

Cycle structure, capacity, and resource costs of phosphorylation memory

Abstract

How many recoverable states can a substrate with n modification sites support when all sites share one kinase and one phosphatase? We study the distributive mass-action network on the n -cube of phosphoforms in which every enzyme–edge interaction has its own bound complex and all bound substrate is retained in the conserved inventory. Let C_n be the largest number of locally asymptotically stable positive equilibria that one list of rate constants can have in one compatibility class. We prove $\max\{n, \lfloor (2^{n-2} + n)/2 \rfloor\} \leq C_n \leq 2^n$, so that $C_n = \Theta(2^n)$; the upper bound counts nonhyperbolic attractors, and the lower bound is attained by hyperbolic sinks. In contrast, when the effective modification ratios are balanced around every elementary square of the cube, the capacity is exactly n , the value for a single sequential chain; more than n stable states therefore require a nonzero effective cycle affinity. The exponential construction combines an exponentially large loading-interpolation space with an independently normally attracting saturated family. We then separate this structural capacity from operation. A finite-molecule theorem for the full continuous-time Markov chain converts any finite set of hyperbolic sinks into labels with explicit retention, recovery and readout guarantees at a computable sufficient system size; a local large-deviation theorem for crowded states gives an exit barrier of order δ^{2n} , with recovery time of order $\delta^{-(2n-2)}$ at separation δ ; and a reversible extension with one common ATP-hydrolysis potential shows that finitely many sinks persist under finite chemical driving. Exact certificates for a three-site source with four sinks and seven equilibria illustrate the theory. For this example the sufficient certificates remain astronomically expensive, and we prove a concentration-ratio obstruction for its design family; these are diagnoses of the constructions, not lower bounds for phosphorylation memory in general. The proofs are conventional mathematics, supported by exact rational and interval computations and by a small amount of machine-checked algebra.

1 Introduction

1.1 Background and question

Reversible phosphorylation at several sites is among the most common mechanisms by which cells regulate protein function [7, 47]. Since Lisman’s proposal that a bistable kinase–phosphatase system could store information for times far longer than the lifetime of its molecules [14, 39], multisite modification has also been discussed as a chemical memory, and the stochastic lifetime of such switches has been studied in detail for particular models [4, 42]. The mathematical starting point is multistationarity. Markevich, Hoek and Kholodenko observed that a two-site substrate acted on distributively by one kinase and one phosphatase can be bistable although its only feedback is competition for the shared enzymes [41]. For the *sequential* mechanism, in which the sites are modified in a fixed order, Wang and Sontag proved that there are at most $2n - 1$ positive equilibria [56]; examples with $2n - 1$ equilibria are known for $n \leq 4$ [25]; and Feliu, Rendall and Wiuf proved that $\lfloor n/2 \rfloor + 1$ of the equilibria can be asymptotically stable [22]. A companion paper [8]

Table 1: Four quantities that should not share the word “capacity” without qualification.

Quantity	Meaning
Positive-equilibrium count	Number of positive steady states, stable or unstable, in one compatibility class
Stable-equilibrium capacity C_n	Largest number of locally asymptotically stable equilibria for one rate list and one class (Definition 1.1)
Operational label count	Number of states meeting a common preparation, recovery, retention and observation contract for a finite number of molecules (Section 5)
Information capacity	A channel-theoretic quantity, which needs specified inputs, observations and an error law; not determined here

shows that the sequential chain has at most n locally asymptotically stable equilibria and that n hyperbolic sinks, together with $n - 1$ saddles, occur for every n . Further results on parameter regions, bistability and dynamics are in [27, 11, 23, 33, 12, 29, 10, 5, 20, 21, 55], and surveys in [13, 17, 18].

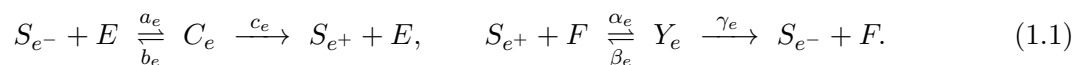
A substrate with n sites has 2^n phosphoforms, not $n + 1$, and when the order of modification is not enforced all of them participate [28, 50]. Thomson and Gunawardena showed that the equilibria of such post-translational modification systems are still described by rational functions of the free-enzyme ratio [53, 40], and their numerical experiments suggested that the number of stable phosphoform distributions grows with n [54]. This raises the question studied here.

For the unordered distributive mechanism with one kinase and one phosphatase, how many stable equilibria can coexist for one set of rate constants and one set of conserved totals, which kinetic structure is responsible for the number, and what does it cost to use these equilibria as the states of a finite-molecule memory?

The last part of the question matters because four quantities are easily conflated (Table 1). An equilibrium can be stable and yet restore disturbances so slowly, or lie so close to its neighbours, that it is useless as a stored state; and a rigorous probability bound for a finite number of molecules can be far more expensive than the underlying dynamics. Throughout this paper the word *capacity* refers to the second row of Table 1 unless it is qualified.

1.2 The model

Vertices of the n -cube are subsets $I \subseteq [n] = \{1, \dots, n\}$, and S_I denotes the phosphoform whose phosphorylated sites are exactly those in I . An *edge* is a pair $e = (e^-, e^+)$ with $e^+ = e^- \cup \{i\}$, $i \notin e^-$; we call e^- the tail and e^+ the head. There are $q = 2^n$ vertices and $m = n2^{n-1}$ edges. For every edge the kinase E and the phosphatase F act by the distributive mass-action mechanism



All $6m = 3n2^n$ rate constants are positive. The two enzymes are shared by all edges, so the binding reactions compete for the free enzymes, but the complexes C_e, Y_e are edge-specific. A mechanism in which one bound complex can release several different products is a different network, and our results do not automatically transfer to it. There are $q + 2m + 2 = (n + 1)2^n + 2$ species and three conserved totals,

$$E_T = E + \sum_e C_e, \quad F_T = F + \sum_e Y_e, \quad S_T = \sum_I S_I + \sum_e (C_e + Y_e). \quad (1.2)$$

The bound substrate is part of S_T ; no quasi-steady-state or low-enzyme approximation is made anywhere in the paper. The set of nonnegative concentrations with prescribed positive totals is a *compatibility class*. It is a compact convex polytope of dimension $d = q + 2m - 1$, invariant under the dynamics, and every statement about stability refers to the dynamics within a class. For $n = 3$ there are 34 species, 12 edges, 72 rate constants and $d = 31$. The *sequential chain* is the analogous network on a path $S_0, \dots, S_n$ with n edges [56, 8].

Definition 1.1. A *sink* is a positive equilibrium that is locally asymptotically stable relative to its compatibility class; it need not be hyperbolic. A *hyperbolic sink* is one at which the Jacobian restricted to the class is Hurwitz. The *stable-equilibrium capacity* C_n is the supremum, over all positive rate constants and all positive totals, of the number of sinks in the class. Equilibria of different rate lists or different classes are never combined.

1.3 Results

Put $p_e = a_e/(b_e + c_e)$, $q_e = \alpha_e/(\beta_e + \gamma_e)$ (the *loadings*) and $k_e^+ = c_e p_e$, $k_e^- = \gamma_e q_e$ (the *effective rates*), and let $\rho_e = k_e^+/k_e^-$. The rates are *square balanced* if

$$\rho(I, I \cup \{i\}) \rho(I \cup \{i\}, I \cup \{i, j\}) = \rho(I, I \cup \{j\}) \rho(I \cup \{j\}, I \cup \{i, j\}) \quad (1.3)$$

for every I and every pair $i \neq j$ outside I ; the logarithm of the quotient of the two sides is the *effective affinity* of that square.

Theorem 1.2 (Architecture). *For every $n \geq 2$,*

$$\max\left\{n, \left\lfloor \frac{2^{n-2} + n}{2} \right\rfloor\right\} \leq C_n \leq 2^n,$$

and $C_3 \geq 4$. Consequently $2^{n-3} \leq C_n \leq 2^n$ for $n \geq 3$, and $C_n = \Theta(2^n)$. The upper bound holds for sinks in the sense of Definition 1.1, including nonhyperbolic ones; every lower bound is attained by hyperbolic sinks of (1.1) with one common list of positive rational rate constants and one triple of positive rational totals.

Theorem 1.3 (Balanced ceiling). *If the rates are square balanced, every compatibility class contains at most $2n - 1$ positive equilibria and at most n sinks, and there are square-balanced rates and totals with exactly n hyperbolic sinks. Hence more than n sinks require a nonzero effective affinity on some square.*

Theorem 1.3 says that a balanced cube has the capacity of a single sequential chain, however large its 2^n -vertex state space is. Balance is weaker than permutation symmetry of the sites: it constrains only equilibrium geometry. Under the stronger symmetric choice of rates the cube reduces *exactly* to a chain, both as an ordinary differential equation and as a Markov chain on molecule counts (Propositions 3.2 and 5.1). The converse of the last sentence of Theorem 1.3 is false: an unbalanced square is not sufficient for additional sinks, and an effective affinity is a property of competing kinetic pathways, not a chemical free energy (Section 7).

The remaining results concern operation and cost; their scopes are summarized in Table 2.

- *Finite-molecule memory* (Theorem 5.2). For the full Markov chain of (1.1) at system size Ω , a rational quadratic Lyapunov function at each hyperbolic sink yields explicit bounds for the probability of leaving a neighbourhood before time t and for failing to return to a preparation set, uniformly over declared sets of initial counts, together with a decoder for a bounded-error readout and nonempty sets of actual integer preparations. The bounds are sufficient, not necessary.

Table 2: Results, type of evidence, and essential scope. “Exact replay” refers to the rational and interval computations described in Appendix C.

Result	Evidence	Scope and limits
$C_n \leq 2^n$ (Thm. 2.6)	Proof: elimination, Sard, degree	All positive rates and totals; nonhyperbolic sinks included; not claimed sharp
Balanced capacity n (Thm. 1.3)	Proof; attainment from [8]	Balance of effective ratios only; equilibrium statement, no dynamical reduction
Exact symmetric reduction (Props. 3.2, 5.1)	Proof: ODE and generator identities	Symmetric lifted rates; observations that factor through the aggregation
$C_n \geq \lfloor (2^{n-2} + n)/2 \rfloor$ (Thm. 4.4)	Proof: interpolation rank, normal hyperbolicity, singular perturbation	Fixed n ; no bound on rates, conditioning or resources uniform in n
Finite-molecule memory (Thm. 5.2)	Proof: exact generator identity	Sufficient bounds; local disturbance sets; preparation, not writing
Crowded-state barrier (Thm. 6.1)	Proof: foliation, Hamilton–Jacobi subsolutions, change of measure	One family; declared local tube; iterated limit $\Omega \rightarrow \infty$, then $\delta \rightarrow 0$
Finite driving (Thm. 7.1, Prop. 7.3)	Proof; explicit threshold	Chemostatted extension; sufficient, very conservative threshold
Four sinks, seven equilibria (Prop. 8.1)	Exact replay	One rational three-site source and class
Operating contract and Table 4 (Prop. 8.2)	Exact replay	Sufficient resources for the stated protocol; exploratory physical ceilings
Ratio obstruction (Prop. 8.3)	Proof and exact replay	Fixed effective rates, target and branch of the example

- *Crowding* (Theorem 6.1). In the family of [8] in which the $2n - 1$ equilibria of a chain (or of its symmetric lift to the cube) are separated by a small parameter δ , the reaction-path action needed to leave a declared tube around a sink is $c\delta^{2n}(1 + O(\delta))$, and the exit time satisfies $\log \tau / (\Omega\delta^{2n}) \rightarrow c$ in probability when first $\Omega \rightarrow \infty$ and then $\delta \rightarrow 0$. Since the recovery time is of order $\delta^{-(2n-2)}$ and the readout gap of order δ , barrier, recovery and resolution are tied together in this family.
- *Observation* (Propositions 6.3 and 6.4). A bounded-error readout with d_{obs} channels of range R_j and tolerance ε_j distinguishes at most $\prod_j (1 + \lfloor R_j/2\varepsilon_j \rfloor)$ labels; a passive observation changes neither the chemistry nor the exit law; and a bounded-gain observation cannot separate coalescing states.
- *Finite driving* (Theorem 7.1). Every finite set of hyperbolic sinks persists in a reversible mass-action extension that satisfies detailed balance at unit activities and is driven by one common ATP-hydrolysis potential, provided the potential is large enough; an explicit sufficient threshold is given in Proposition 7.3.
- *A certified example* (Section 8). An explicit rational three-site source has exactly seven positive equilibria in one class, four hyperbolic sinks and three unstable equilibria, with well separated readouts (normalized gaps above 0.11). Every claim about it is verified in exact arithmetic. Its certified sufficient resources are nevertheless astronomically large, and Proposition 8.3 shows that in its design family the kinase-to-substrate ratio on one retained branch is below 7×10^{-11} whatever positive loadings are used.

Table 3: Parameters with distinct roles.

Symbol	Role
$u = E/F, r = E_T/F_T$	Free-enzyme ratio (the coordinate on the equilibrium curve) and total-enzyme ratio
$p_e, q_e; k_e^\pm; \rho_e$	Loadings; effective rates; effective ratios (Section 2)
κ, η, ϵ_b	Loading scale, enzyme scale $(E_T, F_T) = (r\eta, \eta)$, and binding relaxation time in the construction of Section 4
δ	Separation of the equilibria in the crowded family of Section 6
Ω	System size: counts are Ω times concentrations (Section 5)
ϑ	ADP and phosphate activity in the reversible extension; the driving potential is $-2\log \vartheta$ in units of $k_B T$ (Section 7)
ε_j, R_j	Observation tolerance and range of readout channel j
t_{rec}, T	Certified recovery time and storage horizon

1.4 What is not claimed

The upper bound 2^n is not claimed to be sharp, and the value of C_n is open for every $n \geq 3$. The exponential construction is not economical: the number of species and of rate constants grows like $n2^n$, and its rates are badly conditioned. The resource figures in Section 8 are *sufficient* for the stated guarantees; they are neither estimates of the minimum resources of the example nor lower bounds for other designs, and we do not claim that phosphorylation memory is infeasible. No writing protocol that switches between labels is analysed; labels are prepared. Bounds on periodic or other non-equilibrium attractors are outside the scope; such attractors exist in sequential chains with $n \geq 3$ [9]. Finally, the analytic theorems are not formalized in a proof assistant; only a few algebraic identities are (Appendix C).

1.5 Organization and notation

Section 2 parametrizes the equilibria, proves a regularity lemma in which only the totals are perturbed, and derives $C_n \leq 2^n$. Section 3 treats square balance and the symmetric lift. Section 4 proves the exponential lower bound. Section 5 contains the finite-molecule theorem, Section 6 the crowding and observation results, Section 7 the reversible extension and the robustness estimate, and Section 8 the certified example. Section 9 discusses interpretation and open problems. Appendix A recalls what is used from [8], Appendix B proves Theorem 6.1, and Appendix C documents the computations. Several small parameters occur and are not interchangeable; Table 3 lists them.

2 Equilibrium geometry and the universal ceiling

2.1 Rational parametrization

Let x denote the vector of all $q + 2m + 2$ concentrations, $H(x)$ the mass-action vector field of (1.1), and $\mathcal{T}$ the linear map $x \mapsto (E_T, F_T, S_T)$ of (1.2). At an equilibrium the complex equations give

$$C_e = p_e E S_{e-}, \quad Y_e = q_e F S_{e+}, \quad (2.1)$$

so the catalytic fluxes are $c_e C_e = k_e^+ E S_{e-}$ and $\gamma_e Y_e = k_e^- F S_{e+}$. It is convenient to use the *retained pools*

$$Z_I = S_I + \sum_{e^- = I} C_e + \sum_{e^+ = I} Y_e, \quad \sum_I Z_I = S_T, \quad (2.2)$$

because $\dot{Z}_I$ contains only catalytic fluxes. Let U and V be the $q \times q$ generator matrices of the upward and downward effective rates: $U_{e^+, e^-} = k_e^+$, $V_{e^-, e^+} = k_e^-$, with diagonal entries making all column sums zero. With $u = E/F$, substituting (2.1) into $\dot{Z} = 0$ gives

$$F(uU + V)S = 0. \quad (2.3)$$

For $u > 0$ the matrix $uU + V$ is the generator of an irreducible Markov chain on the cube. By the matrix-tree theorem [44, 53] its kernel is spanned by the vector $\tau(u)$ of *tree weights*: $\tau_I(u)$ is the sum, over spanning trees directed towards I , of the products of the edge labels uk_e^+ (upward) and k_e^- (downward). Hence $S_I = s\tau_I(u)$ with a scalar $s > 0$. A tree directed towards a vertex of level $|I| = k$ contains at least k upward edges, on the path from $\emptyset$, and at least $n - k$ downward edges, on the path from $[n]$. Therefore τ_I is a polynomial in u with nonnegative coefficients,

$$u^{|I|} \mid \tau_I, \quad \deg \tau_I \leq q - 1 - n + |I|, \quad \tau_\emptyset(0) > 0. \quad (2.4)$$

Define

$$A(u) = \sum_I \tau_I, \quad B(u) = \sum_e p_e \tau_{e^-}, \quad D(u) = \frac{1}{u} \sum_e q_e \tau_{e^+}, \quad (2.5)$$

polynomials with $\deg A \leq q - 1$, $\deg B, \deg D \leq q - 2$ and $A(0), B(0) > 0$. Then $\sum_e C_e = sEB$ and $\sum_e Y_e = sFuD$, and the three totals read

$$E_T = uF(1 + sB), \quad F_T = F(1 + suD), \quad S_T = sA + suF(B + D). \quad (2.6)$$

Write $r = E_T/F_T$, $L = B - rD$ and $M = B - uD = L + (r - u)D$. Dividing the first two equations of (2.6) gives $suL(u) = r - u$. Hence, when $u \neq r$, necessarily $L(u) \neq 0$ and

$$s = \frac{r - u}{uL(u)}, \quad F = \frac{F_T L(u)}{M(u)}, \quad E = uF, \quad (2.7)$$

and the state is determined by u . An equilibrium requires $s > 0$ and $F > 0$, that is, $L(u)$, $M(u)$ and $r - u$ have the same sign. Inserting (2.7) into the third equation of (2.6) and multiplying by uLM yields, after an elementary rearrangement, the *eliminant*

$$\mathcal{P}(u) = (S_T - E_T - F_T)uLM - (r - u)AM + (u + 1)F_T uL^2 = 0. \quad (2.8)$$

Proposition 2.1 (Equilibrium count). *For all positive rates and totals, every positive equilibrium has a ratio $u = E/F$ that is a positive root of $\mathcal{P}$, distinct equilibria have distinct ratios, $\mathcal{P}$ has degree at most $2q - 1$ and $\mathcal{P}(0) < 0$. Hence every compatibility class contains at most $2^{n+1} - 1$ positive equilibria.*

Proof. The degree bound follows from (2.5), and $\mathcal{P}(0) = -rA(0)M(0) = -rA(0)B(0) < 0$, so $\mathcal{P} \not\equiv 0$. For $u \neq r$ the claims were shown above. Suppose there is an equilibrium with $u = r$. Then $L(r) = 0$, so $M(r) = 0$ and $\mathcal{P}(r) = 0$. With $u = r$ the first two totals coincide and give $F = F_T/(1 + srD(r))$, and the substrate total becomes $sA(r) + F_T sr(B(r) + D(r))/(1 + srD(r))$, which is strictly increasing in s from 0 to ∞ . So there is exactly one equilibrium with this ratio, and it is a root of $\mathcal{P}$. $\square$

Remark 2.2. The separate treatment of $u = r$ matters: a reconstruction formula with a vanishing denominator is not an exclusion proof. For the sequential chain the same elimination gives a polynomial of degree $2n + 1$; the sharper Wang–Sontag bound $2n - 1$ uses the special structure $\tau_i = t_i u^i$ [56]. We do not know whether $2^{n+1} - 1$ is attained on the cube. The example of Section 8 has $n = 3$, $\deg \mathcal{P} = 15$, and seven positive equilibria.

2.2 Regular totals at fixed rates

The next lemma is the technical core of the two upper bounds. It produces nondegenerate equilibria by perturbing only the *totals*. Perturbing rates, which is the usual device, would destroy square balance and could not be used in Section 3.

Lemma 2.3 (Regularity with fixed rates). *Fix positive rates. (i) At every positive x , the derivative $DH(x)$ has rank $d = q + 2m - 1$. (ii) The set $\mathcal{Z}$ of positive equilibria of all classes is a three-dimensional embedded submanifold, parametrized by $(u, s, F) \in \mathbb{R}_{>0}^3$ through $S = s\tau(u)$, $E = uF$ and (2.1), and $T_x\mathcal{Z} = \ker DH(x)$ for $x \in \mathcal{Z}$. (iii) A point $x \in \mathcal{Z}$ is a regular point of the restricted total map $\mathcal{T}|_{\mathcal{Z}} : \mathcal{Z} \rightarrow \mathbb{R}^3$ if and only if the Jacobian of H restricted to the class through x is nonsingular. (iv) The totals for which some equilibrium in the class is degenerate form a set of Lebesgue measure zero in $\mathbb{R}_{>0}^3$.*

Proof. (i) The image of H lies in $\ker \mathcal{T}$, which has dimension d , so the rank is at most d . For the lower bound select the $2m$ equations for $\dot{C}_e, \dot{Y}_e$ and the $q - 1$ equations for $\dot{Z}_I$, $I \neq \emptyset$ (linear combinations of the rows of DH), and the $2m$ complex columns together with the columns of S_I , $I \neq \emptyset$. Since $\dot{Z}$ depends only on the complexes, the selected $d \times d$ matrix has the block form $\begin{pmatrix} -\Delta & \Pi \\ \Gamma & 0 \end{pmatrix}$, where $\Delta = \text{diag}(b_e + c_e, \beta_e + \gamma_e)$ is positive, Π contains the entries $a_e E$ and $\alpha_e F$ in the columns of S_{e-} and S_{e+} , and Γ contains the catalytic constants with the signs of $\dot{Z}$. Its determinant is $\det(-\Delta) \det(\Gamma \Delta^{-1} \Pi)$, and $\Gamma \Delta^{-1} \Pi$ is the matrix $EU + FV = F(uU + V)$ with the row and the column of $\emptyset$ deleted. By the matrix-tree theorem this principal minor equals $\pm F^{q-1} \tau_{\emptyset}(u) \neq 0$.

(ii) The parametrization is smooth and injective, with smooth inverse $u = E/F$, $s = S_{\emptyset}/\tau_{\emptyset}(E/F)$, and by the discussion preceding (2.4) it covers $\mathcal{Z}$. Differentiating $H = 0$ along it gives $T_x\mathcal{Z} \subseteq \ker DH(x)$, and both spaces have dimension three by (i).

(iii) $D(\mathcal{T}|_{\mathcal{Z}})(x)$ is an isomorphism if and only if $T_x\mathcal{Z} \cap \ker \mathcal{T} = 0$, that is, $\ker DH(x) \cap \ker \mathcal{T} = 0$, which says that $DH(x)$ is injective on the tangent space $\ker \mathcal{T}$ of the class. (iv) is Sard's theorem [43] for the smooth map $\mathcal{T}|_{\mathcal{Z}}$ between three-dimensional manifolds. $\square$

2.3 Degree and the index of a sink

Lemma 2.4 (Boundary and degree). *For positive totals, the compatibility class contains no equilibrium on its relative boundary. If all equilibria in the class are nondegenerate, they are finite in number, and*

$$\sum_{x^*} \text{ind}(x^*) = 1, \quad \text{ind}(x^*) = \text{sign det}(-J(x^*)), \quad (2.9)$$

where J is the Jacobian restricted to the class.

Proof. Let x be an equilibrium in the class. If $E = 0$, then $\dot{C}_e = -(b_e + c_e)C_e = 0$ for all e , so $E_T = 0$, a contradiction; hence $E, F > 0$. If every S_I vanished, the complex equations would give $C_e = Y_e = 0$ and $S_T = 0$. Thus $S \neq 0$ is a nonnegative solution of (2.3) with an irreducible generator, so $S > 0$, and then all complexes are positive by (2.1). The class is compact and convex, and H points inward or is tangent on each face $\{x_i = 0\}$. For an interior point x_0 , the homotopy $(1 - t)(-H(x)) + t(x - x_0)$ has no zero on the boundary: for $0 < t < 1$ a zero would satisfy $H_i(x) = \frac{t}{1-t}(x - x_0)_i < 0$ on some face $x_i = 0$, and $t = 0$ is excluded by the first part. Hence the Brouwer degree of $-H$ on the interior of the class is one [16, 32]. Nondegenerate equilibria are isolated, and they cannot accumulate in the compact class because the limit would be a nonisolated equilibrium in its interior. The degree is the sum of their indices. $\square$

Lemma 2.5 (Index of a sink). *A sink is an isolated equilibrium of its class, and the index of $-H$ at a sink is $+1$, whether or not the sink is hyperbolic.*

Proof. All trajectories starting near a sink converge to it, so no other equilibrium is nearby. For the index, take a compact ball N in the basin. For large t the time- t map φ_t sends N into its interior, so the fixed-point index of φ_t on N is one; since no point of ∂N is periodic, the homotopy $t \downarrow 0$ transfers this to the index of $-H$. This is the classical theorem of Krasnosel'skiĭ on asymptotically stable equilibria [37]; see also [8]. $\square$

Theorem 2.6 (Universal ceiling). *For all positive rates and totals, the class contains at most 2^n sinks. Thus $C_n \leq 2^n$.*

Proof. First let the totals be regular in the sense of Lemma 2.3(iv). Let p_+ and p_- be the numbers of equilibria of index $+1$ and -1 . Then $p_+ - p_- = 1$ by (2.9) and $p_+ + p_- \leq 2q - 1$ by Proposition 2.1, so $p_+ \leq q$.

For arbitrary totals c , suppose there were $q + 1$ sinks. Eliminate $S_\emptyset, E, F$ by (1.2); this identifies all nearby classes with open subsets of one space $\mathbb{R}^d$ (the *class chart*) on which the field depends continuously on the totals. Choose disjoint compact neighbourhoods $N_1, \dots, N_{q+1}$ of the sinks, inside the positive class, each containing no other equilibrium. By Lemma 2.5 the degree of $-H$ on each N_i is one. For totals c' close to c the field has no zero on ∂N_i and the degrees are unchanged. By Lemma 2.3 we may take c' regular; then every N_i contains an equilibrium of index $+1$, so $p_+ \geq q + 1$, a contradiction. $\square$

3 Square balance and the sequential ceiling

3.1 Compression of a balanced cube to a chain

Assume (1.3). Any two monotone paths from $\emptyset$ to I differ by a sequence of swaps of adjacent site additions, and each swap leaves the product of the ratios ρ_e along the path unchanged. Hence there are positive weights with

$$w_\emptyset = 1, \quad w_{e+} = \rho_e w_{e-} \quad \text{for every edge } e. \quad (3.1)$$

Then $S_I = s w_I u^{|I|}$ satisfies $u k_e^+ S_{e-} = k_e^- S_{e+}$ on every edge, so it solves (2.3), and by irreducibility $\tau_I(u)$ is proportional to $w_I u^{|I|}$: at a balanced equilibrium every edge carries zero net effective flux. Group the weights by level,

$$A_k = \sum_{|I|=k} w_I, \quad B_k = \sum_{|e^-|=k} p_e w_{e-}, \quad D_k = \sum_{|e^+|=k+1} q_e w_{e+} \quad (0 \leq k \leq n-1), \quad (3.2)$$

and $A_n = w_{[n]}$. After division by the common factor, $A = \sum_k A_k u^k$, $B = \sum_k B_k u^k$ and $D = \sum_k D_k u^k$.

Lemma 3.1 (Compression). *Let the cube rates be square balanced. The sequential chain with loadings, effective rates and catalytic constants*

$$\bar{p}_{k+1} = \frac{B_k}{A_k}, \quad \bar{q}_{k+1} = \frac{D_k}{A_{k+1}}, \quad \bar{c}_{k+1} = \frac{D_k}{B_k}, \quad \bar{\gamma}_{k+1} = 1 \quad (0 \leq k \leq n-1),$$

and any positive dissociation constants, has tree weights $A_k u^k$ and the same polynomials A, B, D . Consequently, for every triple of totals, $(u, s, F) \mapsto (u, s, F)$ is a bijection between the positive equilibria of the cube and those of the chain.

Proof. The effective ratio of the chain at step $k+1$ is $\bar{c}_{k+1}\bar{p}_{k+1}/(\bar{\gamma}_{k+1}\bar{q}_{k+1}) = A_{k+1}/A_k$, which gives the tree weights $A_k u^k$, and then $\sum_k \bar{p}_{k+1} A_k u^k = B$ and $\sum_k \bar{q}_{k+1} A_{k+1} u^k = D$. The equations (2.6), which determine the equilibria, involve the network only through A, B, D . Association constants $\bar{a} = \bar{p}(\bar{b} + \bar{c})$ and $\bar{\alpha} = \bar{q}(\bar{\beta} + \bar{\gamma})$ realize the loadings. $\square$

The lemma is a statement about equilibria. It does not assert that the dynamics of the balanced cube reduces to that of the chain, and in general it does not.

3.2 The symmetric lift

Under a stronger hypothesis the reduction is exact. Let a chain have the rate constants $a_k, b_k, c_k, \alpha_k, \beta_k, \gamma_k$ for $1 \leq k \leq n$. Its *symmetric lift* to the cube is defined by

$$a_e = \frac{a_{k+1}}{n-k}, \quad \alpha_e = \frac{\alpha_{k+1}}{k+1}, \quad (b_e, c_e, \beta_e, \gamma_e) = (b_{k+1}, c_{k+1}, \beta_{k+1}, \gamma_{k+1}) \quad \text{if } |e^-| = k. \quad (3.3)$$

The lifted ratios ρ_e depend only on the level, so the lift is square balanced. Let Π be the aggregation map that sums the S_I over each level and the C_e and Y_e over all edges with tails at the same level.

Proposition 3.2 (Exact deterministic reduction). *For the symmetric lift, $\Pi x(t)$ together with E, F solves the mass-action equations of the chain with the same totals. The positive equilibria of the cube are exactly the uniform lifts of those of the chain; an equilibrium of the cube is a sink if and only if the corresponding equilibrium of the chain is, and hyperbolic sinks correspond to hyperbolic sinks.*

Proof. A vertex of level k has $n-k$ upward and k downward edges, so $\sum_{|e^-|=k} a_e E S_{e^-} = a_{k+1} E \sum_{|I|=k} S_I$ and $\sum_{|e^+|=k+1} \alpha_e F S_{e^+} = \alpha_{k+1} F \sum_{|I|=k+1} S_I$; all other terms are linear with level-dependent constants. This gives the first claim. Collect the $q+2m$ substrate-carrying species in a vector z . Then $\dot{z} = \mathcal{A}(E, F)z$, where, for $E, F > 0$, $\mathcal{A}$ is the generator of an irreducible Markov chain describing the fate of one substrate molecule, and E, F depend on z only through $y = \Pi z$. Let Θ be the uniform lift, $\Pi\Theta = I$. The symmetry of (3.3) gives $\Pi\mathcal{A} = \bar{\mathcal{A}}\Pi$ and $\mathcal{A}\Theta = \Theta\bar{\mathcal{A}}$, where $\bar{\mathcal{A}}$ is the corresponding matrix of the chain. Hence $w = z - \Theta\Pi z$ satisfies $\dot{w} = \mathcal{A}(E, F)w$ on the invariant subspace $\ker \Pi$, while y is autonomous. The null vector of $\mathcal{A}$ is positive and therefore not in $\ker \Pi$, so $\mathcal{A}|_{\ker \Pi}$ is Hurwitz. Thus every equilibrium has $w = 0$; the Jacobian is block triangular with blocks $\bar{\mathcal{A}}$ and $\mathcal{A}|_{\ker \Pi}$; and asymptotic stability of y^* implies that of $(y^*, 0)$, because w obeys a linear equation whose coefficient matrix converges to a Hurwitz matrix. The converse holds because $\{w = 0\}$ is invariant. $\square$

3.3 Proof of Theorem 1.3

Let the rates be square balanced. By Lemma 3.1 and the Wang–Sontag theorem [56, Theorem 3], every class contains at most $2n-1$ positive equilibria. At regular totals, (2.9) gives $p_+ - p_- = 1$ and $p_+ + p_- \leq 2n-1$, so $p_+ \leq n$. Now repeat the second half of the proof of Theorem 2.6: if some class had $n+1$ sinks, nearby regular totals, *for the same balanced rates*, would give $p_+ \geq n+1$. This is where Lemma 2.3 is needed in the form stated. For attainment, [8, Theorem 1.1] provides a chain with n hyperbolic sinks (Appendix A), and by Proposition 3.2 its symmetric lift has n hyperbolic sinks. $\square$

Remark 3.3 (What balance does and does not mean). (a) Square balance is a condition on the $2m$ effective ratios only. It does not make the network detailed balanced: every balanced equilibrium carries the positive futile-cycle flux $c_e C_e = \gamma_e Y_e$ on each edge. (b) The effective affinity of a square

compares two kinetic pathways between the same phosphoforms. It is not the free energy of ATP hydrolysis, which in the reversible extension of Section 7 is the same on every edge whatever the effective affinities are. (c) Nonzero affinity is necessary for more than n sinks but not sufficient; in exploratory computations, squares with circulating rates and several admissible branches did not by themselves produce additional stable branches.

4 Exponentially many sinks from unrestricted loading

Theorem 1.3 shows that additional sinks must come from unbalanced kinetics. This section constructs them. The idea is to separate two tasks. The *effective rates* $k_e^\pm$ are kept near the independent-site values, which makes a one-parameter family of saturated states normally attracting. The *loadings* p_e, q_e , which can be varied without changing the effective rates, are then used to program the slow drift along this family. The loadings enter linearly, and their span has dimension at least $2^{n-2} + n$.

4.1 The loading space and its positive kernel

Fix the effective rates. By (2.5) the map $(p, q) \mapsto L = B - rD$ is linear, with image

$$\mathcal{W} = \text{span}\{\tau_{e-}(u), \tau_{e+}(u)/u : e \text{ an edge}\}. \quad (4.1)$$

Lemma 4.1 (Positive kernel). *For all u , $u \sum_e k_e^+ \tau_{e-}(u) = \sum_e k_e^- \tau_{e+}(u)$. Consequently $(p, q) = (rk^+, k^-)$ lies in the kernel of $(p, q) \mapsto B - rD$, and any real solution (p, q) of $B - rD = L$ becomes a positive solution after adding a sufficiently large multiple of it.*

Proof. Multiply the I -th row of $(uU + V)\tau = 0$ by $|I|$ and sum. Every upward transition raises $|I|$ by one and every downward transition lowers it by one, which gives the identity; the rest is immediate. $\square$

Lemma 4.2 (Interpolation rank). *Let $d_0 = 2^{n-2} + n$ and let $u_1 < \dots < u_{d_0}$ be positive. The set of effective rates $k^\pm \in \mathbb{R}_{>0}^{2m}$ for which the evaluation map $\mathbb{R}^{2m} \rightarrow \mathbb{R}^{d_0}$, $(p, q) \mapsto (L(u_j))_j$, is surjective is open and dense. In particular it contains rational points arbitrarily close to the independent-site rates $k_e^+ = k_e^- = 1$.*

Proof. A $d_0 \times d_0$ minor of the evaluation matrix is a polynomial in the effective rates, so it suffices to find one point of the closed orthant $\mathbb{R}_{\geq 0}^{2m}$ where some minor is nonzero: a polynomial that is not identically zero is nonzero on an open dense set. We use a sparse witness in which some rates vanish but the graph remains connected, so that τ is still the vector of tree weights.

Split the cube into $K = 2^{n-2}$ square faces by fixing the last $n - 2$ coordinates $b \in \{0, 1\}^{n-2}$; let $h = |b|$ be the height of face b and call its vertex with first two coordinates 00 its *anchor*. Keep the four edges of each face. Join the anchors by a spanning tree of the $(n - 2)$ -cube of anchors, and set the rates of all other edges between faces to zero (Figure 1). On face b take upward rates $\lambda_b(1, 2, 3, 5)$ and downward rates $(7, 11, 13, 17)$ on the edges $00 \rightarrow 10$, $00 \rightarrow 01$, $10 \rightarrow 11$, $01 \rightarrow 11$, with distinct $\lambda_b > 0$. On each tree edge take $k^+ = k^- = 1$.

Removing a tree edge disconnects the graph, so its stationary net flux vanishes and the anchor weights are in the ratios u^h . Within a face the weights are those of the isolated square. A direct computation (or the matrix-tree theorem on a 4-cycle) gives, for the anchor and the two middle vertices of the square,

$$2310 + 1016\lambda_b u, \quad \lambda_b u (330 + 195\lambda_b u), \quad \lambda_b u (420 + 153\lambda_b u). \quad (4.2)$$

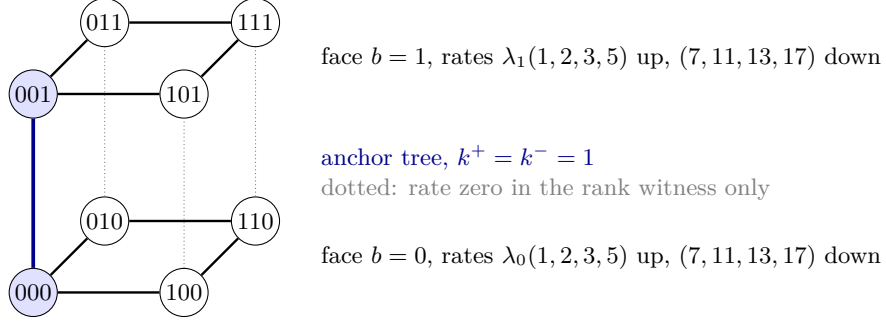


Figure 1: The cube of phosphoforms for $n = 3$ and the sparse witness used in the proof of Lemma 4.2. Each vertex I carries a free form S_I ; each edge carries the two complexes C_e and Y_e of (1.1). The witness only proves that a polynomial minor does not vanish identically; all dynamical statements use the full cube with positive rates on every edge.

Let $\ell_b(u) = 2310 + 1016\lambda_b u$ and $Z = \prod_b \ell_b$. Up to one common factor, the weights are $u^h Z$ at the anchors and $u^h Z$ times the last two expressions in (4.2) divided by ℓ_b at the middle vertices. All these vertices are tails of retained edges. Since $330 \cdot 153 - 420 \cdot 195 = -31410 \neq 0$, the two linear factors in (4.2) span all linear polynomials, so $\mathcal{W}$ contains $u^{h+1}Z$ and $u^{h+1}Z/\ell_b$ for every face, as well as $u^h Z$. Hence $u^k Z \in \mathcal{W}$ for $0 \leq k \leq n-1$. Dividing u^{h+1} by ℓ_b leaves a polynomial of degree $h \leq n-2$ and a nonzero constant remainder, so $Z/\ell_b \in \mathcal{W}$ for every b . The K polynomials Z/ℓ_b have degree $K-1$ and are linearly independent, as evaluation at the K distinct roots of the ℓ_b shows; together with $Z, uZ, \dots, u^{n-1}Z$ they span all polynomials of degree at most $K+n-1$. Evaluation of this d_0 -dimensional space at d_0 distinct points is bijective, and the common factor only rescales the rows of the evaluation matrix. $\square$

4.2 A normally attracting saturated family

For $S \in \mathbb{R}_{>0}^q$ put $A_1(S) = \sum_e k_e^+ S_{e-}$ and $B_1(S) = \sum_e k_e^- S_{e+}$, and consider the *saturated field*

$$\frac{dS}{d\sigma} = G(S) = \frac{US}{A_1(S)} + \frac{VS}{B_1(S)}. \quad (4.3)$$

It describes substrate processed by two fully saturated enzymes: the total upward and the total downward flux are both equal to one. Hence G conserves $S_0 = \sum_I S_I$ and the modification count $\Phi(S) = \sum_I |I| S_I$.

Lemma 4.3 (Normal attraction). (i) *For every $S_0 > 0$, the equilibria of (4.3) with $\sum S_I = S_0$ form the curve $\Gamma = \{S_0 \tau(u)/A(u) : u > 0\}$.* (ii) *Let $J_u \subset (0, \infty)$ be a compact interval. For all effective rates in a neighbourhood of the independent-site rates, depending on J_u , and at every point of Γ with $u \in J_u$, the linearization of G on $\{\sum h_I = 0\}$ has a simple eigenvalue zero, with eigenvector tangent to Γ , and $q-2$ eigenvalues with negative real parts, and $d\Phi/du > 0$ along Γ .*

Proof. (i) $G(S) = B_1^{-1}((B_1/A_1)U + V)S$ vanishes if and only if S is proportional to $\tau(u')$ with $u' = B_1(S)/A_1(S)$, and by Lemma 4.1 this relation holds identically for $S \propto \tau(u')$. (ii) At independent-site rates, $A_1 = nS_0 - \Phi$ and $B_1 = \Phi$ are constant on each invariant slice $\{S_0, \Phi \text{ fixed}\}$, on which G is therefore the linear field of the irreducible generator $G_0 = U/A_1 + V/B_1$. The row vectors $\mathbf{1}^T$ and $(|I|)_I$ span a left-invariant subspace of G_0 with eigenvalues 0 and $-(1/A_1 + 1/B_1)$, so their common kernel, which is the tangent space of the slice, is invariant and carries the remaining $q-2$ eigenvalues,

all with negative real parts. The stationary distribution is binomial, with $\Phi/S_0 = nu/(1+u)$, which is increasing. The curve Γ is explicit for all effective rates and depends smoothly on them, and Φ is conserved for all of them; by continuity of the spectrum on the compact set J_u , the conclusion persists for nearby rates. $\square$

The point of the lemma is that transverse stability is obtained independently of the scalar equation that will determine the position of the equilibria along Γ .

4.3 Realization with retained enzymes

Fix $J_u = [u_1, u_{d_0}]$ and $r > u_{d_0}$, effective rates as in Lemmas 4.2 and 4.3, a small $\eta > 0$, and totals $E_T = r\eta$, $F_T = \eta$, $S_T > \eta(r+1)$. For real *corrections* $v^\pm \in \mathbb{R}^m$, a large κ and a small $\epsilon_b > 0$ define

$$\begin{aligned} p_e &= \kappa r k_e^+ + v_e^+, & c_e &= k_e^+/p_e, & a_e &= p_e/\epsilon_b, & b_e &= 1/\epsilon_b - c_e, \\ q_e &= \kappa k_e^- + v_e^-, & \gamma_e &= k_e^-/q_e, & \alpha_e &= q_e/\epsilon_b, & \beta_e &= 1/\epsilon_b - \gamma_e. \end{aligned} \quad (4.4)$$

For large κ and small ϵ_b all constants are positive, the loadings are p_e, q_e and the effective rates are $k_e^\pm$, independently of κ , $v^\pm$ and ϵ_b . In particular the *equilibria* of (1.1) do not depend on ϵ_b .

Fast binding. As $\epsilon_b \rightarrow 0$ the binding reactions equilibrate. The fast subsystem is a detailed-balanced mass-action system whose $2m$ reaction vectors are linearly independent, since each involves its own complex. In reaction-extent coordinates its linearization at the binding equilibrium is $-\Phi_f \Gamma_f^T X^{-1} \Gamma_f$, where Γ_f is the fast stoichiometric matrix, Φ_f the diagonal matrix of equilibrium fluxes and X the diagonal matrix of concentrations; it is similar to a negative definite matrix. Hence the manifold (2.1) of binding equilibria is normally attracting, and by the standard theory of singularly perturbed systems [24, 36], at an equilibrium the Jacobian of (1.1) has, for small ϵ_b , $2m$ eigenvalues with real parts of order $-1/\epsilon_b$ and $q-1$ eigenvalues converging to those of the *reduced system*

$$\dot{Z} = EUS + FVS, \quad E = \frac{r\eta}{1 + \sum_e p_e S_{e-}}, \quad F = \frac{\eta}{1 + \sum_e q_e S_{e+}}, \quad (4.5)$$

in which S is a function of the pools Z of (2.2). The map $S \mapsto Z$ is a local diffeomorphism: in logarithmic coordinates $\sigma_I = \log S_I$ its derivative is

$$\mathbf{M} = \text{diag}(S + c + y) - \frac{cc^T}{E_T} - \frac{yy^T}{F_T}, \quad c_I = \sum_{e^- = I} C_e, \quad y_I = \sum_{e^+ = I} Y_e, \quad (4.6)$$

and the Cauchy–Schwarz inequality gives $(c^T z)^2 \leq (\sum_I c_I) \sum_I c_I z_I^2 \leq E_T \sum_I c_I z_I^2$, and similarly for y , so $z^T \mathbf{M} z \geq \sum_I S_I z_I^2 > 0$.

Saturation. Put $\varepsilon = 1/\kappa$ and use the time $\sigma = \eta t/\kappa$. Then (4.5) becomes

$$\frac{dZ}{d\sigma} = \hat{E}US + \hat{F}VS, \quad \hat{E} = \frac{r}{\varepsilon + rA_1(S) + \varepsilon \sum_e v_e^+ S_{e-}}, \quad \hat{F} = \frac{1}{\varepsilon + B_1(S) + \varepsilon \sum_e v_e^- S_{e+}}, \quad (4.7)$$

with $C_e = \eta(rk_e^+ + \varepsilon v_e^+) \hat{E} S_{e-}$ and $Y_e = \eta(k_e^- + \varepsilon v_e^-) \hat{F} S_{e+}$. This is a family of vector fields that is smooth in (S, ε) up to $\varepsilon = 0$. At $\varepsilon = 0$ the field is the saturated field G written in the coordinates $Z = S + \eta(\cdots)$, a near-identity change of variables; both enzymes are fully bound, so $\sum_I S_I = S_0 := S_T - \eta(r+1)$; and $\Phi_Z = \sum_I |I| Z_I$ is conserved. For η sufficiently small, Lemma 4.3 therefore holds for the field at $\varepsilon = 0$ in the coordinates Z , with $d\Phi_Z/du > 0$ along Γ .

4.4 Programming the drift

For $\varepsilon > 0$ the count Φ_Z is no longer conserved. From (4.7),

$$\frac{d\Phi_Z}{d\sigma} = \hat{E}A_1 - \hat{F}B_1 = \varepsilon \left[\frac{1 + \sum_e v_e^- S_{e+}}{B_1(S)} - \frac{1 + \sum_e v_e^+ S_{e-}}{rA_1(S)} \right] + O(\varepsilon^2), \quad (4.8)$$

uniformly on compact subsets of $\mathbb{R}_{>0}^q$; note that the term of order zero vanishes for *every* S , not only on Γ . On Γ we have $S = (S_0/A(u))\tau(u)$ and $B_1 = uA_1$, and the bracket equals $\mathcal{N}(u)/(rB_1)$ with the *programmed numerator*

$$\mathcal{N}(u) = r - u - \frac{uS_0}{A(u)} L_v(u), \quad L_v(u) = \sum_e v_e^+ \tau_{e-}(u) - \frac{r}{u} \sum_e v_e^- \tau_{e+}(u). \quad (4.9)$$

Theorem 4.4 (Exponential lower bound). *For every $n \geq 2$ there are positive rational rate constants and totals for which (1.1) has at least $\lfloor (2^{n-2} + n)/2 \rfloor$ hyperbolic sinks in one compatibility class.*

Proof. Step 1: corrections. The map $v \mapsto (L_v(u_j))_j$ is the evaluation map of Lemma 4.2 up to the signs and factors r of some columns, hence surjective. Choose v with $\mathcal{N}(u_j) = (-1)^{j+1}$ for $1 \leq j \leq d_0$. Let e_0 be an edge and $g(u) = uS_0\tau_{e_0}^-(u)/A(u) > 0$. Replacing $v_{e_0}^+$ by $v_{e_0}^+ + \theta$ replaces $\mathcal{N}$ by $\mathcal{N} - \theta g$, whose zeros are the solutions of $\mathcal{N}/g = \theta$. By Sard's theorem, for almost every θ all zeros in J_u are simple, and for small $|\theta|$ the signs at the u_j are unchanged. We fix such a rational v . Each of the $\lfloor d_0/2 \rfloor$ intervals (u_j, u_{j+1}) with j odd then contains a zero u_* with $\mathcal{N}'(u_*) < 0$.

Step 2: small ε . Near the compact arc of Γ over J_u , and within the class $\sum_I Z_I = S_T$, use coordinates (ϕ, ζ) with $\phi = \Phi_Z$ and $\zeta \in \mathbb{R}^{q-2}$ vanishing on Γ . The family (4.7) takes the form

$$\phi' = \varepsilon g(\phi, \zeta, \varepsilon), \quad \zeta' = \Xi(\phi, \zeta, \varepsilon), \quad \Xi(\phi, 0, 0) = 0, \quad \partial_\zeta \Xi(\phi, 0, 0) \text{ Hurwitz},$$

and $g(\phi, 0, 0) = \mathcal{N}(u)/(rB_1)$ at the point of Γ with $\Phi_Z = \phi$. The implicit function theorem solves $\Xi = 0$ by $\zeta = \zeta^*(\phi, \varepsilon) = O(\varepsilon)$, and then each simple zero ϕ_* of $g(\cdot, 0, 0)$ continues to an equilibrium for small $\varepsilon > 0$. Its Jacobian has $q - 2$ eigenvalues near those of $\partial_\zeta \Xi$, and one eigenvalue $\varepsilon(\partial_\phi g - \partial_\zeta g(\partial_\zeta \Xi)^{-1} \partial_\phi \Xi) + O(\varepsilon^2)$. Since $\Xi(\phi, 0, 0) \equiv 0$, we have $\partial_\phi \Xi = O(\varepsilon)$ at the equilibrium, so this eigenvalue is

$$\varepsilon \partial_\phi g(\phi_*, 0, 0) + O(\varepsilon^2), \quad \text{sign } \partial_\phi g(\phi_*, 0, 0) = \text{sign } \mathcal{N}'(u_*) < 0, \quad (4.10)$$

because $d\Phi_Z/du > 0$ and $rB_1 > 0$. Thus for one sufficiently large rational κ the reduced system (4.5) has $\lfloor d_0/2 \rfloor$ hyperbolic sinks, and all p_e, q_e are positive.

Step 3: small ε_b . By the discussion of fast binding, for sufficiently small rational $\varepsilon_b > 0$ these equilibria are hyperbolic sinks of (1.1). The choices were made in the order $n; u_j, r; k^\pm; \eta, S_T; v^\pm; \kappa; \varepsilon_b$. Each condition is open, so all parameters can be taken rational, and none is chosen depending on a later one. $\square$

Proof of Theorem 1.2. The upper bound is Theorem 2.6. The lower bound n is the attainment part of Theorem 1.3, the exponential bound is Theorem 4.4, and $C_3 \geq 4$ is Proposition 8.1. For $n \geq 3$, $\lfloor (2^{n-2} + n)/2 \rfloor \geq 2^{n-3}$. $\square$

Remark 4.5 (Cost of the construction). (a) *Speed.* With the effective rates, the corrections and η fixed, the slow eigenvalue (4.10) is of order $1/\kappa$ in the time $\sigma = \eta t/\kappa$, hence of order η/κ^2 in the original time, while the transverse rates are of order η/κ . To transfer this asymptotic statement to (1.1) the relaxation parameter must be chosen as $\varepsilon_b(\kappa)$ with spectral error $o(\eta/\kappa^2)$; a fixed ε_b is

not covered. (b) *Positivity*. If v solves the interpolation problem, the loadings $v + \kappa(rk^+, k^-)$ are positive exactly when $\kappa > \max_i(-v_i/k_i)$; large corrections thus force large κ and slow restoration. Optimizing over the whole kernel of the evaluation map, rather than along one kernel direction, can reduce the burden; this is how the example of Section 8 was improved. (c) *Conditioning*. For a scaled evaluation matrix with singular triplets (σ_j, u_j, v_j) , every solution of $Mv = y$ satisfies $\|v\| \geq |u_j^T y|/\sigma_j$. The burden depends on the alignment of the target y with the small singular directions, not on the condition number alone. (d) *Description*. The labels provide $n - 3 \leq \log_2 C_n \leq n$ bits of deterministic codebook for $n \geq 3$, but the network has $(n + 1)2^n + 2$ species and $3n2^n$ rate constants; two enzyme types do not make the kinetic description small. None of (a)–(d) is a lower bound for other constructions.

5 Finite-molecule operation

Stable equilibria of the rate equations become usable labels only if a finite number of molecules stays near them, returns after a disturbance, and can be told apart. This section formulates such a contract for the full stochastic model of (1.1) and proves computable sufficient conditions. No diffusion approximation and no stochastic quasi-steady-state reduction is used; such reductions need additional hypotheses [35].

5.1 The Markov chain and exact lumping

At system size $\Omega > 0$ the state is a vector N of molecule counts, and $X = N/\Omega$ is the vector of concentrations. A binding reaction with constant k and reactant counts N_i, N_j has propensity kN_iN_j/Ω ; the two reactants are always distinct species, so no combinatorial correction arises. A unimolecular reaction has propensity kN_i . With $a_\rho(x)$ the mass-action rate of reaction ρ and ν_ρ its reaction vector, the generator is

$$\mathcal{L}_\Omega \varphi(x) = \Omega \sum_\rho a_\rho(x) [\varphi(x + \nu_\rho/\Omega) - \varphi(x)] \quad (5.1)$$

[2]. The totals (1.2) are conserved, so for Ω such that $\Omega E_T, \Omega F_T, \Omega S_T$ are integers the chain lives on a finite set, and it is nonexplosive. We use the class chart of the proof of Theorem 2.6: the d coordinates are the S_I with $I \neq \emptyset$ and all complexes, and $S_\emptyset, E, F$ are recovered from the totals. In this chart $\nu_\rho \in \mathbb{R}^d$, H denotes the vector field, J its Jacobian, and $\|\cdot\|$ the Euclidean norm.

Proposition 5.1 (Exact stochastic reduction). *For the symmetric lift (3.3), and for every initial distribution, the aggregated process $\Pi N(t)$ is a Markov chain whose generator is that of the sequential chain at the same Ω . Consequently any event defined by the aggregated path has the same probability in both models.*

Proof. The reactions of the cube that induce a given jump of ΠN are those of one type at one level. Their propensities add up to $\sum_{|e^-|=k} \frac{a_{k+1}}{n-k} N_E N_{S_{e^-}}/\Omega = a_{k+1} N_E (\sum_{|I|=k} N_{S_I})/\Omega$ for kinase binding, similarly for phosphatase binding, and to a level constant times an aggregated complex count for the unimolecular reactions. These sums depend on N only through ΠN , which is the criterion for strong lumpability [34]. $\square$

The reduction is exact, not asymptotic, but it is limited to observations that factor through Π and to the lifted rates; independently perturbed site-specific rates or site-specific readouts are not covered.

5.2 Retention and recovery

Theorem 5.2 (Computable finite memory). *Let x^* be a positive hyperbolic sink with chart Jacobian J . Let P be a symmetric matrix and $\underline{p}, \bar{p}, \mu, K, \varrho > 0$ constants such that*

$$\underline{p}I \leq P \leq \bar{p}I, \quad -(J^T P + P J) \geq \mu I, \quad \|H(x^* + h) - Jh\| \leq K\|h\|^2 \quad (\|h\| \leq \varrho), \quad (5.2)$$

the ball $\|h\| \leq \varrho$ lies in the positive class, and $\varrho \leq \mu/(4\bar{p}K)$. Put

$$V(x) = h^T P h \quad (h = x - x^*), \quad v = \underline{p}\varrho^2, \quad \lambda = \frac{\mu}{2\bar{p}}, \quad D = \{V < v\},$$

and let $C \geq \sup_{x \in D} \sum_{\rho} a_{\rho}(x) \nu_{\rho}^T P \nu_{\rho}$. Let τ be the first exit time of X from D . Then for every Ω , every $X_0 \in D$ and every $t \geq 0$,

$$\Pr(\tau \leq t) \leq \frac{V(X_0) + Ct/\Omega}{v}, \quad (5.3)$$

$$\mathbb{E}[V(X_t) \mathbf{1}_{\{\tau > t\}}] \leq e^{-\lambda t} V(X_0) + \frac{C}{\lambda \Omega}. \quad (5.4)$$

In particular, with the preparation set $\mathbf{B} = \{V \leq v/4096\}$ and the disturbance set $\mathbf{A} = \{V \leq v/1024\}$, for every $X_0 \in \mathbf{A}$ and $t \geq 8/\lambda$,

$$\Pr(\tau \leq t \text{ or } X_t \notin \mathbf{B}) \leq \frac{1}{1024} + \frac{Ct}{\Omega v} + \frac{1}{256} + \frac{4096 C}{\lambda \Omega v}. \quad (5.5)$$

Proof. Since V is quadratic, (5.1) gives the exact identity

$$\mathcal{L}_{\Omega} V(x) = 2h^T P H(x) + \frac{1}{\Omega} \sum_{\rho} a_{\rho}(x) \nu_{\rho}^T P \nu_{\rho}, \quad (5.6)$$

with no remainder. On D we have $\underline{p}\|h\|^2 \leq V < \underline{p}\varrho^2$, so $\|h\| < \varrho$, and by (5.2) $2h^T P H \leq -\mu\|h\|^2 + 2\bar{p}K\|h\|^3 \leq -\frac{\mu}{2}\|h\|^2 \leq -\lambda V$. Hence $\mathcal{L}_{\Omega} V \leq -\lambda V + C/\Omega$ on D . Dynkin's formula for the stopped process gives $\mathbb{E}V(X_{t \wedge \tau}) \leq V(X_0) + Ct/\Omega$. At the exit time the chain has jumped out of D , so $V(X_{\tau}) \geq v$ whatever the overshoot, and Markov's inequality gives (5.3). Applying Dynkin's formula to $e^{\lambda s} V(X_s)$ gives

$$\mathbb{E}[e^{\lambda(t \wedge \tau)} V(X_{t \wedge \tau})] \leq V(X_0) + \frac{C}{\Omega} \int_0^t e^{\lambda s} ds,$$

and $e^{\lambda t} V(X_t) \mathbf{1}_{\{\tau > t\}}$ is bounded by the integrand on the left, which gives (5.4). Finally, by Markov's inequality, $\Pr(\tau > t, V(X_t) > v/4096) \leq (4096/v)(e^{-\lambda t} v/1024 + C/(\lambda \Omega))$, and $4e^{-8} < 1/256$; adding (5.3) with $V(X_0) \leq v/1024$ gives (5.5). $\square$

The scalar form of the identity (5.6) and the multiplicity identity used in Proposition 5.1 are among the machine-checked algebraic facts (Appendix C).

Corollary 5.3 (Operational contract). *Let $x_1^*, \dots, x_{N_{\text{lab}}}^*$ be hyperbolic sinks in one class with data $(P_i, \varrho_i, v_i, \lambda_i, C_i)$ as in Theorem 5.2. Let ℓ be a linear readout and $\varepsilon_{\text{obs}} \geq 0$ a bound for the observation error, and assume that the intervals $[\ell(x_i^*) - \|\ell\|\varrho_i - \varepsilon_{\text{obs}}, \ell(x_i^*) + \|\ell\|\varrho_i + \varepsilon_{\text{obs}}]$ are pairwise disjoint. Put $t_{\text{rec}} = \max_i 8/\lambda_i$, $T = 100 t_{\text{rec}}$, and let Ω be compatible with the totals and satisfy*

$$\Omega \geq \max_i \max \left\{ \frac{1000 C_i T}{v_i}, \frac{10^6 C_i}{\lambda_i v_i}, \frac{1024 d \bar{p}_i}{v_i}, 1 \right\}. \quad (5.7)$$

Then: (a) each B_i contains a lattice point of the class, obtained by rounding the chart coordinates of x_i^* ; (b) from every state of B_i the chain stays in D_i during $[0, T]$ with probability at least $1 - 0.0013$, and on this event the interval decoder returns i at all times; (c) from every state of A_i the chain is in B_i at time t_{rec} , without having left D_i , with probability at least $1 - 0.009$; (d) recovery from A_i followed by storage for time T fails with probability less than 0.011.

Proof. (a) Rounding each chart coordinate of Ωx_i^* to a nearest integer changes x by at most $1/(2\Omega)$ per coordinate, so $V_i \leq \bar{p}_i d / (4\Omega^2) \leq v_i / 4096$ by the third term of (5.7) and $\Omega \geq 1$; the dependent counts are then integers by compatibility and positive because the ball lies in the positive class. When x_i^* is irrational one rounds a sufficiently accurate enclosure instead. (b) and (c) follow from (5.3) and (5.5): the terms are at most $1/4096 + 10^{-3}$ and $1/1024 + 10^{-5} + 1/256 + 0.0041$. On $\{\tau_i > T\}$ the readout differs from $\ell(x_i^*)$ by at most $\|\ell\|_{\varrho_i}$. (d) is the Markov property at time t_{rec} and a union bound. $\square$

Remark 5.4 (Three limits of the contract). (a) Retention is continuous: an excursion out of D_i counts as a failure even if the chain returns. (b) Recovery is from the local disturbance sets A_i , not from arbitrary states. (c) Labels are *prepared*. No protocol that switches the system from one label to another with a common finite control is analysed, and the corollary does not describe a complete write–store–read device. The individual guarantees (b) and (c) exceed 0.99; the composite guarantee (d) is what may be quoted for recovery followed by storage.

Proposition 5.5 (A common clock does not change the budget). *Multiply all rate constants by $\varsigma > 0$ and keep the totals, the matrices P_i and the radii ϱ_i . Then the hypotheses (5.2) hold with $(\mu_i, K_i, C_i, \lambda_i)$ replaced by $\varsigma(\mu_i, K_i, C_i, \lambda_i)$, the times t_{rec} and T are divided by ς , and the right-hand side of (5.7) is unchanged.*

Proof. The generator becomes $\varsigma \mathcal{L}_\Omega$, so the new process is the old one run at time ςt . The ratio μ_i/K_i , and hence the admissible radius and v_i , are unchanged, and $C_i T$ and C_i/λ_i are invariant. $\square$

Thus slowing a design to respect physical upper limits on rate constants changes the reported times but not the molecule budget of the matched protocol $T = 100 t_{\text{rec}}$; at a fixed absolute storage time the bound would change, and recovery would have to be charged on the same clock.

6 Crowding, restoration, and observation

6.1 The exit barrier of crowded states

The chain attains its n sinks in [8] by splitting one equilibrium of multiplicity $2n - 1$. For $n \geq 2$ and real numbers $\xi_0 < \dots < \xi_{2n-2}$, that paper constructs rate constants $\kappa(\delta)$, analytic in a small parameter $\delta \geq 0$, with fixed totals, such that for $\delta > 0$ the class contains exactly the $2n - 1$ equilibria $X_j(\delta)$, with $E = 1 + \delta \xi_j/2$ at $X_j(\delta)$, all converging to one point $X_\circ$ as $\delta \rightarrow 0$; the X_j with j even are hyperbolic sinks, and the Jacobian at $X_j(\delta)$ has $3n - 1$ eigenvalues with real parts below a negative constant and one eigenvalue

$$\lambda_j(\delta) = -a \delta^{2n-2} \prod_{k \neq j} (\xi_j - \xi_k) (1 + O(\delta)), \quad a > 0 \quad (6.1)$$

(Appendix A). We call this the *crowded family*. By Propositions 3.2 and 5.1 everything below applies verbatim to the symmetric lift of the family to the cube and to events defined through the aggregation Π .

For a path $\phi : [0, t] \rightarrow$ class chart, the reaction-path action is $\int_0^t \mathcal{I}(\phi, \dot{\phi}) ds$, where $\mathcal{I}(x, \cdot)$ is the Legendre transform of the Hamiltonian

$$\mathcal{H}(x, \theta) = \sum_{\rho} a_{\rho}(x) (e^{\theta \cdot \nu_{\rho}} - 1), \quad (6.2)$$

the large-deviation rate function of the chain (5.1) [49, 26, 1].

Theorem 6.1 (Local action and iterated exit law). *Consider the crowded family for some $n \geq 2$, and let i be even. For every $k \geq 1$ and all sufficiently small δ there are C^k coordinates $(z, w) \in \mathbb{R} \times \mathbb{R}^{3n-1}$ on a fixed neighbourhood of $X_{\circ}$ in the class, depending C^k on δ , such that the rate equations read*

$$\dot{z} = f(z, \delta) = -a \prod_{j=0}^{2n-2} (z - \delta \xi_j) (1 + O(|z| + \delta)), \quad \dot{w} = B(z, w, \delta) w, \quad (6.3)$$

with a as in (6.1) and $B(0, 0, 0)$ Hurwitz; $X_j(\delta)$ is the point $(\delta \xi_j, 0)$. Let $l < \xi_i < r_c$ be cuts with $\xi_{i-1} < l$ and $r_c < \xi_{i+1}$ (when these neighbours exist), let P_t be a positive definite matrix with $B(0, 0, 0)^T P_t + P_t B(0, 0, 0) < 0$, let $1 < \alpha_t < n$, and put

$$D_{\delta} = \{\delta l < z < \delta r_c, \quad w^T P_t w < \delta^{2\alpha_t}\}.$$

(i) The infimum I_{δ} of the action over all paths from $X_i(\delta)$ to the complement of D_{δ} satisfies

$$I_{\delta} = c_i \delta^{2n} (1 + O(\delta)), \quad c_i = \frac{2a}{b_0} \min_{s \in \{l, r_c\}} \int_{\xi_i}^s \prod_j (y - \xi_j) dy > 0, \quad (6.4)$$

where $b_0 = \omega^T Q(X_{\circ}) \omega$, $Q(x) = \sum_{\rho} a_{\rho}(x) \nu_{\rho} \nu_{\rho}^T$, and $\omega = dz$ at $X_{\circ}$ is the left null covector of the Jacobian normalized by the choice of z . (ii) Let τ_{δ} be the exit time from D_{δ} of the chain started at lattice points $X_0^{\Omega} \rightarrow X_i(\delta)$. Then for every $\zeta > 0$,

$$\lim_{\delta \downarrow 0} \liminf_{\Omega \rightarrow \infty} \Pr\left(\left|\frac{\log \tau_{\delta}}{\Omega \delta^{2n}} - c_i\right| < \zeta\right) = 1. \quad (6.5)$$

The proof is in Appendix B. The inner limit is taken at fixed δ ; the theorem is not uniform along joint sequences (δ, Ω) , and it concerns the declared tube D_{δ} , not the basin of attraction. The action in (i) is the full reaction-path action: paths may vary all enzyme, substrate and complex coordinates, and the transverse bound $\alpha_t < n$ is what excludes a cheaper escape through those coordinates.

Corollary 6.2 (Barrier, recovery and resolution). *In the crowded family, the linear recovery time $t_{\text{rec}} = 1/|\lambda_i|$, the gap ΔR between neighbouring values of any observable with nonzero derivative along the equilibrium curve, and the barrier satisfy*

$$t_{\text{rec}} = \Theta(\delta^{-(2n-2)}), \quad \Delta R = \Theta(\delta), \quad I_{\delta} = \Theta(\delta^{2n}) \implies I_{\delta} = \Theta(t_{\text{rec}}^{-n/(n-1)}) = \Theta((\Delta R)^{2n}),$$

with constants depending on the family.

Proof. The first relation is (6.1), the second is [8, Corollary 8.4] or follows from the analyticity of $X_j(\delta)$, and the third is (6.4). $\square$

The linear-noise estimate $\text{Var}(z) \simeq b_0/(2\Omega|\lambda_i|)$ gives a squared separation-to-noise ratio of order $\Omega \delta^{2n}$, in agreement with the exponent in (6.4); it is a consistency check, not a substitute for the theorem. These relations describe *one* mechanism by which many states can be expensive, namely crowding. They do not say that every multi-label source is crowded: the four labels of Section 8 are well separated, and their cost has other origins.

6.2 Observation

Proposition 6.3 (Bounded-error packing). *Suppose label i has ideal observation $y_i \in \prod_{j=1}^{d_{\text{obs}}} [\alpha_j, \alpha_j + R_j]$, the observed vector is $y_i + e$ with an arbitrary error satisfying $|e_j| \leq \varepsilon_j$, $\varepsilon_j > 0$, and a decoder identifies the label for every admissible error. Then the number of labels satisfies*

$$N_{\text{lab}} \leq \prod_{j=1}^{d_{\text{obs}}} \left(1 + \left\lfloor \frac{R_j}{2\varepsilon_j} \right\rfloor\right).$$

Proof. The closed error boxes around different ideal observations must be disjoint. Let $m_j = 1 + \lfloor R_j/2\varepsilon_j \rfloor > R_j/2\varepsilon_j$ and cut the j -th range into m_j consecutive intervals of length $R_j/m_j < 2\varepsilon_j$, assigning endpoints to one side. Two ideal observations in the same product cell differ by less than $2\varepsilon_j$ in every coordinate, so their boxes share the midpoint. Hence each cell contains at most one ideal observation. $\square$

Variation of the readout within the safe sets D_i enlarges the boxes. The bound is of order $(R/\varepsilon)^{d_{\text{obs}}}$, which is why several independent channels can help; it is an upper bound, not an attainment result, and correlated channels, or phosphoform fractions constrained to a simplex, give less.

Proposition 6.4 (Passive and bounded-gain observation). (a) *If an observation is a function of the chemical path, possibly randomized, that does not feed back on the chemistry, then the law of every chemical exit time, and hence every barrier of Theorem 6.1, is the same for all such observations.* (b) *If equilibria satisfy $\|x_i(\delta) - x_j(\delta)\| = O(\delta)$ and the observation map has a Lipschitz constant bounded independently of δ , then the observations differ by $O(\delta)$, for any number of channels.*

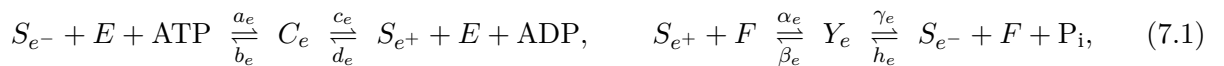
Proof. (a) The exit time is a functional of the chemical path, whose law is unchanged. (b) is the Lipschitz inequality. $\square$

A richer decoder can change which excursions count as failures; that defines a new event, which needs its own retention proof. A reporter that binds the substrate is not passive: it sequesters substrate and belongs in the reaction network. Thus multichannel readout is a promising direction for well-spread states, but it does not repair crowding.

7 Finite thermodynamic driving and robustness

The catalytic steps of (1.1) are irreversible, which corresponds to infinite driving by ATP hydrolysis. We show that the sinks survive in a thermodynamically consistent model with finite driving, and that effective affinities present no obstruction; see [30, 45, 46] for background and [38] for a study of finite driving in other biochemical switches.

Theorem 7.1 (Finite-driving persistence). *Let ATP, ADP and inorganic phosphate be chemostatted with activities 1, ϑ and ϑ , and replace (1.1) on every edge by*



with $d_e = a_e c_e / b_e$ and $h_e = \alpha_e \gamma_e / \beta_e$. (i) *There are standard chemical potentials for which every reaction of (7.1) satisfies local detailed balance; hence all Wegscheider cycle conditions hold, including those for cycles through different paths of the cube, and the chemical potential difference of ATP hydrolysis, $-2 \log \vartheta$ in units of $k_B T$, is the same for every turnover on every edge.* (ii)

The vector field of (7.1) is $H + \vartheta G$ with a polynomial G , and the totals (1.2) are still conserved. Consequently every finite set of hyperbolic sinks of (1.1) in one class continues, for all sufficiently small $\vartheta > 0$, to hyperbolic sinks of (7.1) in the same class, and disjoint readout neighbourhoods remain disjoint.

Proof. (i) Set the standard potentials, in units of $k_B T$, to zero for all free species S_I, E, F and for the three chemostatted species, to $\log(b_e/a_e)$ for C_e and to $\log(\beta_e/\alpha_e)$ for Y_e . Then each of the four reaction pairs has ratio of forward to backward constant equal to $\exp(-\Delta\mu^\circ)$: for the binding steps by definition, and for the catalytic steps because $c_e/d_e = b_e/a_e$ and $\gamma_e/h_e = \beta_e/\alpha_e$. Around any closed cycle of reactions the increments of a potential telescope, so the product of the ratios is one [48]. With the stated activities the hydrolysis potential is $\log[1/(\vartheta \cdot \vartheta)]$. (ii) The forward propensities are unchanged, and the new reverse propensities are $\vartheta d_e E S_{e+}$ and $\vartheta h_e F S_{e-}$. The C^1 convergence on compact sets is immediate, and hyperbolic equilibria and their spectra persist by the implicit function theorem. $\square$

Remark 7.2. (a) The three notions that occur in this paper are different: effective square balance (1.3), which concerns the irreversible network; detailed balance of the closed reversible network, which holds for (7.1) by construction; and the maintained hydrolysis potential. An arbitrary effective affinity is compatible with (i) because it compares kinetic pathways, not equilibrium constants. (b) The choice of zero standard potentials for the phosphoforms is a normalization of convenience; other potentials change d_e, h_e by bounded factors and leave (ii) unchanged. (c) The theorem concerns a chemostatted system. A finite closed fuel pool needs a depletion argument, and a quantitative design needs feasible chemostat concentrations and ATP throughput. (d) The finite-molecule constants of Theorem 5.2 do not transfer automatically: the added reactions change the drift, the noise and the equilibria, and the contract must be recomputed.

The following estimate makes “sufficiently small” explicit, and also gives a robustness margin for the rate constants.

Proposition 7.3 (Explicit perturbation bound). *Assume the hypotheses of Theorem 5.2 and let $D' = \{V \leq v/4\}$. Let G be a C^1 field on D' , tangent to the class, with $\|G\| \leq b_G$ and $\|DG\| \leq d_G$ there, and let $\ell_0 = \min\{v/(4\bar{p}), 1\}$. If*

$$0 \leq \vartheta \leq \min\left\{\frac{\mu}{8\bar{p}d_G}, \frac{\mu\ell_0}{8\bar{p}b_G}\right\}, \quad (7.2)$$

then $H + \vartheta G$ has exactly one equilibrium in D' , and it is a hyperbolic sink.

Proof. On D' we have $\|h\| \leq \varrho/2$. For the mass-action field the remainder is a quadratic form bounded by $K\|h\|^2$ through a bound on its bilinear form, so $\|DH(x^* + h) - J\| \leq 2K\|h\| \leq K\varrho$, and $DH^T P + PDH \leq -\mu I + 2\bar{p}K\varrho I \leq -\frac{\mu}{2}I$ on D' . Integrating along the segment from x^* gives $h^T PH(x^* + h) \leq -\frac{\mu}{4}\|h\|^2$. On $\partial D'$, $\|h\|^2 \geq v/(4\bar{p})$, so $\|h\| \geq \ell_0$ because $\min\{t, 1\} \leq \sqrt{t}$, and $\vartheta|h^T PG| \leq \vartheta\bar{p}b_G\|h\| \leq \frac{\mu}{8}\|h\|^2$. Thus the perturbed field points strictly inward on $\partial D'$, and a zero exists by degree. Moreover $D(H + \vartheta G)^T P + PD(H + \vartheta G) \leq -\frac{\mu}{2}I + 2\vartheta\bar{p}d_G I \leq -\frac{\mu}{4}I$ on the convex set D' , so the flow contracts the metric P there: the zero is unique and its Jacobian is Hurwitz by Lyapunov's theorem. $\square$

For independent relative errors $|\Delta k_\rho| \leq \vartheta k_\rho$ of the rate constants one may take $b_G = \sum_\rho \sup a_\rho \|\nu_\rho\|$ and the analogous bound for d_G ; for the reversible extension, G consists of the reverse catalytic reactions. In both cases rational upper bounds on D' are computable, and the perturbed equilibrium stays in D' , hence within the decoder interval of Corollary 5.3.

8 A certified three-site source with four labels

This section applies the theory to one explicit source. Every statement labelled “computer-assisted” is verified by exact rational arithmetic or outward-rounded interval arithmetic from the data files that accompany the paper, by the program described in Appendix C; approximate decimals are given only for orientation.

8.1 The source and its equilibria

Let $n = 3$. Encode $I \subseteq \{1, 2, 3\}$ by its bit mask $0, \dots, 7$, and order the twelve edges by increasing tail mask and then by increasing added site. The effective rates are

$$k^+ = (8, 1, 7, 1, 4, 4, 9, 8, 8, 8, 5, 6), \quad k^- = (5, 5, 7, 8, 6, 2, 2, 9, 7, 6, 2, 5), \quad (8.1)$$

which are not square balanced, and the totals are $(E_T, F_T, S_T) = (10, 1, 264\,490\,005\,424)$, so $r = 10$. The tree weights τ_I are the 7×7 cofactors of the matrix $uU + V$; they are integer polynomials of degrees 4, 5, 5, 6, 5, 6, 6, 7. The loadings $(p, q) \in \mathbb{Q}_{>0}^{24}$ are chosen in the affine space of solutions of

$$L(u) = B(u) - 10 D(u) = \prod_{j=1}^6 (u - j), \quad (8.2)$$

and the rate constants are $(a_e, b_e, c_e, \alpha_e, \beta_e, \gamma_e) = (p_e + k_e^+, 1, k_e^+/p_e, q_e + k_e^-, 1, k_e^-/q_e)$, which have the prescribed loadings and effective rates. We consider two solutions of (8.2). The *original source* adds a large multiple of the kernel vector of Lemma 4.1 to one signed solution. The *final source* was obtained by minimizing the largest loading over the whole solution space by linear programming and then repairing the result in exact arithmetic; its loadings lie between 2.0×10^{-4} and 5.4×10^{-2} , and its rate constants between 1 and 9999.0001. No optimality is claimed or needed. The two sources have the same polynomials A, L , and therefore by (2.7) the same free substrate concentrations at a given ratio u .

The design principle is visible in (2.7): $s = (10 - u)/(uL(u))$ must be positive, so for $u < 10$ equilibria can only lie in the intervals $(0, 1)$, $(2, 3)$, $(4, 5)$ and $(6, 10)$, where $L > 0$, and the substrate total $g(u)$ along the equilibrium curve tends to infinity at every zero of L (Figure 2). For large S_T this gives two equilibria in each of the first three intervals and one in the last.

Proposition 8.1 (Computer-assisted: seven equilibria, four sinks). *For the final source:*

- (a) *The eliminant (2.8) has degree 15 and exactly eight positive roots, all simple. Seven of them are admissible,*

$$u \approx 1.529 \times 10^{-7}, 0.99988, 2.00464, 2.96146, 4.14463, 4.83103, 6.05627,$$

and the eighth, $u \approx 10.000002 > r$, is not. Hence the class contains exactly seven positive equilibria, and all are nondegenerate.

- (b) *The first, third, fifth and seventh are hyperbolic sinks: for each there is an explicit rational symmetric matrix P_i and a rational lower-triangular matrix R_i with positive diagonal such that $R_i P_i R_i^T$ and $-(J^T P_i + P_i J)$ are strictly diagonally dominant with positive diagonal, with margins above 9/10, for every Jacobian J of the 31-dimensional class chart evaluated on an isolating interval of the root.*

- (c) *The other three equilibria have index -1 and are unstable.*

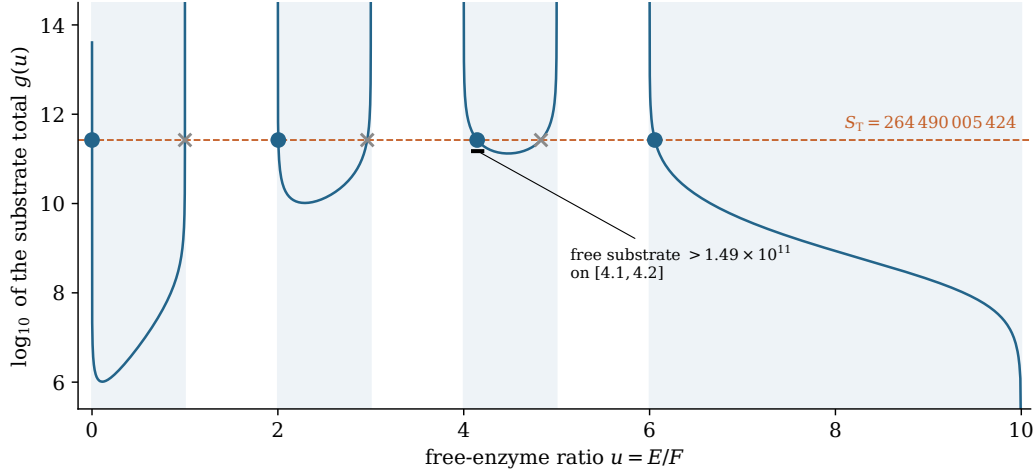


Figure 2: The final source. The substrate total $g(u)$ along the curve of equilibria with $(E_T, F_T) = (10, 1)$, on the four admissible intervals (shaded), and the seven equilibria at $S_T = 264\,490\,005\,424$: sinks (dots) on the decreasing parts and unstable equilibria (crosses) on the increasing parts. The short black bar is the lower bound of Proposition 8.3 for the free substrate on $4.1 \leq u \leq 4.2$. The curve is drawn from high-precision floating-point evaluation; the positions and the stability of the equilibria are certified separately.

In particular $C_3 \geq 4$. The same statements hold for the original source.

Proof. (a) and (b) are verified by the replay program: real-root isolation for the integer eliminant, the sign conditions of (2.7) at the roots, and interval evaluation of the state, the Jacobian and the two matrix inequalities. Diagonal dominance of $R_i P_i R_i^T$ gives $P_i > 0$ since R_i is invertible, and Lyapunov's theorem gives that J is Hurwitz. For nondegeneracy, note that along the curve of equilibria with E_T, F_T fixed, $\mathcal{P}(u) = (S_T - g(u))uLM$, so a root with $LM \neq 0$ is simple exactly when $g'(u) \neq 0$. In the coordinates (u, s, F) of Lemma 2.3 the derivative of (E_T, F_T) with respect to (s, F) has determinant $uFM \neq 0$ by (2.6), so $\mathcal{T}|_{\mathcal{Z}}$ is regular at the equilibrium if and only if $g'(u) \neq 0$. (c) By Lemma 2.4 the seven indices add up to one; four of them are $+1$, so the remaining three, which are ± 1 , are all -1 . Then $\det(-J) < 0$, so J has a positive real eigenvalue. $\square$

High-precision eigenvalue computations (diagnostics, not used in the proof) agree: the three unstable equilibria have exactly one unstable direction. Double-precision eigenvalues are unreliable for this source because of its extreme separation of time scales, which is why the certificates use exact arithmetic throughout.

8.2 Readout and operating contract

As readout take the total amount of fully phosphorylated substrate, $R = S_{[3]} + \sum_{e+= [3]} Y_e$. At the four sinks

$$R/S_T \approx 3.31 \times 10^{-21}, \quad 0.316355, \quad 0.558771, \quad 0.670451, \quad (8.3)$$

with all gaps larger than 0.1116 (Figure 3). The first value is positive, not zero. The separation of the four labels is therefore *moderate*: this example is not crowded in the sense of Section 6, and its cost must not be attributed to four states being squeezed into a scalar interval.

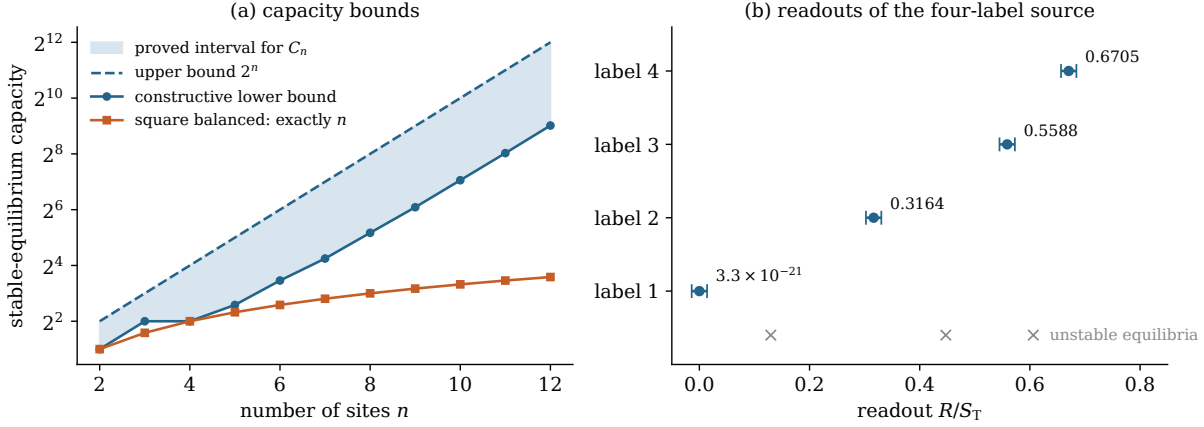


Figure 3: (a) Theorem 1.2: the proved interval for C_n , the constructive lower bound (including $C_3 \geq 4$) and the exact capacity n of square-balanced rates. No uncomputed value of C_n is plotted. (b) Readouts (8.3) of the four sinks of the final source with the observation tolerance of Proposition 8.2 (one eighth of the smallest gap), and the readouts of the three unstable equilibria.

Proposition 8.2 (Computer-assisted: operating contract). *For the final source with all rate constants divided by 50000000/10001 (the model clock), let P_i be as in Proposition 8.1 and define $\bar{p}_i, \underline{p}_i, \mu, K, \varrho_i$ by the rational formulas of Appendix C, with ε_{obs} equal to one eighth of the smallest readout gap and C_i the local bound (8.4). Then the hypotheses of Theorem 5.2 and Corollary 5.3 hold with*

$$\Omega = \lceil \text{right-hand side of (5.7)} \rceil + 1 \approx 1.31 \times 10^{100}.$$

Every retention bound is below 0.00105 and every recovery bound below 0.008987. Four explicit vectors of 34 nonnegative integer counts with the exact totals $(10\Omega, \Omega, S_T\Omega)$ lie in the preparation sets $\mathbf{B}_i$.

The constant C_i in Theorem 5.2 may be bounded crudely, using $S_I \leq S_T$, $E, C_e \leq E_T$, $F, Y_e \leq F_T$, or locally, using that on the ellipsoid D_i every species satisfies

$$x_j \leq \bar{x}_j^* + \|\mathcal{T}_j\|_1 \varrho_i, \quad (8.4)$$

where $\bar{x}_j^*$ is an upper bound of the equilibrium value and $\mathcal{T}_j$ the row of the affine chart map for species j . The local bound changes nothing else in the proof of Theorem 5.2, which already works with the stopped process. It lowers the sufficient Ω by a factor 5.0×10^6 for the original source and 5.4×10^9 for the final source.

8.3 Physical units and certified resources

To compare with laboratory scales, let the concentration unit be $C_0 = 10^{-6} \text{ M}/S_T$, so that the total substrate concentration is $1 \mu\text{M}$, and let the time unit be 1 s. An association constant a then corresponds to a/C_0 and a unimolecular constant b to $b\text{ s}^{-1}$. We slow each source uniformly until its largest association constant is $10^6 \text{ M}^{-1}\text{ s}^{-1}$ and its largest unimolecular constant at most 100 s^{-1} ; for the final source the factor is 2.4×10^{12} . By Proposition 5.5 this changes all times and leaves Ω unchanged. The volume is $\Omega/(N_A C_0)$ with $N_A = 6.02214076 \times 10^{23} \text{ mol}^{-1}$. Table 4 lists the result.

Relative to the first column, the final source lowers the sufficient number of molecules by a factor of about 10^{32} and the certified recovery time by 3.6×10^7 . These are large improvements of

Table 4: Certified *sufficient* resources for the contract of Corollary 5.3, with total substrate $1\mu\text{M}$ and the common upper limits on rate constants. Entries are base-10 logarithms. Each storage horizon is $100t_{\text{rec}}$ of its own source, so the horizons are not the same absolute duration. None of the columns describes a feasible assay.

Quantity	Original, global C_i	Original, local C_i	Final, local C_i
Substrate molecules	143.534	136.835	111.541
Kinase molecules	133.111	126.413	101.118
Phosphatase molecules	132.111	125.413	100.118
Volume (litres)	125.754	119.055	93.761
Recovery time t_{rec} (seconds)	38.706	38.706	31.152
Storage horizon T (seconds)	40.706	40.706	33.152

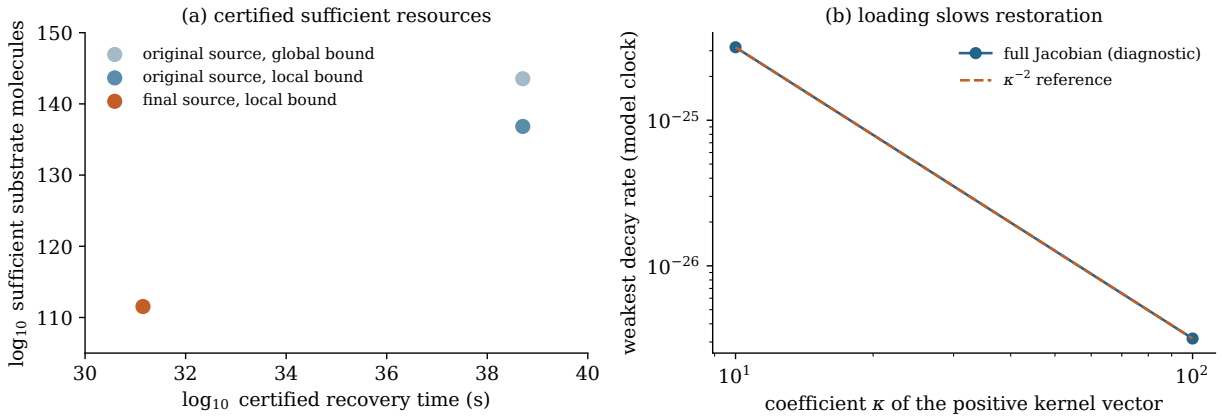


Figure 4: (a) The three columns of Table 4; the points are certified sufficient values under the protocol $T = 100t_{\text{rec}}$, not an optimized trade-off curve. (b) Floating-point diagnostic for Remark 4.5(a): with the effective rates and the target (8.2) fixed, increasing the multiple κ of the kernel vector from 10 to 100 lowers the weakest decay rate at the third sink by a factor 99.999. Neither panel is a stochastic simulation or a calibrated prediction.

an existence certificate, and they leave it far from any experiment: 3.5×10^{111} substrate molecules, 5.8×10^{93} litres and a recovery time of 1.4×10^{31} seconds. For comparison, $20\mu\text{L}$ at $1\mu\text{M}$ contains 1.2×10^{13} substrate molecules. This is a comparison with a sufficient bound; it does not show that a small volume cannot hold four labels in another design. Units cannot remove the problem: at $1\mu\text{M}$ substrate the phosphatase concentration is $1\mu\text{M}/S_T \approx 3.8 \times 10^{-18}\text{M}$, and after the slowing the largest unimolecular constant is about $4 \times 10^{-9}\text{s}^{-1}$.

As an exploratory screen for a cell-free assay one may ask for enzyme-to-substrate ratios between 0.01 and 1, association constants between 10^3 and $10^6\text{M}^{-1}\text{s}^{-1}$, unimolecular constants between 0.01 and 100s^{-1} , recovery within 10^3s , and a normalized readout error of 0.01; in vitro studies of ERK2 and its phosphatases work at comparable enzyme and substrate concentrations over minutes [51]. These are design assumptions, not measured ranges for the 72 constants of this source, and no such study provides a calibrated four-label error law. Both sources fail the screen by many orders of magnitude in the ratios, in the lower kinetic limits and in the recovery time.

8.4 Where the cost comes from

Several losses are distinct, and only some of them are artefacts of the analysis.

Intrinsic slowness. High-precision eigenvalues (diagnostics) give weakest decay rates of 1.4×10^{-17} , 1.3×10^{-27} , 9.1×10^{-29} and $4.4 \times 10^{-28} \text{ s}^{-1}$ at the four sinks of the final source in the slowed physical clock. This is the η/κ^2 mechanism of Remark 4.5, made worse by the association limit (Figure 4b). No Lyapunov function can certify faster recovery than the dynamics has.

Certificate losses. The certified rates λ_i are smaller than the intrinsic ones by factors between 59 and 246. The bounds $\bar{p}_i/\underline{p}_i$ on the condition numbers of the metrics P_i lie between 10^{19} and 10^{31} , and the radii ϱ_i between 10^{-22} and 10^{-17} in the chart; both enter (5.7) through $v_i = \underline{p}_i \varrho_i^2$. Global propensity bounds cost a further factor between 5×10^6 and 3×10^{10} in C_i , which (8.4) removes. Better metrics, or exponential rather than quadratic test functions, could reduce these losses. They cannot repair the first item or the next.

Concentration geometry. In both sources the enzymes are almost completely bound and nearly all substrate is free, and $E_T/S_T = 3.8 \times 10^{-11}$. For the family of designs to which both sources belong this is forced:

Proposition 8.3 (Ratio obstruction). *Fix the effective rates (8.1), the target (8.2) and $(E_T, F_T) = (10, 1)$. For every choice of positive loadings satisfying (8.2), and every S_T , an equilibrium with $4.1 \leq u \leq 4.2$ has free substrate*

$$\sum_I S_I = \frac{(10-u)A(u)}{u \prod_{j=1}^6 (u-j)} \geq \frac{2260898025826008319}{15168384} > 1.49 \times 10^{11}, \quad \text{hence} \quad \frac{E_T}{S_T} < 6.8 \times 10^{-11}.$$

Proof. By (2.7), $\sum_I S_I = sA$ depends on the loadings only through L , which is fixed. On $[4.1, 4.2]$ we have $L > 0$ and $L(u) \leq \prod_j \max\{|4.1-j|, |4.2-j|\}$, while $A(u) \geq A(4.1)$ because A has nonnegative coefficients, and $(10-u)/u \geq 5.8/4.2$. The resulting rational number is the one displayed, and $S_T \geq \sum_I S_I$. $\square$

The third sink of both sources lies in this interval. Thus no choice of loadings, and in particular no shift along the kernel of Lemma 4.1, brings this family into the range of enzyme-to-substrate ratios of the screen. The proposition is an obstruction for the stated family only. The effective rates, the position of the roots of the target relative to the tree weights, the retained branch and the enzyme totals are all outside its hypotheses; changing the enzyme totals, for instance, alters the shape of $g(u)$ and hence the set of coexisting equilibria, and has not been explored. Multiplying all tree weights and the target by a common factor changes nothing.

8.5 Robustness and driving thresholds for the example

Proposition 8.4 (Computer-assisted). *For the final source, with rational bounds b_G, d_G computed on the ellipsoids D_i : (a) every rate list whose 72 constants differ from the given ones by relative errors at most 10^{-99} has four hyperbolic sinks, one in each D'_i , with pairwise disjoint decoder intervals; (b) the reversible extension (7.1) has four such sinks for every $0 < \vartheta \leq 10^{-102}$, that is, for a hydrolysis potential of $204 \ln 10 \approx 469.7$ in units of $k_B T$.*

Proof. Proposition 7.3 with the certified minima 4.2×10^{-99} and 4.0×10^{-102} of the right-hand side of (7.2) over the four sinks. $\square$

These margins measure the conditioning of the example and the conservatism of Proposition 7.3. They are not experimentally meaningful tolerances, the threshold is not a minimum free-energy requirement (cellular ATP hydrolysis provides roughly $20 k_B T$ [45]), and the finite-molecule contract has not been recomputed for the perturbed or the reversible source.

9 Discussion

Which kinetic structure permits many states. The capacity of the distributive cube is governed by the cycle structure of its *effective* rates. If every elementary square is balanced, the equilibria are those of a single chain and at most n of them are stable, regardless of the 2^n phosphoforms and of the loadings. If squares are unbalanced, the loadings become an exponentially large space of design parameters for the scalar equation on the equilibrium curve, and $\Theta(2^n)$ stable states are possible. The proofs keep these two ingredients apart: the number of crossings is programmed linearly through the loadings, while transverse stability comes from the saturated family and does not depend on the programming. Rank alone, a scalar root count alone, or a Hamiltonian path through the cube would not establish stable capacity for the full mass-action network.

Structure is not operation. Theorem 5.2 shows that any finite set of hyperbolic sinks becomes a set of reliable labels at some finite system size, so stable-equilibrium capacity is an upper bound on what can be stored in equilibrium labels and is attainable in principle. The price is not controlled by the capacity theorems. In crowded families it is controlled by Theorem 6.1 and Corollary 6.2; in the loading construction, by the η/κ^2 slowing and the positivity burden of Remark 4.5; and in the worked example, additionally by very conservative certificates. We have tried to keep three kinds of statement apart: theorems with explicit scope, exact computations for one source, and floating-point diagnostics. In particular the figures 10^{11} molecules and 10^{31} seconds are sufficient for our certificate and are not claimed to be necessary, even for this source.

What the example teaches for design. The obstruction of Proposition 8.3 and the diagnostics of Section 8 suggest that the next design step is not a finer readout, a larger volume or a sharper probability bound, but a different equilibrium geometry: effective rates and a target whose retained branches have moderate free-substrate levels at moderate enzyme totals, found by continuation of the common-total branches while monitoring the smallest readout gap, the weakest eigenvalue and the size of the loadings. A low-parameter family with four robust sinks would be more informative than another isolated point among 72 rate constants. Whether a three-site substrate with shared enzymes can hold four robust, readable and quickly recovering states at physical rates and inventories is open, and nothing in this paper suggests that it is impossible.

Open problems. (1) Determine C_n ; already $4 \leq C_3 \leq 8$ is open, as is the maximal number of positive equilibria, between 7 and 15 for $n = 3$. (2) Find an intermediate structural condition, weaker than square balance, that bounds the capacity polynomially in n . (3) Prove an exit law for crowded states that is uniform along joint sequences (δ, Ω) , or a global basin-exit law. (4) Replace the quadratic certificates by test functions adapted to the slow direction; preliminary computations with exponential test functions improve the retention part of the budget by about nine orders of magnitude but not the recovery part. (5) Recompute the finite-molecule contract for the reversible extension at a declared hydrolysis potential, with ATP turnover accounted for. (6) Analyse a writing protocol. (7) Extend the capacity results to mechanisms with shared complexes, processive steps or several enzymes.

Data and code availability. The source bundle of this paper contains the exact rational data of the example, the certificate replay program, the figure scripts and the three Lean files; see Appendix C.

A The sequential input

We recall what is used from the companion paper [8] on the sequential chain $S_0, \dots, S_n$. Its Theorem 1.1 states that every compatibility class contains at most n sinks (nonhyperbolic ones included), and that for every $n \geq 1$ there are positive rational rate constants and totals with exactly $2n - 1$ positive equilibria, all hyperbolic, of which n are sinks and $n - 1$ are saddles with a one-dimensional unstable manifold. The bound $2n - 1$ on the number of equilibria is due to Wang and Sontag [56].

The construction is as follows. For auxiliary numbers $x_0, \dots, x_{2n-2} > 1$ a positive polynomial recurrence produces a chain whose $2n - 1$ equilibria, at the totals $(2r, 2, 2r + 2)$, have the free enzyme concentrations

$$E = \frac{x_j - 1}{2}, \quad F = \frac{4}{x_j + 1}, \quad u = \frac{x_j^2 - 1}{8}. \quad (\text{A.1})$$

Positivity of all coefficients holds whenever $r > \max_j u(x_j)$ and $4r + \sqrt{1 + 8r} \geq \sum_j x_j$, by an interlacing argument. The loadings and effective rates are determined by the equilibrium data, while n catalytic scales and the dissociation constants remain free. Stability is obtained at the *coalesced* system $x_j \equiv 3$, where all equilibria coincide at a point X_o with $E = F = 1$: after an acceleration of binding as in (4.4), the slow dynamics on the retained pools is governed by an $n \times n$ matrix whose leading $(n - 1)$ -block converges, as $r \rightarrow \infty$, to a path matrix with a rank-one positive feedback of gain strictly less than one, hence is Hurwitz, and a small final catalytic current keeps the remaining eigenvalue at zero. The outcome is [8, Lemma 7.2]: for suitable positive constants the class Jacobian at X_o has an algebraically simple eigenvalue 0 and $3n - 1$ eigenvalues in the open left half-plane. With these constants frozen, the *split family* takes $x_j = 3 + \delta \xi_j$ for real $\xi_0 < \dots < \xi_{2n-2}$. All data are real-analytic in δ , the totals are fixed, and [8, Theorem 1.2] gives the splitting law (6.1), with the constant $a > 0$ of [8, (7.2)], from an exact formula for the determinant of the class Jacobian. Since $\prod_{k \neq j} (\xi_j - \xi_k)$ has the sign $(-1)^j$, the even-indexed equilibria are the sinks. By (A.1), $E = 1 + \delta \xi_j / 2$ at $X_j(\delta)$. For $n = 1$ the equilibrium is unique and globally attracting [3], and [8] exhibits rate constants for which it is hyperbolic.

B Proof of Theorem 6.1

B.1 Phase and fibre coordinates

Extend the split family by the equation $\dot{\delta} = 0$. At $(X_o, 0)$ the extended linearization has a two-dimensional centre space, on which it vanishes because the zero eigenvalue is simple and δ is a parameter, and a Hurwitz complement. For every k there is a C^k centre manifold; it is normally attracting and contains all equilibria $X_j(\delta)$ for small δ [6]. On a small neighbourhood the flow on the centre manifold and its inverse grow at most like $e^{\varepsilon_0 |t|}$, while the stable directions contract at a rate $\beta_0 > 0$, and $(k + 1)\varepsilon_0 < \beta_0$ for a small enough neighbourhood. Under this spectral gap condition the stable foliation of the centre manifold is C^k [24, 31, 19]. If a compact invariant base is required by the cited theorem, one modifies the nonlinear terms outside the neighbourhood, which can be done with arbitrarily small C^1 norm because they vanish to first order, and restricts afterwards. Let π be the projection along the stable fibres onto the centre manifold of parameter δ . We take

$$z = 2(E \circ \pi) - 2.$$

This is an admissible coordinate on the one-dimensional centre manifold because the centre direction at X_o is tangent to the equilibrium curve (A.1), along which $dE/dx = 1/2 \neq 0$; and $z = \delta \xi_j$ at

$X_j(\delta)$ exactly. Choosing fibre coordinates w that vanish on the centre manifold gives (6.3) with $\dot{z} = f(z, \delta)$ independent of w , by invariance of the foliation, and with $B(0, 0, 0)$ Hurwitz.

Since $f(\delta\xi_j, \delta) = 0$ for all j , repeated division (Hadamard's lemma, with k large) gives $f(z, \delta) = g(z, \delta) \prod_j (z - \delta\xi_j)$ with g continuous. The eigenvalue of the class Jacobian along the centre manifold at $X_j(\delta)$ is $\partial_z f(\delta\xi_j, \delta) = g(\delta\xi_j, \delta) \delta^{2n-2} \prod_{k \neq j} (\xi_j - \xi_k)$, and comparison with (6.1) gives $g(0, 0) = -a$. This is a factorization of the *dynamics* on the centre manifold, not only a statement about its zero set.

B.2 Lower bound for the action

Put $\omega(x) = D(z \circ \pi)(x)$, $b(x) = \omega(x)Q(x)\omega(x)^T$ and $b_0 = b(X_0)$. Near the interior point X_0 all propensities are positive, and the reaction vectors span the class, so Q is positive definite there and $b_0 > 0$; moreover the cone generated by the reaction vectors is the whole class, since every binding step has its dissociation and a substrate transfer can be reversed through the opposite enzyme. Hence $\mathcal{H}(x, \cdot)$ in (6.2) is strictly convex and superlinear, and

$$\mathcal{H}(x, \theta) = H(x) \cdot \theta + \frac{1}{2}\theta^T Q(x)\theta + O(|\theta|^3), \quad \mathcal{I}(x, H(x) + \chi) = \frac{1}{2}\chi^T Q(x)^{-1}\chi + O(|\chi|^3)$$

uniformly near X_0 . Let $U(z) = -\int_{\delta\xi_i}^z f(y, \delta) dy$. Because i is even, $U \geq 0$ between the cuts, and by (6.3)

$$U(\delta s) = a \delta^{2n} \int_{\xi_i}^s \prod_j (y - \xi_j) dy (1 + O(\delta)), \quad s \in \{l, r_c\}. \quad (\text{B.1})$$

Let $W_1 = 2U(z \circ \pi)/b_\delta$ with $b_\delta = b_0 + C_1\delta$. Then $DW_1 = -(2f/b_\delta)\omega$, and since $\omega \cdot H = f$ exactly,

$$\mathcal{H}(x, DW_1) = -\frac{2f^2}{b_\delta} \left(1 - \frac{b(x)}{b_\delta}\right) + O(|f|^3) \leq 0$$

on D_δ if C_1 is large, because $b(x) = b_0 + O(\delta)$ and $|f| = O(\delta^{2n-1})$ there. The exact autonomy of the phase is essential for this inequality at points with $w \neq 0$. For every path ϕ the Fenchel inequality $\mathcal{I}(\phi, \dot{\phi}) \geq DW_1(\phi) \cdot \dot{\phi} - \mathcal{H}(\phi, DW_1(\phi))$ gives an action of at least the increase of W_1 ; a path leaving D_δ through a phase cut therefore costs at least $2U(\delta s)/b_\delta$. For exits through the transverse boundary use $W_2 = \kappa_t w^T P_t w$. Its derivative along the flow is at most $-c\kappa_t \|w\|^2$ for small δ , and the quadratic and higher terms of $\mathcal{H}(x, DW_2)$ are at most $C\kappa_t^2 \|w\|^2$; with κ_t small, $\mathcal{H}(x, DW_2) \leq 0$, and a transverse exit costs at least $\kappa_t \delta^{2\alpha_t}$, which exceeds δ^{2n} in order because $\alpha_t < n$. Every exit from D_δ is of one of the two kinds.

B.3 A matching path

Consider the controlled field $H + \chi$ with $\chi(x) = -(2f(z, \delta)/b(x))Q(x)\omega(x)^T$. Its phase satisfies $\dot{z} = f - 2f = -f$ exactly, so the phase moves from near $\delta\xi_i$ to either cut. The fibre equation is forced by a term of order $|f| = O(\delta^{2n-1})$, so $\|w\| = O(\delta^{2n-1})$ stays inside the transverse tube because $2n-1 > \alpha_t$. The cost rate is $\frac{1}{2}\chi^T Q^{-1}\chi + O(|\chi|^3) = 2f^2/b(x) + O(|f|^3)$, and with $dt = dz/(-f)$ the total cost is $2U(\delta s)(1 + O(\delta))/b_0$. The path is started at a point arbitrarily close to the equilibrium, which it leaves in finite time; joining that point to the equilibrium costs arbitrarily little because Q is positive definite. Together with the lower bound and (B.1), this proves (6.4).

B.4 Exit times from the generator

Fix $\delta > 0$, and let $J_- = \min\{2U(\delta l)/b_\delta, 2U(\delta r_c)/b_\delta, \kappa_t \delta^{2\alpha_t}\}$ and $J_+ = I_\delta$.

Lower tail. Enlarging C_1 and decreasing κ_t slightly makes $\mathcal{H}(x, DW_j)$ strictly negative on $\{W_j \geq \varepsilon\} \cap \overline{D_\delta}$, for any fixed $\varepsilon > 0$. A Taylor expansion on the set of states reachable by one jump from D_δ gives $e^{-\Omega W_j} \mathcal{L}_\Omega e^{\Omega W_j} = \Omega \mathcal{H}(x, DW_j) + O_\delta(1)$, which is nonpositive on $\{W_j \geq \varepsilon\}$ for large Ω and at most $C_\delta \Omega$ elsewhere. Dynkin's formula for the stopped process and Markov's inequality, applied to $e^{\Omega W_1}$ and $e^{\Omega W_2}$ separately, give for initial states with $W_j(X_0) \leq \varepsilon$

$$\Pr(\tau_\delta \leq t) \leq C'_\delta (1 + C_\delta \Omega t) e^{-\Omega(J_- - 2\varepsilon)}; \quad (\text{B.2})$$

the overshoot changes the boundary values of W_j by $O_\delta(1/\Omega)$, which is absorbed in C'_δ . Hence $\Pr(\tau_\delta \leq e^{\Omega(J_- - 4\varepsilon)}) \rightarrow 0$.

Upper tail. The rate function has the flux representation

$$\mathcal{I}(x, y) = \inf \left\{ \sum_\rho \left(q_\rho \log \frac{q_\rho}{a_\rho(x)} - q_\rho + a_\rho(x) \right) : q_\rho \geq 0, \sum_\rho q_\rho \nu_\rho = y \right\}$$

[49]. Approximate the path of the previous subsection by a path on a finite time interval with bounded positive fluxes $q_\rho(t)$, cost at most $J_+ + \varepsilon$, that ends at a positive distance outside D_δ . Under the change of measure with intensities Ωq_ρ , the compensated counts divided by Ω have quadratic variation $O(1/\Omega)$, so by Doob's and Gronwall's inequalities the process follows the path with probability tending to one, and the logarithm of the likelihood ratio, divided by Ω , converges in probability to minus the cost, all logarithms being bounded on the compact positive tube. Hence, from every starting point in a small neighbourhood G_0 of the sink, the original chain exits within the duration of the path with probability at least $\frac{1}{2}e^{-\Omega(J_+ + 2\varepsilon)}$ for large Ω . Every deterministic trajectory from $\overline{D_\delta}$ enters G_0 within a common finite time, since the phase field points inward and has the single zero $\delta \xi_i$ between the cuts and the fibres contract; by the same martingale estimate the chain either exits or reaches G_0 within that time with probability at least $1/2$. The strong Markov property then gives, for a block length h_δ ,

$$\Pr(\tau_\delta > t) \leq \left(1 - \frac{1}{4}e^{-\Omega(J_+ + 2\varepsilon)}\right)^{\lfloor t/h_\delta \rfloor}, \quad (\text{B.3})$$

which tends to zero for $t = e^{\Omega(J_+ + 3\varepsilon)}$. This direct argument avoids applying a general exit theorem for unconstrained diffusions or jump processes to a process confined to a conserved class; the general large-deviation theory of reaction networks has hypotheses that would have to be verified [1].

Conclusion. By (6.4) and (B.1), $J_\pm = c_i \delta^{2n} (1 + O(\delta))$. Given $\zeta > 0$, choose first δ small, then $\varepsilon < \zeta \delta^{2n}/10$, and then Ω large in (B.2) and (B.3); lattice points converging to $X_i(\delta)$ satisfy $W_j(X_0^\Omega) \leq \varepsilon$ and lie in G_0 eventually. This proves (6.5). No constant above is controlled uniformly as $\delta \rightarrow 0$, which is why the limit is iterated. $\square$

C Certificates, replay, and machine-checked algebra

C.1 Constants of the operating contract

All quantities are rational. The chart has $d = 31$ coordinates. Let P_i, R_i be the saved matrices of Proposition 8.1, let $c = 50000000/10001$ be half of the largest rate constant, and divide all rate constants by c . Then

- $\bar{p}_i = \|P_i\|_\infty$, and $\underline{p}_i = \frac{9}{10} (\|R_i\|_\infty \|R_i\|_1)^{-1}$, because $R_i P_i R_i^T \geq \frac{9}{10} I$ implies $P_i \geq \frac{9}{10} (R_i^T R_i)^{-1}$;
- $\mu = \frac{9}{10}/c$, from the margin of $-(J^T P_i + P_i J)$ at the unscaled rates;

- $K = \sum 12 k \|\nu\|_1 g$, summed over the 24 binding reactions with constant k , where $g = 31$ if the substrate reactant is $S_\emptyset$ and $g = 1$ otherwise. This bounds the bilinear form of the quadratic part of H : the deviation of E or F is a sum of 12 chart coordinates and that of $S_\emptyset$ a sum of 31;
- $\varrho_i = \min\{\min_j x_j^*/(4d), \text{gap}/16, \mu/(4\bar{p}_i K)\}$, which keeps the ball in the positive class and the readout, a sum of four chart coordinates, within $\text{gap}/8$ of its value at the sink;
- C_i by (8.4); $v_i, \lambda_i, t_{\text{rec}}, T$ and Ω as in Theorem 5.2 and Corollary 5.3.

The integer preparations are verified at 220 digits after refining the isolating intervals by exact bisection. The bounds b_G, d_G of Proposition 8.4 are sums over reactions of $k\|\nu\|_1$ times products of the bounds (8.4), and of the corresponding derivatives.

C.2 Files and replay

The data files are: `postproof_candidate_interface.json` (loadings, rate constants, integer eliminant); `postproof_candidate_sinks.json` (isolating intervals, P_i, R_i); `postproof_candidate_operating.json` and `postproof_candidate_local_certificate.json` (contract constants, Ω , robustness thresholds); `postproof_candidate_preparations.json` (integer count vectors); `tree_weights.json`; and the corresponding files for the original source. The program `check_paper.py` recomputes, without reading any intermediate result, the tree weights from (8.1) by Laplacian cofactors, the identity (8.2), the eliminant from (2.8), the root census, the four interval Lyapunov certificates, all constants of the previous subsection, the budget (5.7), the error bounds, the inventories and preparation inequalities of the four count vectors, the entries of Table 4, Proposition 8.3 and the thresholds of Proposition 8.4. It also checks symbolically the elimination identity behind (2.8), the square weights (4.2) and the expansion (4.8). It performs no search, optimization, eigenvalue computation or simulation; every decision is made in exact rational arithmetic or in integer interval arithmetic with outward rounding. It needs Python with SymPy and runs in about one minute (745 checks); the option `--full` (776 checks) also replays the four sink certificates of the original source. The script `diagnostics.py` produces the floating-point diagnostics quoted in Section 8, which are marked as such, and `make_figures.py` the figures. How the certificates were found (high-precision eigenvectors for P_i , a linear program for the loadings) is irrelevant for their validity.

C.3 Machine-checked algebra

Three small Lean 4 files [15, 52] check algebraic identities used in the text: **TreeTotals** (the relations $M = L + (r - u)D$ and $suL = r - u$, and the positivity of the derivative used at $u = r$ in Proposition 2.1); **UniformLoading** (the kernel identity of Lemma 4.1 in scalar form, the cancellation of the zero-order term and the form of the numerator in (4.8)–(4.9), and the determinant -31410 in Lemma 4.2); and **SymmetricAggregation** (the multiplicity identity of Propositions 3.2 and 5.1 and the scalar form of (5.6)). They compile without `sorry` and with warnings treated as errors. They are routine identities. Sard’s theorem, degree theory, the stability constructions, invariant foliations, the large-deviation estimates and the operating contract are *not* formalized, and the interval certificates are checked by the Python program above, not by the Lean kernel.

References

- [1] A. Agazzi, A. Dembo, and J.-P. Eckmann. Large deviations theory for Markov jump models of chemical reaction networks. *Ann. Appl. Probab.*, 28(3):1821–1855, 2018. doi:[10.1214/17-AAP1344](https://doi.org/10.1214/17-AAP1344).
- [2] D. F. Anderson and T. G. Kurtz. *Stochastic Analysis of Biochemical Systems*. Mathematical Biosciences Institute Lecture Series. Springer, Cham, 2015.
- [3] D. Angeli and E. D. Sontag. Translation-invariant monotone systems, and a global convergence result for enzymatic futile cycles. *Nonlinear Anal. Real World Appl.*, 9(1):128–140, 2008. doi:[10.1016/j.nonrwa.2006.09.006](https://doi.org/10.1016/j.nonrwa.2006.09.006).
- [4] W. Bialek. Stability and noise in biochemical switches. In *Advances in Neural Information Processing Systems 13*, pages 103–109. MIT Press, 2001. arXiv:[cond-mat/0005235](https://arxiv.org/abs/cond-mat/0005235).
- [5] F. Bihan, A. Dickstein, and M. Giaroli. Lower bounds for positive roots and regions of multistationarity in chemical reaction networks. *J. Algebra*, 542:367–411, 2020. doi:[10.1016/j.jalgebra.2019.10.002](https://doi.org/10.1016/j.jalgebra.2019.10.002); arXiv:[1807.05157](https://arxiv.org/abs/1807.05157).
- [6] J. Carr. *Applications of Centre Manifold Theory*, volume 35 of *Applied Mathematical Sciences*. Springer, New York, 1981.
- [7] P. Cohen. The regulation of protein function by multisite phosphorylation – a 25 year update. *Trends Biochem. Sci.*, 25(12):596–601, 2000. doi:[10.1016/S0968-0004\(00\)01712-6](https://doi.org/10.1016/S0968-0004(00)01712-6).
- [8] Sharp equilibrium and stable-state capacities of multisite phosphorylation. Companion manuscript, 2026.
- [9] Attracting oscillations in sequential distributive multisite phosphorylation. Companion manuscript, 2026.
- [10] C. Conradi, E. Feliu, M. Mincheva, and C. Wiuf. Identifying parameter regions for multistationarity. *PLoS Comput. Biol.*, 13(10):e1005751, 2017. doi:[10.1371/journal.pcbi.1005751](https://doi.org/10.1371/journal.pcbi.1005751); arXiv:[1608.03993](https://arxiv.org/abs/1608.03993).
- [11] C. Conradi, D. Flockerzi, and J. Raisch. Multistationarity in the activation of a MAPK: parametrizing the relevant region in parameter space. *Math. Biosci.*, 211(1):105–131, 2008. doi:[10.1016/j.mbs.2007.10.004](https://doi.org/10.1016/j.mbs.2007.10.004).
- [12] C. Conradi and M. Mincheva. Catalytic constants enable the emergence of bistability in dual phosphorylation. *J. R. Soc. Interface*, 11(95):20140158, 2014. doi:[10.1098/rsif.2014.0158](https://doi.org/10.1098/rsif.2014.0158).
- [13] C. Conradi and A. Shiu. Dynamics of posttranslational modification systems: recent progress and future directions. *Biophys. J.*, 114(3):507–515, 2018. doi:[10.1016/j.bpj.2017.11.3787](https://doi.org/10.1016/j.bpj.2017.11.3787); arXiv:[1705.10913](https://arxiv.org/abs/1705.10913).
- [14] F. Crick. Memory and molecular turnover. *Nature*, 312:101, 1984. doi:[10.1038/312101a0](https://doi.org/10.1038/312101a0).
- [15] L. de Moura and S. Ullrich. The Lean 4 theorem prover and programming language. In *Automated Deduction – CADE 28*, volume 12699 of *Lecture Notes in Comput. Sci.*, pages 625–635. Springer, 2021. doi:[10.1007/978-3-030-79876-5_37](https://doi.org/10.1007/978-3-030-79876-5_37).

- [16] K. Deimling. *Nonlinear Functional Analysis*. Springer, Berlin, 1985.
- [17] A. Dickenstein. Biochemical reaction networks: an invitation for algebraic geometers. In *Mathematical Congress of the Americas*, volume 656 of *Contemp. Math.*, pages 65–83. Amer. Math. Soc., 2016. doi:10.1090/conm/656/13076.
- [18] A. Dickenstein. Algebraic geometry tools in systems biology. *Notices Amer. Math. Soc.*, 67(11), 2020. doi:10.1090/noti2188.
- [19] J. Eldering, M. Kvalheim, and S. Revzen. Global linearization and fiber bundle structure of invariant manifolds. *Nonlinearity*, 31(9):4202–4245, 2018. doi:10.1088/1361-6544/aaca8d.
- [20] E. Feliu, N. Kaihnsa, T. de Wolff, and O. Yürük. The kinetic space of multistationarity in dual phosphorylation. *J. Dynam. Differential Equations*, 34(2):825–852, 2022. doi:10.1007/s10884-020-09889-6; arXiv:2001.08285.
- [21] E. Feliu, N. Kaihnsa, T. de Wolff, and O. Yürük. Parameter region for multistationarity in n -site phosphorylation networks. *SIAM J. Appl. Dyn. Syst.*, 22(3):2024–2053, 2023. doi:10.1137/22M1504548; arXiv:2206.08908.
- [22] E. Feliu, A. D. Rendall, and C. Wiuf. A proof of unlimited multistability for phosphorylation cycles. *Nonlinearity*, 33(11):5629–5658, 2020. doi:10.1088/1361-6544/ab9a1e; arXiv:1904.02983.
- [23] E. Feliu and C. Wiuf. Enzyme-sharing as a cause of multi-stationarity in signalling systems. *J. R. Soc. Interface*, 9(71):1224–1232, 2012. doi:10.1098/rsif.2011.0664; arXiv:1101.5799.
- [24] N. Fenichel. Geometric singular perturbation theory for ordinary differential equations. *J. Differential Equations*, 31(1):53–98, 1979. doi:10.1016/0022-0396(79)90152-9.
- [25] D. Flockerzi, K. Holstein, and C. Conradi. N -site phosphorylation systems with $2N - 1$ steady states. *Bull. Math. Biol.*, 76(8):1892–1916, 2014. doi:10.1007/s11538-014-9984-0; arXiv:1312.4774.
- [26] M. I. Freidlin and A. D. Wentzell. *Random Perturbations of Dynamical Systems*, volume 260 of *Grundlehren der mathematischen Wissenschaften*. Springer, Heidelberg, 3rd edition, 2012.
- [27] J. Gunawardena. Multisite protein phosphorylation makes a good threshold but can be a poor switch. *Proc. Natl. Acad. Sci. USA*, 102(41):14617–14622, 2005. doi:10.1073/pnas.0507322102.
- [28] J. Gunawardena. Distributivity and processivity in multisite phosphorylation can be distinguished through steady-state invariants. *Biophys. J.*, 93(11):3828–3834, 2007. doi:10.1529/biophysj.107.110866.
- [29] J. Hell and A. D. Rendall. A proof of bistability for the dual futile cycle. *Nonlinear Anal. Real World Appl.*, 24:175–189, 2015. doi:10.1016/j.nonrwa.2015.02.004; arXiv:1404.0394.
- [30] T. L. Hill. *Free Energy Transduction and Biochemical Cycle Kinetics*. Springer, New York, 1989.
- [31] M. W. Hirsch, C. C. Pugh, and M. Shub. *Invariant Manifolds*, volume 583 of *Lecture Notes in Mathematics*. Springer, Berlin, 1977.

- [32] J. Hofbauer. An index theorem for dissipative semiflows. *Rocky Mountain J. Math.*, 20(4):1017–1031, 1990. doi:[10.1216/rmjm/1181073059](https://doi.org/10.1216/rmjm/1181073059).
- [33] K. Holstein, D. Flockerzi, and C. Conradi. Multistationarity in sequential distributed multisite phosphorylation networks. *Bull. Math. Biol.*, 75(11):2028–2058, 2013. doi:[10.1007/s11538-013-9878-6](https://doi.org/10.1007/s11538-013-9878-6); arXiv:[1304.6661](https://arxiv.org/abs/1304.6661).
- [34] J. G. Kemeny and J. L. Snell. *Finite Markov Chains*. Van Nostrand, Princeton, NJ, 1960.
- [35] J. K. Kim, K. Josić, and M. R. Bennett. The validity of quasi-steady-state approximations in discrete stochastic simulations. *Biophys. J.*, 107(3):783–793, 2014. doi:[10.1016/j.bpj.2014.06.012](https://doi.org/10.1016/j.bpj.2014.06.012).
- [36] P. Kokotović, H. K. Khalil, and J. O’Reilly. *Singular Perturbation Methods in Control: Analysis and Design*, volume 25 of *Classics in Applied Mathematics*. SIAM, Philadelphia, 1999.
- [37] M. A. Krasnosel’skiĭ and P. P. Zabreĭko. *Geometrical Methods of Nonlinear Analysis*, volume 263 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, 1984.
- [38] D. Li, S. Liao, Q. Ouyang, and F. Li. Interplay between ATP and hydrolysis free energy in promoting and restraining biochemical switches. *Phys. Rev. Research*, 6:033050, 2024. doi:[10.1103/PhysRevResearch.6.033050](https://doi.org/10.1103/PhysRevResearch.6.033050).
- [39] J. E. Lisman. A mechanism for memory storage insensitive to molecular turnover: a bistable autophosphorylating kinase. *Proc. Natl. Acad. Sci. USA*, 82(9):3055–3057, 1985. doi:[10.1073/pnas.82.9.3055](https://doi.org/10.1073/pnas.82.9.3055).
- [40] A. K. Manrai and J. Gunawardena. The geometry of multisite phosphorylation. *Biophys. J.*, 95(12):5533–5543, 2008. doi:[10.1529/biophysj.108.140632](https://doi.org/10.1529/biophysj.108.140632).
- [41] N. I. Markevich, J. B. Hoek, and B. N. Kholodenko. Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades. *J. Cell Biol.*, 164(3):353–359, 2004. doi:[10.1083/jcb.200308060](https://doi.org/10.1083/jcb.200308060).
- [42] P. Miller, A. M. Zhabotinsky, J. E. Lisman, and X.-J. Wang. The stability of a stochastic CaMKII switch: dependence on the number of enzyme molecules and protein turnover. *PLoS Biol.*, 3(4):e107, 2005. doi:[10.1371/journal.pbio.0030107](https://doi.org/10.1371/journal.pbio.0030107).
- [43] J. W. Milnor. *Topology from the Differentiable Viewpoint*. University Press of Virginia, Charlottesville, 1965.
- [44] I. Mirzaev and J. Gunawardena. Laplacian dynamics on general graphs. *Bull. Math. Biol.*, 75(11):2118–2149, 2013. doi:[10.1007/s11538-013-9884-8](https://doi.org/10.1007/s11538-013-9884-8).
- [45] H. Qian. Phosphorylation energy hypothesis: open chemical systems and their biological functions. *Annu. Rev. Phys. Chem.*, 58:113–142, 2007. doi:[10.1146/annurev.physchem.58.032806.104550](https://doi.org/10.1146/annurev.physchem.58.032806.104550).
- [46] R. Rao and M. Esposito. Nonequilibrium thermodynamics of chemical reaction networks: wisdom from stochastic thermodynamics. *Phys. Rev. X*, 6(4):041064, 2016. doi:[10.1103/PhysRevX.6.041064](https://doi.org/10.1103/PhysRevX.6.041064).
- [47] C. Salazar and T. Höfer. Multisite protein phosphorylation – from molecular mechanisms to kinetic models. *FEBS J.*, 276(12):3177–3198, 2009. doi:[10.1111/j.1742-4658.2009.07027.x](https://doi.org/10.1111/j.1742-4658.2009.07027.x).

- [48] S. Schuster and R. Schuster. A generalization of Wegscheider’s condition. Implications for properties of steady states and for quasi-steady-state approximation. *J. Math. Chem.*, 3:25–42, 1989. doi:[10.1007/BF01171883](https://doi.org/10.1007/BF01171883).
- [49] A. Shwartz and A. Weiss. *Large Deviations for Performance Analysis: Queues, Communications, and Computing*. Chapman & Hall, London, 1995.
- [50] T. Suwanmajo and J. Krishnan. Mixed mechanisms of multi-site phosphorylation. *J. R. Soc. Interface*, 12(107):20141405, 2015. doi:[10.1098/rsif.2014.1405](https://doi.org/10.1098/rsif.2014.1405).
- [51] C. A. Taylor IV, K. W. Cormier, S. E. Keenan, S. Earnest, S. Stippec, C. Wichaidit, Y.-C. Juang, J. Wang, S. Y. Shvartsman, E. J. Goldsmith, and M. H. Cobb. Functional divergence caused by mutations in an energetic hotspot in ERK2. *Proc. Natl. Acad. Sci. USA*, 116(31):15514–15523, 2019. doi:[10.1073/pnas.1905015116](https://doi.org/10.1073/pnas.1905015116).
- [52] The mathlib Community. The Lean mathematical library. In *Proc. 9th ACM SIGPLAN Int. Conf. on Certified Programs and Proofs (CPP 2020)*, pages 367–381. ACM, 2020. doi:[10.1145/3372885.3373824](https://doi.org/10.1145/3372885.3373824); arXiv:[1910.09336](https://arxiv.org/abs/1910.09336).
- [53] M. Thomson and J. Gunawardena. The rational parameterisation theorem for multisite post-translational modification systems. *J. Theoret. Biol.*, 261(4):626–636, 2009. doi:[10.1016/j.jtbi.2009.09.003](https://doi.org/10.1016/j.jtbi.2009.09.003).
- [54] M. Thomson and J. Gunawardena. Unlimited multistability in multisite phosphorylation systems. *Nature*, 460:274–277, 2009. doi:[10.1038/nature08102](https://doi.org/10.1038/nature08102).
- [55] A. Torres and E. Feliu. Symbolic proof of bistability in reaction networks. *SIAM J. Appl. Dyn. Syst.*, 20(1):1–37, 2021. doi:[10.1137/20M1326672](https://doi.org/10.1137/20M1326672); arXiv:[1909.13608](https://arxiv.org/abs/1909.13608).
- [56] L. Wang and E. D. Sontag. On the number of steady states in a multiple futile cycle. *J. Math. Biol.*, 57(1):29–52, 2008. doi:[10.1007/s00285-007-0145-z](https://doi.org/10.1007/s00285-007-0145-z); arXiv:[0704.0036](https://arxiv.org/abs/0704.0036).