} + e^{\epsilon}\omega m' - \omega_p m')^2}{4(3e^{2\epsilon}\omega^2 + 4e^{\epsilon}\omega_p\omega + 3\omega_p)(\omega^2 e^{2\gamma} + \omega m' e^{\gamma} + m'^2)} \\ s_{2,3}^2 &= \frac{\langle 2,3\rangle^2}{\left(\|2\|\,\|3\|\right)^2} = \frac{\omega^4 e^{2\gamma}}{4(\omega^2 + \omega m' + m'^2)(e^{2\gamma}\omega^2 + e^{\gamma}\omega m' + m'^2)}.\end{aligned}$$

We now demonstrate the orthogonality-discrimination relations made in the main text. To do so, we first compute the orthogonality in the high and low discrimination limits, in order to demonstrate that orthogonality is lower ($\sum s_{i,j}^2$ higher) as high discrimination improves. We will then compute the degree to which orthogonality movement between the low and high discrimination limits is monotonic.

In the energetic regime, the discrimination is maximized in the limits

$$\frac{\omega_p}{\omega e^{\epsilon}} \to 0$$

. We must therefore consider: $\omega \to \infty$,$\epsilon \to \infty$, and $\omega_p \to 0$.

a. Energetic Limit 1: $\omega \to \infty$ Note that we replace m' with μ in the below.

$\sum s_{i,j}^2$ is increasing with ω To demonstrate this, we will show the following.

$$\lim_{\omega\to\infty} \sum s_{i,j}^2 > \lim_{\omega\to 0} \sum s_{i,j}^2$$

Analytically, we can see that the in the limit of $\omega \to \infty$, only terms of order ω^4 remain. If we expand and collect the terms together in ω

$$s_{1,2}^2 = \frac{9w^4e^{2\epsilon}+w^3\left(12\omega_p e^{\epsilon}+6\mu e^{2\epsilon}\right)+w^2\left(4\omega_p^2-2\mu\omega_p e^{\epsilon}+\mu^2e^{2\epsilon}\right)+w\left(-4\mu\omega_p^2-2\mu^2\omega_p e^{\epsilon}\right)+\mu^2\omega_p^2}{12w^4e^{2\epsilon}+w^3(16\omega_p e^{\epsilon}+12\mu e^{2\epsilon})+w^2\left(12\omega_p^2+16\mu\omega_p e^{\epsilon}+12\mu^2e^{2\epsilon}\right)+w\left(12\mu\omega_p^2+16\mu^2\omega_p e^{\epsilon}\right)+12\mu^2\omega_p^2}$$

$$s_{1,3}^2 = \frac{9w^4e^{2\gamma+2\epsilon}+w^3\left(6\mu e^{\gamma+2\epsilon}+12e^{\gamma+\epsilon}\omega_p e^{\gamma}\right)+w^2\left(-6\mu\omega_p e^{\gamma+\epsilon}+4\omega_p e^{2\gamma}+4\mu e^{\epsilon}\omega_p e^{\gamma}+\mu^2e^{2\epsilon}\right)+w\left(-4\mu\omega_p\omega_p e^{\gamma}-2\mu^2\omega_p e^{\epsilon}\right)+\mu^2\omega_p^2}{12w^4e^{2\gamma+2\epsilon}+w^3(12e^{2\epsilon}\mu e^{\gamma}+16\omega_p e^{2\gamma+\epsilon})+w^2\left(12e^{2\gamma}\omega_p^2+16\omega_p e^{\epsilon}\mu e^{\gamma}+12\mu^2e^{2\epsilon}\right)+w\left(12\omega_p^2\mu e^{\gamma}+16\mu^2\omega_p e^{\epsilon}\right)+12\mu^2\omega_p^2}$$

$$s_{2,3}^2 = \frac{\omega^4e^{2\gamma}}{4e^{2\gamma}w^4+w^3(4e^{\gamma}\mu+4e^{2\gamma}\mu)+w^2(4e^{\gamma}\mu^2+4e^{2\gamma}\mu^2+4\mu^2)+w(4e^{\gamma}\mu^3+4\mu^3)+4\mu^4}$$

Thus, in the limit $\omega \to \infty$,

$$\lim_{\omega\to\infty} \sum s_{i,j}^2 = \frac{9e^{2\epsilon}}{12e^{2\epsilon}} + \frac{9e^{2\epsilon+2\gamma}}{12e^{2\epsilon+2\gamma}} + \frac{e^{2\gamma}}{4e^{2\gamma}} = \frac{7}{4}$$

in the limit $\omega \to 0$, only the constant terms (those not multiplied by ω) remain. We therefore have

$$\lim_{\omega\to 0} \sum s_{i,j}^2 = \frac{\mu^2\omega_p^2}{12\mu^2\omega_p^2} + \frac{\mu^2\omega_p^2}{12\mu^2\omega_p^2} + 0 = \frac{1}{6}$$

This gives the desired result:

$$\lim_{\omega\to\infty} \sum s_{i,j}^2 = \frac{7}{4} > \lim_{\omega\to 0} \sum s_{i,j}^2 = \frac{1}{6} \quad .$$

The increase in $\sum s_{i,j}^2$ in monotonic To demonstrate that the increase in $\sum s_{i,j}^2$ is monotonic in ω we must show that

$$\frac{d}{d\omega} \sum s_{i,j}^2 > 0$$

We will compute the derivatives of each of the components separately. The easiest is the $s_{2,3}^2$ term.

$$\frac{d}{d\omega} s_{2,3}^2 = \frac{e^{2\gamma}\mu w^3\left(e^{\gamma}w\left(3\mu^2+w^2+2\mu w\right)+e^{2\gamma}w^2(2\mu+w)+\mu\left(4\mu^2+2w^2+3\mu w\right)\right)}{4\left(\mu^2+w^2+\mu w\right)^2\left(\mu^2+e^{2\gamma}w^2+e^{\gamma}\mu w\right)^2}$$

which is greater than zero because all rate constants are positive. This is the desired result.

Now lets turn to the other two terms. It is sufficient to consider the numerator of the derivatives of $\sum s_{1,j}^2$

$$\frac{d}{d\omega}s_{1,j}^2 = 4(e^\epsilon\omega(m'+3\omega)-\omega_p(m'-2\omega))[e^\epsilon{\omega_p}^2(10m'^3+55m'^2\omega+39m'\omega^2+10\omega^3) + e^{2\epsilon}\omega_p m'\omega(10m'^2+45m'\omega+26\omega^2)+3e^{3\epsilon}m'\omega^3(5m'+\omega)+3{\omega_p}^3m'(5m'+4\omega)].$$

This term is positive except for the case

$$\omega_p m' > \omega_p 2\omega + e^\epsilon\omega m' + 2e^\epsilon\omega$$
$$1 > \frac{2\omega}{m'} + \frac{e^\epsilon\omega}{\omega_p} + \frac{2e^\epsilon\omega^2}{\omega_p m'}$$

which is only satisfied outside of the proofreading regime $\frac{\omega_p}{e^\epsilon\omega} > 1$.

b. Energetic Limit 2: $\epsilon \to \infty$

$\sum s_{i,j}^2$ is increasing with ϵ To demonstrate this, we will show the following.

$$\lim_{\epsilon\to\infty}\sum s_{i,j}^2 > \lim_{\epsilon\to-\infty}\sum s_{i,j}^2$$

First note that the term $s_{2,3}^2$ is not a function of ϵ. If we rearrange the other two $s_{i,j}^2$ terms to collect w.r.t ϵ we get,

$$s_{1,2}^2 = \frac{4w^2\omega_p^2+e^\epsilon\left(12w^3\omega_p-2\mu w^2\omega_p-2\mu^2w\omega_p\right)+e^{2\epsilon}\left(9w^4+6\mu w^3+\mu^2w^2\right)-4\mu w\omega_p^2+\mu^2\omega_p^2}{12w^2\omega_p^2+e^\epsilon(16w^3\omega_p+16\mu w^2\omega_p+16\mu^2w\omega_p)+e^{2\epsilon}(12w^4+12\mu w^3+12\mu^2w^2)+12\mu w\omega_p^2+12\mu^2\omega_p^2}$$
$$s_{1,3}^2 = \frac{4w^2\omega_p e^{2\gamma}+e^\epsilon\left(12e^\gamma w^3\omega_p e^\gamma-6e^\gamma\mu w^2\omega_p+4\mu w^2\omega_p e^\gamma-2\mu^2w\omega_p\right)+e^{2\epsilon}\left(9e^{2\gamma}w^4+6e^\gamma\mu w^3+\mu^2w^2\right)-4\mu w\omega_p\omega_p e^\gamma+\mu^2\omega_p^2}{12e^{2\gamma}w^2\omega_p^2+e^\epsilon(16e^{2\gamma}w^3\omega_p+16w^2\omega_p\mu e^\gamma+16\mu^2w\omega_p)+e^{2\epsilon}(12e^{2\gamma}w^4+12w^3\mu e^\gamma+12\mu^2w^2)+12w\omega_p^2\mu e^\gamma+12\mu^2\omega_p^2}.$$

In the limit of $\epsilon \to \infty$ we have

$$\lim_{\epsilon\to\infty} s_{1,2}^2 = \frac{\left(9w^4+6\mu w^3+\mu^2w^2\right)}{(12w^4+12\mu w^3+12\mu^2w^2)}$$
$$\lim_{\epsilon\to\infty} s_{1,3}^2 = \frac{\left(9e^{2\gamma}w^4+6e^\gamma\mu w^3+\mu^2w^2\right)}{(12e^{2\gamma}w^4+12w^3\mu e^\gamma+12\mu^2w^2)}.$$

In contrast, as $\epsilon \to -\infty$ we have:

$$\lim_{\epsilon\to-\infty} s_{1,2}^2 = \frac{(\mu-2w)^2}{12(\mu^2+w^2+\mu w)}$$
$$\lim_{\epsilon\to-\infty} s_{1,3}^2 = \frac{(\mu-2we^\gamma)^2}{12(\mu^2+e^{2\gamma}w^2+w\mu e^\gamma)}.$$

To understand the behavior of these expressions, we introduce the ratio variable $\sigma = \frac{w}{\mu}$:

$$\lim_{\epsilon\to\infty} s_{1,2}^2 = \frac{\left(9\sigma^4+6\sigma^3+\sigma^2\right)}{(12\sigma^4+12\sigma^3+12\sigma^2)}$$
$$\lim_{\epsilon\to\infty} s_{1,3}^2 = \frac{\left(9e^{2\gamma}\sigma^4+6e^\gamma\sigma^3+\sigma^2\right)}{(12e^{2\gamma}\sigma^4+12\sigma^3e^\gamma+12\sigma^2)},$$

and:

$$\lim_{\epsilon\to-\infty} s_{1,2}^2 = \frac{(1-2\sigma)^2}{12(1+\sigma^2+\sigma)}$$
$$\lim_{\epsilon\to-\infty} s_{1,3}^2 = \frac{(1-2\sigma e^\gamma)^2}{12(1+e^{2\gamma}\sigma^2+\sigma\mu e^\gamma)}.$$

In the limit of large σ, we have:

$$\lim_{\epsilon\to\infty}\sum s_{i,j}^2 \propto \frac{3}{4} > \lim_{\epsilon\to-\infty}\sum s_{i,j}^2 \propto \frac{1}{3}$$

The increase in $\sum s_{i,j}^2$ is monotonic Again the $s_{2,3}^2$ term is not a function of ϵ, so considering only the terms of type $s_{1,j}^2$

$$\frac{d}{d\epsilon}s_{1,j}^2 = 40e^\epsilon\omega_p\omega\left(m'^2+m'\omega+\omega^2\right)\left[3e^\epsilon\omega_p\omega^2(2m'+\omega)+e^{2\epsilon}m'\omega^2(m'+3\omega) + {\omega_p}^2\left(-m'^2+m'\omega+2\omega^2\right)\right]$$

As expected, these terms are monotonically increasing except when $-m'^2{\omega_p}^2$ dominates all other (positive) terms in the square bracket, which requires m' large, and $\frac{\omega_p}{e^\epsilon \omega} > 1$, far from the proofreading limit. Putting these sums back into the equation for orthogonality we can verify that orthogonality is decreasing as ϵ increases in the proofreading limit ($\sigma \approx 50$)

$$\lim_{\epsilon\to-\infty} \Theta = -0.3394 > \lim_{\epsilon\to\infty} \Theta = -0.8635.$$

c. Energetic Limit 3: $\omega_p \to 0$ Again it is instructive to rearrange $s^2_{i,j}$ to collect the ω_p terms. Again the third term is not a function of ω_p, This gives

$$s^2_{1,2} = \frac{9w^4e^{2\epsilon}+6\mu w^3e^{2\epsilon}+\omega_p^2\left(\mu^2+4w^2-4\mu w\right)+\mu^2w^2e^{2\epsilon}+\omega_p\left(12w^3e^{\epsilon}-2\mu w^2e^{\epsilon}-2\mu^2we^{\epsilon}\right)}{12w^4e^{2\epsilon}+12\mu w^3e^{2\epsilon}+\omega_p^2(12\mu^2+12w^2+12\mu w)+12\mu^2w^2e^{2\epsilon}+\omega_p(16w^3e^{\epsilon}+16\mu w^2e^{\epsilon}+16\mu^2we^{\epsilon})}$$

$$s^2_{1,3} = \frac{9w^4e^{2\gamma+2\epsilon}+6\mu w^3e^{\gamma+2\epsilon}+12w^3e^{\gamma+\epsilon}\omega_pe^{\gamma}+\omega_p\left(-6\mu w^2e^{\gamma+\epsilon}-4\mu w\omega_pe^{\gamma}-2\mu^2we^{\epsilon}\right)+4w^2\omega_pe^{2\gamma}+4\mu w^2e^{\epsilon}\omega_pe^{\gamma}+\mu^2w^2e^{2\epsilon}+\mu^2\omega_p^2}{12w^4e^{2\gamma+2\epsilon}+12w^3e^{2\epsilon}\mu e^{\gamma}+\omega_p^2(12\mu^2+12e^{2\gamma}w^2+12w\mu e^{\gamma})+12\mu^2w^2e^{2\epsilon}+\omega_p(16w^3e^{2\gamma+\epsilon}+16w^2e^{\epsilon}\mu e^{\gamma}+16\mu^2we^{\epsilon})}$$

and we must show that the sums are decreasing in ω_p, i.e.

$$\lim_{\omega_p\to0}\sum s^2_{i,j} > \lim_{\omega_p\to\infty}\sum s^2_{i,j}$$

In the limit of $\omega_p \to 0$ we have

$$\lim_{\omega_p\to0} s^2_{1,2} = \frac{(\mu+3w)^2}{12(\mu^2+w^2+\mu w)}$$
$$\lim_{\omega_p\to0} s^2_{1,3} = \frac{(\mu+3e^{\gamma}w)^2}{12(\mu^2+w(\mu e^{\gamma}+e^{2\gamma}w))},$$

while in the limit of $\omega_p \to \infty$

$$\lim_{\omega_p\to\infty} s^2_{1,2} = \frac{\mu^2+4w^2-4\mu w}{12\mu^2+12w^2+12\mu w}$$
$$\lim_{\omega_p\to\infty} s^2_{1,3} = \frac{(\mu-2e^{\gamma}w)^2}{12(\mu^2+e^{2\gamma}w^2+w\mu e^{\gamma})}.$$

Making the same substitutions as before ($\sigma = w/\mu$) gives:

$$\lim_{\omega_p\to0} s^2_{1,2} = \frac{(1+3\sigma)^2}{12(1+\sigma^2+\sigma)}$$
$$\lim_{\omega_p\to0} s^2_{1,3} = \frac{(1+3e^{\gamma}\sigma)^2}{12(1+\sigma e^{\gamma}+e^{2\gamma}\sigma^2)}$$

and

$$\lim_{\omega_p\to\infty} s^2_{1,2} = \frac{1+4\sigma^2-4\sigma}{12+12\sigma^2+12\sigma}$$
$$\lim_{\omega_p\to\infty} s^2_{1,3} = \frac{(1-2e^{\gamma}\sigma)^2}{12(1+e^{2\gamma}\sigma^2+\sigma e^{\gamma})}.$$

As previously we have the desired result directly:

$$\lim_{\omega_p\to0}\sum s^2_{i,j} \propto \frac{3}{4} > \lim_{\omega_p\to\infty}\sum s^2_{i,j} \propto \frac{1}{3}.$$

The increase in $\sum s^2_{i,j}$ is monotonic We compute

$$\frac{d}{d\omega_p}s^2_{1,j} = -40e^{\epsilon}\omega\left(m'^2+m'\omega+\omega^2\right)\left[3e^{\epsilon}\omega_p\omega^2(2m'+\omega)+e^{2\epsilon}m'\omega^2(m'+3\omega)\right.$$
$$\left.+{\omega_p}^2\left(-m'^2+m'\omega+2\omega^2\right)\right].$$

These terms are monotonically decreasing except for when $-m'^2{\omega_p}^2$ dominates all other (positive) terms

in the square bracket, which requires m' large, and $\frac{\omega_p}{e^{\epsilon}\omega} > 1$.

Putting these sums back into the equation for orthogonality we can verify that orthogonality is increasing as ω_p increases in the proofreading limit ($\sigma \approx 50$)

$$\lim_{\omega_p \to 0} \Theta = -0.8635 < \lim_{\omega_p \to \infty} \Theta = -0.3395$$

d. The Hopfield Network in the Kinetic Regime

Derivation of the kinetic regime error rate We first derive an expression for the error rate in the kinetic regime of the Ninio-Hopfield scheme, ξ_{kinetic}. We then determine the appropriate proofreading limits in the kinetic regime.

We compute that:

$$\begin{aligned}
\xi_{\text{kinetic}} &= \frac{(e^{\epsilon+\epsilon_i}\omega\omega_i + \omega\omega_p + e^{\epsilon_i}\omega_i\omega_p)(e^{2\delta}\omega\omega_i + e^{\delta+\epsilon_p}\omega\omega_p + e^{\epsilon_i+\epsilon_p}\omega_i\omega_p)}{(e^{2\delta+\epsilon+\epsilon_i}\omega\omega_i + e^{\delta}\omega\omega_p + e^{\epsilon_i}\omega_i\omega_p)(\omega\omega_i + e^{\epsilon_p}\omega\omega_p + e^{\epsilon_i+\omega_p}\omega_i\omega_p)} \\
&= \frac{(e^{\epsilon+\epsilon_i}a + b + c)(e^{2\delta}a + e^{\epsilon_p+\delta}b + e^{\epsilon_p}c)}{(e^{2\delta+\epsilon+\epsilon_i}a + e^{\delta}b + c)(a + e^{\epsilon_p}b + e^{\epsilon_p}c)} \\
&= \frac{(e^{2\delta}a + e^{\delta+\epsilon_p}b + e^{\epsilon_p}c)(e^{\epsilon+\epsilon_i}a + b + c)}{(e^{2\delta+\epsilon+\epsilon_i}a + e^{\delta}b + c)(a + e^{\epsilon_p}b + e^{\epsilon_p}c)}
\end{aligned} \tag{E1}$$

where we have let $a = \omega\omega_i,\ b = \omega\omega_p,\ c = \omega_i\omega_p e^{\epsilon_i}$.

When the total dissipation $\epsilon_i + \epsilon_p + \epsilon$ is high, the terms $e^{\epsilon+\epsilon_i}$ in Equation E1 will dominate. We therefore have that

$$\xi_{\text{kinetic}} \approx \frac{e^{2\delta}a + e^{\delta+\epsilon_p}b + e^{\epsilon_p}c}{e^{2\delta}a + e^{2\delta+\epsilon_p}b + e^{2\delta+\epsilon_p}c}$$

from which it is clear that proofreading requires that e^{ϵ_p}/a be very large. Moreover, the error fraction is minimized when c/b is very large. Note that proofreading can still occur when $b/c >> 1$, but the error fraction is not minimized in this regime. Translating these conditions into Kramer's form parameters gives the necessary limits for maximum discrimination

$$e^{\epsilon_p} \to \infty,\ \frac{\omega_i e^{\epsilon_i}}{\omega} \to \infty.$$

As in the energetic regime, we take $m' = \mu = \omega_i e^{\epsilon_i}$, and write the limits as:

$$e^{\epsilon_p} \to \infty,\ \frac{\mu}{\omega} \to \infty \to \infty.$$

We now investigate orthogonality in these discriminatory limits.

Orthogonality is increasing with μ Recall that increasing orthogonality requires $\sum s_{i,j}^2$ decreasing. Lets begin by rewriting the elements of $\sum s_{i,j}^2$ w.r.t μ

$$\begin{aligned}
s_{1,2}^2 &= \frac{\left(\omega e^{\delta+\epsilon}(\mu+2\omega) + e^{-\delta_p}\omega_p(\omega-\mu) + \omega\omega_p + \omega^2 e^{\epsilon}\right)^2}{2(\mu^2+\mu\omega+\omega^2)\left(\left((e^{\delta}+1)\omega e^{\epsilon} + e^{-\delta_p}\omega_p + \omega_p\right)^2 + \omega^2 e^{2(\delta+\epsilon)} + e^{-2\delta_p}\omega_p^2 + \omega_p^2 + \omega^2 e^{2\epsilon}\right)} \\
s_{1,3}^2 &= \frac{\left(\omega\left(\omega\left(-e^{\delta+\epsilon}\right) - e^{-\delta_p}\omega_p - \omega_p - \omega e^{\epsilon}\right) + \mu\omega_p + \omega e^{\epsilon}(-\mu-\omega)\right)^2}{(\mu^2+(\mu+\omega)^2+\omega^2)\left(\left(\omega e^{\delta+\epsilon} + e^{-\delta_p}\omega_p + \omega_p + \omega e^{\epsilon}\right)^2 + \omega^2 e^{2\delta+2\epsilon} + e^{-2\delta_p}\omega_p^2 + \omega_p^2 + \omega^2 e^{2\epsilon}\right)} \\
s_{2,3}^2 &= \frac{\omega^4 e^{2\delta}}{4(\mu^2+\mu\omega+\omega^2)(\omega^2 e^{2\delta} + \mu\omega e^{\delta} + \mu^2)}.
\end{aligned}$$

Because $s_{2,3}^2$ has μ in the denominator but not in the numerator it must go to zero as $\mu \to \infty$. The expressions for the remaining $s_{1,i}^2$ terms are,

$$\lim_{\mu\to 0} s_{1,2}^2 = \frac{\left(2\omega e^{\delta+\delta_p+\epsilon}+e^{\delta_p}\omega_p+\omega e^{\delta_p+\epsilon}+\omega_p\right)^2}{4\left(\omega^2 e^{2(\delta+\delta_p+\epsilon)}+\omega^2 e^{\delta+2(\delta_p+\epsilon)}+\omega\omega_p e^{\delta+\delta_p+\epsilon}+\omega\omega_p e^{\delta+2\delta_p+\epsilon}+e^{\delta_p}\omega_p^2+e^{2\delta_p}\omega_p^2+\omega^2 e^{2(\delta_p+\epsilon)}+\omega\omega_p e^{\delta_p+\epsilon}+\omega\omega_p e^{2\delta_p+\epsilon}+\omega_p^2\right)}$$

$$\lim_{\mu\to 0} s_{1,3}^2 = \frac{\left(\omega e^{\delta+\delta_p+\epsilon}+e^{\delta_p}\omega_p+2\omega e^{\delta_p+\epsilon}+\omega_p\right)^2}{4\left(\omega^2 e^{2(\delta+\delta_p+\epsilon)}+\omega^2 e^{\delta+2(\delta_p+\epsilon)}+\omega\omega_p e^{\delta+\delta_p+\epsilon}+\omega\omega_p e^{\delta+2\delta_p+\epsilon}+e^{\delta_p}\omega_p^2+e^{2\delta_p}\omega_p^2+\omega^2 e^{2(\delta_p+\epsilon)}+\omega\omega_p e^{\delta_p+\epsilon}+\omega\omega_p e^{2\delta_p+\epsilon}+\omega_p^2\right)}$$

and

$$\lim_{\mu\to\infty} s_{1,2}^2 = \frac{\left(\omega_p-\omega e^{\delta+\delta_p+\epsilon}\right)^2}{4\left(\omega^2 e^{2(\delta+\delta_p+\epsilon)}+\omega^2 e^{\delta+2(\delta_p+\epsilon)}+\omega\omega_p e^{\delta+\delta_p+\epsilon}+\omega\omega_p e^{\delta+2\delta_p+\epsilon}+e^{\delta_p}\omega_p^2+e^{2\delta_p}\omega_p^2+\omega^2 e^{2(\delta_p+\epsilon)}+\omega\omega_p e^{\delta_p+\epsilon}+\omega\omega_p e^{2\delta_p+\epsilon}+\omega_p^2\right)}$$

$$\lim_{\mu\to\infty} s_{1,3}^2 = \frac{e^{2\delta_p}\left(\omega_p-\omega e^{\epsilon}\right)^2}{4\left(\omega^2 e^{2(\delta+\delta_p+\epsilon)}+\omega^2 e^{\delta+2(\delta_p+\epsilon)}+\omega\omega_p e^{\delta+\delta_p+\epsilon}+\omega\omega_p e^{\delta+2\delta_p+\epsilon}+e^{\delta_p}\omega_p^2+e^{2\delta_p}\omega_p^2+\omega^2 e^{2(\delta_p+\epsilon)}+\omega\omega_p e^{\delta_p+\epsilon}+\omega\omega_p e^{2\delta_p+\epsilon}+\omega_p^2\right)}.$$

Here we will again make a ratio substitution, $\tau = \omega e^{\epsilon}/\omega_p$ and send $\tau \to \infty$ and required in the kinetic discriminatory regime. In this limit, we have:

$$\lim_{\mu\to 0} s_{1,2}^2 = \frac{1+4e^{\delta}+4e^{2\delta}}{4+4e^{\delta}+4e^{2\delta}}$$

$$\lim_{\mu\to 0} s_{1,3}^2 = \frac{4+4e^{\delta}+e^{2\delta}}{4+4e^{\delta}+4e^{2\delta}}$$

and

$$\lim_{\mu\to\infty} s_{1,2}^2 = \frac{e^{2\delta}}{4+4e^{\delta}+4e^{2\delta}}$$

$$\lim_{\mu\to\infty} s_{1,3}^2 = \frac{1}{4+4e^{\delta}+4e^{2\delta}}.$$

This gives the desired result,

$$\lim_{\mu\to 0}\sum s_{i,j}^2 > \lim_{\mu\to\infty}\sum s_{i,j}^2.$$

Putting these sums back into the equation for orthogonality we can verify that orthogonality is increasing as μ increases in the proofreading limit ($\tau \approx 10^4$):

$$\lim_{\mu\to 0}\Theta = -0.819 < \lim_{\mu\to\infty}\Theta = 0.3612.$$

e. The increase in $\sum s_{i,j}^2$ is monotonic Again it is easiest to start with the $s_{2,3}^2$ term. An application of the quotient rule reveals that $\frac{d}{d\mu}s_{2,3}^2 < 0$. The remaining derivatives are given by

$$\frac{d}{d\mu}s_{1,2}^2 = -\frac{\omega\left(\omega(\mu+2\omega)e^{\delta+\delta_p+\epsilon}+e^{\delta_p}\omega\omega_p+\omega^2 e^{\delta_p+\epsilon}+\omega_p(\omega-\mu)\right)\left(3\mu\omega e^{\delta+\delta_p+\epsilon}+e^{\delta_p}\omega_p(2\mu+\omega)+\omega e^{\delta_p+\epsilon}(2\mu+\omega)+3\omega_p(\mu+\omega)\right)}{4(\mu^2+\mu\omega+\omega^2)^2\left(\omega^2 e^{2(\delta+\delta_p+\epsilon)}+\omega^2 e^{\delta+2(\delta_p+\epsilon)}+\omega\omega_p e^{\delta+\delta_p+\epsilon}+\omega\omega_p e^{\delta+2\delta_p+\epsilon}+e^{\delta_p}\omega_p^2+e^{2\delta_p}\omega_p^2+\omega^2 e^{2(\delta_p+\epsilon)}+\omega\omega_p e^{\delta_p+\epsilon}+\omega\omega_p e^{2\delta_p+\epsilon}+\omega_p^2\right)}$$

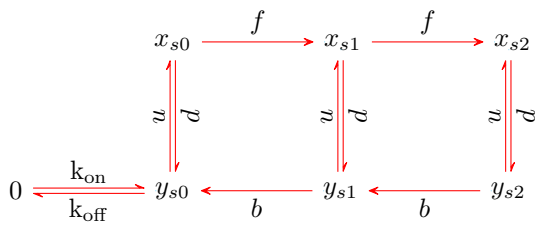
$$\frac{d}{d\mu}s_{1,3}^2 = -\frac{\omega\left(\omega(2\mu+\omega)e^{\delta+\delta_p+\epsilon}+3e^{\delta_p}\omega_p(\mu+\omega)+3\mu\omega e^{\delta_p+\epsilon}+\omega_p(2\mu+\omega)\right)\left(\omega e^{\delta_p+\epsilon}\left(\left(e^{\delta}+2\right)\omega+\mu\right)+\omega_p\left(e^{\delta_p}(\omega-\mu)+\omega\right)\right)}{4(\mu^2+\mu\omega+\omega^2)^2\left(\omega^2 e^{2(\delta+\delta_p+\epsilon)}+\omega e^{\delta+\delta_p+\epsilon}\left(e^{\delta_p}(\omega_p+\omega e^{\epsilon})+\omega_p\right)+e^{\delta_p}\omega_p^2+e^{2\delta_p}\omega_p^2+\omega^2 e^{2(\delta_p+\epsilon)}+\omega\omega_p e^{\delta_p+\epsilon}+\omega\omega_p e^{2\delta_p+\epsilon}+\omega_p^2\right)}$$

Which are both strictly negative. We conclude that $\sum s_{i,j}^2$ is a monotonically decreasing function of μ, thus orthogonality is monotonically increasing.

Appendix F: Expressions for Error Rate in the Ladder Graph

We wish to derive expressions for the error rate of the ladder discrimination scheme in the kinetic and energetic regimes.

A single side of the ladder has structure:



where we have dropped the superscripts d^S, u^S for

clarity.

We will use the Matrix-Tree theorem (MTT), which provides an expression for steady states in terms of *spanning trees* [28]. Recall that a *spanning tree* of a graph G is is a subgraph which includes every vertex of G and has no cycles (when edge directions ignored). A spanning tree is said to be *rooted* at node i if node i is the only vertex of the subgraph without any outgoing edges.

The MTT provides an expression for the steady state of node i in terms of the sum of the product of the rates of each spanning tree rooted at i. That is:

$$\rho_i = \sum_{T \in S_i(G)} \left(\prod_{j \xrightarrow{\alpha} k \in T} a \right), \tag{F1}$$

where $S_i(G)$ is the set of all spanning trees of graph G rooted at i.

We will exploit the structure of our ladder network in order to simplify this expression. Our ladder consists of two subgraphs joined at a single node, 0: $G = G_W \oplus_0 G_R$ (G_R, G_W, corresponding to subgraphs for the right, wrong products, respectively). Equation F1 implies that

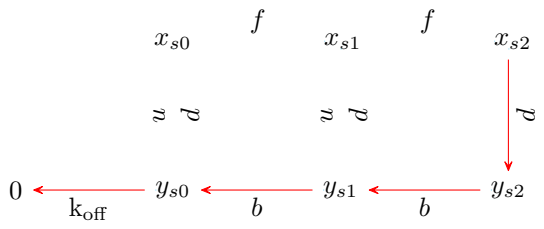
$$\rho_i(G_W \oplus_0 G_R) = \begin{cases} \rho_i(G_W)\rho_0(G_R), & \text{if } i \in G_W \\ \rho_i(G_R)\rho_0(G_W), & \text{if } i \in G_R. \end{cases}$$

This gives for the error

$$\xi = \frac{\rho_W}{\rho_R} = \frac{\rho_W(G_W)\rho_0(G_R)}{\rho_R(G_R)\rho_0(G_W)}, \tag{F2}$$

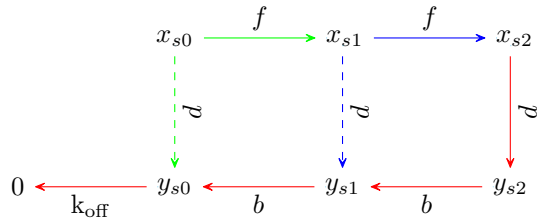
where ρ_W, ρ_R represent the kernel elements corresponding to the discriminatory nodes in the upper corner of the wrong, right subgraphs (respectively).

We therefore need only determine analytical expressions for the sums of (products of rate constants of) spanning trees rooted at the top corner and 0 nodes. Let's count the trees rooted at $\rho_0(G_S)$ first. In order for the tree to be rooted at 0, there are a number of essential arrows:



without any of which it is impossible to produce a spanning tree rooted at 0. The necessity of these arrows comes from the unidirectionality of the f, b.

What other arrows are necessary for a spanning tree? Consider the diagram



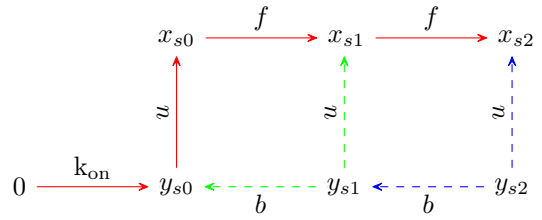
It is necessary and sufficient for a spanning tree rooted at 0 to contain all of the red arrows, and exactly one of the green arrows and one of the blue arrows. This holds in general; each loop in a ladder must contribute either a factor of f or d to a spanning tree rooted at 0.

We can thus compute:

$$\begin{aligned} \rho_0(G_S) &= k_{\text{off}} b^\alpha d \sum_{k=0}^{\alpha} \binom{\alpha}{k} f^{\alpha-k} d^k \\ &= k_{\text{off}} b^\alpha d (f+d)^\alpha \end{aligned}$$

where the second line follows from the Binomial theorem, and where we have set the number of square loops in the ladder portion of the graph to be α.

We can now repeat this procedure with spanning trees rooted in the upper corner, with red, blue, and green as before:



Which gives us:

$$\begin{aligned} \rho_S(G_S) &= k_{\text{on}} f^\alpha u \sum_{k=0}^{\alpha} \binom{\alpha}{k} b^{\alpha-k} u^k \\ &= k_{\text{on}} f^\alpha u (b+u)^\alpha. \end{aligned}$$

Note that in comparison to the last expression, we have merely made the substitutions: $b \to f$, $f \to b$, $d \to u$, $u \to d$. Plus $k_{\text{off}} \to k_{\text{on}}$, of course.

Returning to our expression for the error gives

$$\begin{aligned} \xi &= \frac{\rho_W}{\rho_R} = \frac{\rho_W(G_W)\rho_0(G_R)}{\rho_R(G_R)\rho_0(G_W)} \\ &= \frac{k_{\text{on}} k_{\text{off}}\ f_W^\alpha\ u_W\ b_R^\alpha\ d_R\ (u_W + b_W)^\alpha (f_R + d_R)^\alpha}{k_{\text{on}} k_{\text{off}}\ f_R^\alpha\ u_R\ b_W^\alpha\ d_W\ (u_R + b_R)^\alpha (f_W + d_W)^\alpha} \end{aligned}$$

where we have denoted variables coming from the 'right' and 'wrong' sides of the ladder with subscripts R and W, respectively. We can do some cancellation (b $= b_R = b_W$ and $f = f_R = f_W$) to arrive at:

$$\xi = \frac{d_R u_W (u_W + b)^\alpha (f + d_R)^\alpha}{d_W u_R (u_R + b)^\alpha (f + d_W)^\alpha}.$$

In the energetic regime we have that $u_R = u_W$, and that $d_W = d_R e^\gamma$:

$$\xi_{\text{energetic}} = \frac{(f + d_R)^\alpha}{e^\gamma (f + d_R e^\gamma)^\alpha}.$$

In the kinetic regime, we have that $d_R = d_W e^\delta$, $u_R = u_W e^\delta$, giving

$$\xi_{\text{kinetic}} = \frac{(u + b)^\alpha (f + d e^\delta)^\alpha}{(u e^\delta + b)^\alpha (f + d)^\alpha}.$$

Appendix G: Orthogonality and Error in the Ladder Graph

We first derive the discriminatory limit in the energetic regime. Recall that

1. Energetic regime

$$\xi_{\text{energetic}} = \frac{(f + d_R)^\alpha}{e^\gamma (f + d_R e^\gamma)^\alpha}.$$

The substitution $\eta = \frac{d_R}{f}$ gives

$$\xi_{\text{energetic}} = \frac{(1 + \eta)^\alpha}{e^\gamma (1 + \eta e^\gamma)^\alpha}$$

from which read off that proofreading requires η to be large. This corresponds to the intuition that the rate of discards must be large with respect the reaction speed.

We must now demonstrate that orthogonality is decreasing as η becomes large.

As in the Ninio-Hopfield case, we will use the notation $\sum s_{i,j}^2$ to denote the squared, normalized inner product between columns i, j in Matrix $\mathcal{L}^{a,b}$ formed by deleting the columns corresponding to the discriminatory nodes a, b from the full Laplacian for this graph.

For any given loop of the ladder, these terms are given by

$$\langle x_{si}, y_{s(i+1)} \rangle^2 = \frac{(bd + fu)^2}{4\,(b^2 + bu + u^2)\,(d^2 + df + f^2)}$$
$$\langle y_{si}, y_{s(i+1)} \rangle^2 = \frac{b^2 (b + u)^2}{4\,(b^2 + bu + u^2)^2}$$
$$\langle x_{si}, x_{s(i+1)} \rangle^2 = \frac{f^2 (d + f)^2}{4\,(d^2 + df + f^2)^2}$$
$$\langle x_{si}, y_{si} \rangle^2 = \frac{(bd + 2du + fu)^2}{4\,(b^2 + bu + u^2)\,(d^2 + df + f^2)}.$$

For N loops, there will be N of each of these terms except for $\langle x_{si}, x_{s(i+1)} \rangle^2$ for which there will be $(N-1)$ for each side of the ladder. In addition, there will be two terms that originate from the reactant node (note in this case we are considering a slightly altered graph, where $k_{off} = b$ and $k_{on} = f$, and k_{on} connects 0 to x_{s0}). These are given as

$$\langle 0, x_{s0} \rangle^2 = \frac{(2b - u)^2}{12\,(b^2 + bu + u^2)}$$
$$\langle 0, y_{s0} \rangle^2 = \frac{(d + f)^2}{6\,(d^2 + (d + f)^2 + f^2)}.$$

Recall that in the energetic regime, our effective parameter of interest if $\eta = d/f$, noting that $\langle y_{si}, y_{s(i+1)}$ and $\langle 0, x_{s0} \rangle^2$ are not functions of η and making this substitution along with the another substitute $\phi = u/b$ gives

$$\langle x_{si}, y_{s(i+1)} \rangle^2 = \frac{(\eta + \phi)^2}{4\,(1 + \phi + \phi^2)\,(1 + \eta + \eta^2)}$$
$$\langle x_{si}, x_{s(i+1)} \rangle^2 = \frac{(\eta + 1)^2}{4\,(1 + \eta + \eta^2)^2}$$
$$\langle x_{si}, y_{si} \rangle^2 = \frac{(\eta + 2\eta\phi + \phi)^2}{4\,(1 + \phi + \phi^2)\,(1 + \eta + \eta^2)}$$
$$\langle 0, y_{s0} \rangle^2 = \frac{(\eta + 1)^2}{12\,(1 + \eta + \eta^2)}$$

We will set $\phi \to 0$ for convenience. In this limit we have:

$$\langle x_{si}, y_{s(i+1)} \rangle^2 = \frac{\eta^2}{4\,(1 + \eta + \eta^2)}$$
$$\langle x_{si}, x_{s(i+1)} \rangle^2 = \frac{(\eta + 1)^2}{4\,(1 + \eta + \eta^2)^2}$$
$$\langle x_{si}, y_{si} \rangle^2 = \frac{\eta^2}{4\,(1 + \eta + \eta^2)}$$
$$\langle 0, y_{s0} \rangle^2 = \frac{(\eta + 1)^2}{12\,(1 + \eta + \eta^2)}$$

which take values 0, 1/4, 0, and 1/12 in the limit $\eta \to 0$ and 1/4, 0, 1/4, and 1/12 in the limit $\eta \to \infty$ For N loops, we will have N terms of the first and third type, and $N-1$ terms of the second type. The last term is unchanged in these limits. This gives the desired result,

$$\lim_{\eta\to 0} \sum s_{i,j}^2 \propto \frac{N-1}{4} < \lim_{\eta\to\infty} \sum s_{i,j}^2 \propto \frac{2N}{4}.$$

Finally, we consider the case when $\phi \to \infty$. Note that $\langle x_{si}, x_{s(i+1)}\rangle^2$ terms are not functions of ϕ. The two remaining terms to consider are,

$$\langle x_{si}, y_{s(i+1)}\rangle^2 = \frac{(\eta+\phi)^2}{4\left(1+\phi+\phi^2\right)\left(1+\eta+\eta^2\right)}$$
$$\langle x_{si}, y_{si}\rangle^2 = \frac{(\eta+2\eta\phi+\phi)^2}{4\left(1+\phi+\phi^2\right)\left(1+\eta+\eta^2\right)}$$

which in the $\phi \to \infty$ limit become,

$$\langle x_{si}, y_{s(i+1)}\rangle^2 = \frac{1}{4\eta^2+4\eta+4}$$
$$\langle x_{si}, y_{si}\rangle^2 = \frac{4\eta^2+4\eta+1}{4\eta^2+4\eta+4}$$

Combining these term yields,

$$\lim_{\eta\to 0} \sum s_{i,j}^2 \propto \frac{2N-1}{4} < \lim_{\eta\to\infty} \sum s_{i,j}^2 \propto \frac{4N}{4}.$$

We can directly compute that $\sum s_{i,j}^2$ is monotonically increasing in both the $\phi \to 0$ and $\phi \to \infty$ limits.

a. *ϕ does not affect orthogonality in the $f \ll d$ limit*

Before turning to the kinetic regime, we demonstrate that ϕ does not affect orthogonality in the energetic discrimination limit.

We examine the elements $s_{i,j}^2$ that depend on ϕ in the $\eta \to \infty$ discriminatory limit. Before taking the limit, we have

$$\langle x_{si}, y_{s(i+1)}\rangle^2 = \frac{(\eta+\phi)^2}{4\left(1+\phi+\phi^2\right)\left(1+\eta+\eta^2\right)}$$
$$\langle x_{si}, y_{si}\rangle^2 = \frac{(\eta+2\eta\phi+\phi)^2}{4\left(1+\phi+\phi^2\right)\left(1+\eta+\eta^2\right)}$$
$$\langle y_{si}, y_{s(i+1)}\rangle^2 = \frac{(\phi+1)^2}{4\left(\phi^2+\phi+1\right)^2}.$$

In the $\eta \to \infty$ limit these become

$$\langle x_{si}, y_{s(i+1)}\rangle = \sqrt{\frac{1}{4\phi^2+4\phi+4}}$$
$$\langle x_{si}, y_{si}\rangle = \sqrt{\frac{4\phi^2+4\phi+1}{4\phi^2+4\phi+4}}$$
$$\langle y_{si}, y_{s(i+1)}\rangle = \frac{(\phi+1)}{2\left(\phi^2+\phi+1\right)}.$$

Now we must evaluate these in the limits of $\phi \to 0$ and $\phi \to \infty$, the first term goes from 1/2 to 0 as $\phi \to \infty$. The second term goes from 1/2 to 1 and the third term goes from 1/2 to 0. Because each loop consists of two of the second type term and one each of the first and third type term, the sum is the same in each limit.

In the full expression for orthogonality, we do observe a small non-constant dependence on ϕ, but this is marginal and strictly decreases the orthogonality, thereby reinforcing our notion that ϕ cannot be used to increased realizable pathways in the low f regime.

2. Kinetic regime

We now need to demonstrate that

$$\xi_{\text{kinetic}} = \frac{(u+b)^\alpha (f+de^\delta)^\alpha}{(ue^\delta+b)^\alpha (f+d)^\alpha}.$$

Define $\eta = d/f,\ \phi = u/b$ as before.

$$\xi_{\text{kinetic}} = \frac{(\phi+1)^\alpha (1+\eta e^\delta)^\alpha}{(\phi e^\delta+1)^\alpha (\eta+1)^\alpha}.$$

which attains its minimum of $e^{-\alpha\delta}$ in the limit $\phi \to \infty,\ \eta \to 0$. The previous sections demonstrated that orthogonality is increasing in these limits.

Appendix H: Supplemental Information

Figure S1 demonstrates that orthogonality tends to increase as we add connections of equal order of magnitude to a graph. The figure was generated by first creating an all-to-all connected graph having 16 nodes. Rate constants were chosen randomly from the distribution $Exp[\mathcal{N}(0, \frac{1}{3})]$. For each of the c = 'connectivity fractions' in Figure S1, a random set of $1 - c * (16^2 - 16)/2$ connections was then chosen for deletion and removed bidirectionally. These random deletion sets were chosen 1000 times for each connectivity fraction considered. The mean and standard deviation of these 1000 samples is plotted. Graph sparsity 0 corresponds to a 16×16 grid graph.

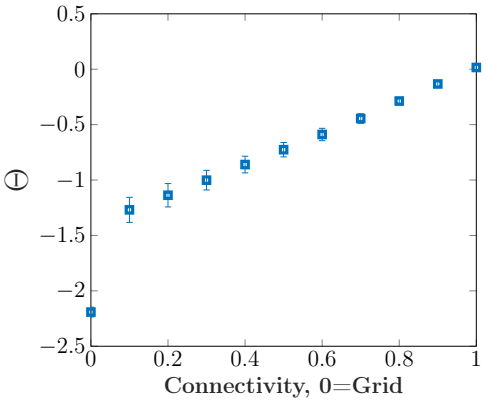


FIG. S1: As a graph becomes more connected, orthogonality increases. Orthogonality is plotted against varying connectivities of a 16 node graph, generated as described in the main text of this section. Zero connectivity corresponds to a 16×16 grid graph, by convention.

Figure S2 shows the relationship between orthogonality and error for the Ninio-Hopfield model in the kinetic regime ($\gamma = 0, \delta = \delta_p = 1$). High orthogonality and high dissipation are necessary for low error.

1. Tables of parameter values

Table I gives the values of the rate constants in the irreversible style ladder graph model, and used to generate the plots in Figure 4 Table II gives the values used to derive Kramer's form rate constants for the reversible ladder graph and to generate the plots shown in Figure 4. Table III gives the val-

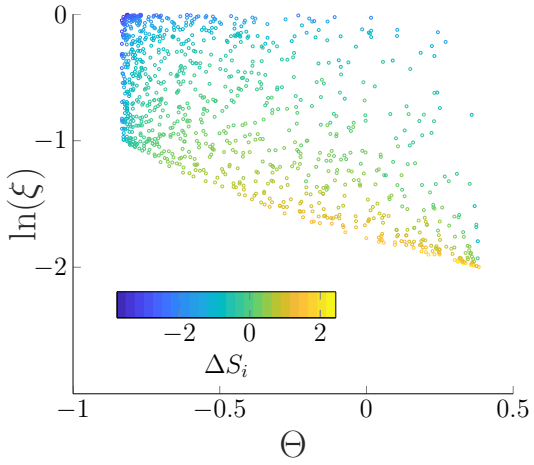


FIG. S2: Orthogonality is required to achieve the minimum error rate in the kinetic regime ($\gamma = 0, \delta = 1$). The log of the error rate ($\ln(\xi)$) as a function of the orthogonality (Θ) is plotted for simulations of the triangle graph (Hopfield-Ninio) with Kramer's form rate constants for 1,000 randomly chosen values of ω_i, ω_p, ϵ_i, and ϵ_p. Other parameters were fixed ($\omega = 1$, $\epsilon = 10$). Color shows the dissipation ΔS_i at steady state.

Parameter	Fig4(b) Energetic	Fig4(b) Kinetic
f	0.1	2
b	2	0.1
u	0.1	3
d	$f(x)$	$f(x)$
γ	1	0
δ	0	1

TABLE I: Parameters used to generate different figures for the irreversible ladder graph. $f(x)$ indicate that this parameter was used as an independent variable for plotting.

Parameter	Fig4(c)	Fig4(d)
ω_f	$f(x)$	0.0874
ω_b	2.3565	2.3565
ω_d	15.33	$f(x)$
ϵ_f	3	3
ϵ_b	3	3
ϵ_u	3	3
γ	1	1
δ	0	0

TABLE II: Parameters used to generate different figures for the reversible ladder graph. $f(x)$ indicate that this parameter was used as an independent variable for plotting.

Parameter	Fig2(b)	Fig2(c)	Fig2(d) Energetic	Fig2(d) Kinetic	Fig S1
ω	1	1	1	1	1
ϵ	10	10	10	10	10
γ	1	1	1	0	0
δ	0	0	0	1	1
δ_p	0	0	0	1	1
ω_i	$Exp[\mathcal{N}(0, \frac{1}{2})]$	0.55	0.55	2.27	$Exp[\mathcal{N}(0, \frac{1}{2})]$
ϵ_i	$\mathcal{N}(0, 2)$	$f(x)$	$f(x)$	$f(x)$	$\mathcal{N}(0, 2)$
ω_p	$Exp[\mathcal{N}(0, \frac{1}{2})]$	0.7318	0.7318	0.8982	$Exp[\mathcal{N}(0, \frac{1}{2})]$
ϵ_p	$\mathcal{N}(0, 2)$	4.2245	4.2245	2.5553	$\mathcal{N}(0, 2)$

TABLE III: Parameters used to generate different figures for the Hopfield-Ninio Model. $f(x)$ indicate that this parameter was used as a dependent variable for plotting.

ues used to derive Kramer's form rate constants for the Hopfield-Ninio model and to generate the plots shown in Figure 2.

MATLAB code used to generate the figures can be found at
https://github.com/davex0r/GeometricRequirementsDiscrimination
LaTeX code for the manuscript can be found and text revisions can be tracked at
https://github.com/davex0r/GeometricDiscriminationManuscript