Complete positivity and equilibrium realizations of reversible observations

Samuil Petkov
Department of Physics, École normale supérieure, Université PSL, Paris, France
samuil.petkov@phys.ens.psl.eu

5 September 2026

Abstract

Complete positivity of mirrored observation matrices is necessary for equilibrium realization, but does not ensure compatible equilibrium dynamics. We study stationary finite-state continuous-time Markov chains with instantaneous deterministic labels on a fixed physical clock. For reflection-positive laws with d active labels and reflected rank $r \le d+1$, we construct a detailed-balance realization on the minimal number r of states. Binary rank-four laws with two prediction dimensions per label compress to four states when a finite detailed-balance realization already exists. At $r = d+2$, a rational ten-state chain with six labels passes every mirrored matrix-CP test through one common nonnegative Gram factorization, yet admits no equilibrium realization. A rational seven-time statistic separates its law from all finite equilibrium observation laws, uniformly over hidden dimensions, rates and topology. It gives a positive entropy-production floor for every exact finite reciprocal realization, with conservative constants too small for practical statistical use. The law also has no ordinary symmetric Markov extension on any standard Borel state space at the original clock, despite factorability of its reflected half-path operator. For laws that fail a matrix-CP test, a separate deletion-deficit theorem gives a computable nonlinear dissipation bound.

Mathematics Subject Classification (2020). Primary 60J27; Secondary 82C05, 15B48.
Keywords. equilibrium realization; detailed balance; hidden Markov models; reflection positivity; complete positivity; reflected rank; entropy production.

hexagon:2611.00658v1

1 What reversible observations can establish about equilibrium

An observed trajectory can be invariant under time reversal while its microscopic Markov mechanism violates detailed balance, as established for aggregated and hidden Markov models [30, 31]. The nonuniqueness of aggregated Markov presentations of one observed law is classical [43, 44]. The realization question asks whether some finite detailed-balance mechanism produces exactly the observed law. An irreversible presentation alone cannot answer this question: the same law may also admit an equilibrium presentation. We study exact realization with instantaneous deterministic labels on a fixed continuous clock.

At equilibrium, two observed histories read outward from a stationary midpoint are conditionally independent and identically distributed given the hidden midpoint. Every common finite classification of those histories therefore gives a completely positive matrix, a Gram sum of nonnegative vectors. This classical conditional-iid geometry [17, 18] is stronger than entrywise nonnegative quadratic positivity [7–9]. It also connects the realization question to reflection-positive reconstruction. Reflection positivity supplies a self-adjoint observable semigroup [26, 28], but does not generally supply a larger symmetric Markov process: Arveson's discrete-time nonextendable examples already establish that broader distinction [27].

A related but distinct finite-memory obstruction comes from Fanizza et al. [38]: some discrete-time processes with finite linear memory have no finite-dimensional quantum realization. There, complete positivity refers to quantum evolution maps; here it refers to entrywise nonnegative Gram factorizations of mirrored classical observation matrices. Neither representation contract settles whether the same continuous-time observed law has a finite ordinary detailed-balance realization with instantaneous labels and the original clock.

The first result is an exact low-rank guarantee. Let d be the number of active labels and r the rank of the full reflected event kernel. Reflection positivity and $r \le d+1$ suffice for an irreducible detailed-balance realization on exactly r states, preserving all labels and every cylinder on the original clock. The

construction draws on established low-dimensional positive-realization and binary aggregated-model ideas [24, 25, 29, 32, 33] and enforces the label and generator constraints simultaneously.

A complementary compression theorem treats binary rank four when each label has two prediction dimensions: any existing finite detailed-balance realization can be replaced by a four-state one, irrespective of its original hidden-state count. The two replacements must be selected jointly to preserve nonnegative cross rates.

The general low-rank guarantee nevertheless ends at $r = d + 1$. A rational ten-state model with six labels and $r = d + 2$ passes every mirrored matrix-CP test using one common nonnegative factorization, but has no finite detailed-balance realization of its complete observed law. A seven-time rational statistic separates it from all finite equilibrium observations and certifies a positive dissipation floor across exact compatible models. The obstruction concerns symmetric nonnegative dynamics, rather than incompatible choices of Gram factors for separate windows.

For thermodynamic inference, a violation of one CP test has a quantitative consequence. Observed path-relative entropy vanishes for a reversible law [1, 2], while counts, waiting times and constraints on microscopic statistics can retain information [3–5, 14]. Equilibrium restrictions on binary dwell laws already provide nonequilibrium tests [23]; Skinner and Dunkel also give reversible observed trajectories with positive certified hidden dissipation [3]. Thus reversible observations concealing dissipation are a precedent, not the novelty claim here. Hidden entropy production under coarse graining and for a marginal observer has been treated in stochastic thermodynamics [39, 40], and fits of hidden Markov models to observed data give lower bounds on it [41]. Transition-indexed waiting-time methods [6, 16] and recently resolved excursion or cycle methods [34, 35] use complementary observation contracts. Here the completely positive deletion deficit of an arbitrary matched classifier gives a nonlinear entropy-production bound independent of the finite hidden dimension and topology. Its convex weight and copositive dual have classical geometric antecedents [19]; the contribution is their link to microscopic Markov path divergence. Finite-observable certificates for dissipation plus a positive price per hidden state require global model calibration [46]. Our explicit separation concerns the unpenalized union of finite equilibrium architectures.

Section 2 fixes the observation contract. Section 3 states the low-rank realization and binary compression theorems, followed by the all-window-CP counterexample. Section 4 gives the finite-window dissipation theorem, followed by one reversible benchmark and fixed-rule sampling guarantees. All full proofs, the rational counterexample and the additional robustness and kinetic comparisons are included after the references. Exact finite equilibrium nonexistence and a positive dissipation infimum remain distinct throughout. *Scope.* Classical ingredients are credited to the cited sources, and all results are stated for the observation contract of Section 2 only.

2 Observation model and physical conventions

Let X be a finite, irreducible, stationary continuous-time Markov chain with row generator Q, stationary law π and reciprocal support: $q_{xy} > 0$ if and only if $q_{yx} > 0$. Missing reciprocal pairs are allowed. Every microscopic state is even under physical time reversal. Write

$$\sigma(Q) = \sum_{x<y} (\pi_x q_{xy} - \pi_y q_{yx}) \log \frac{\pi_x q_{xy}}{\pi_y q_{yx}}. \tag{1}$$

An absent pair contributes zero. We set $k_{\mathrm{B}} = 1$ and use natural logarithms. Identification with physical dissipation requires thermodynamically resolved transitions and the appropriate local detailed-balance interpretation [2].

Observe $Y_t = h(X_t)$ for an arbitrary instantaneous label map. We call a realization equilibrium when its microscopic generator satisfies ordinary detailed balance. Exact realization means equality of the complete stationary cadlag observed law at every physical time. No state-count or topology constraint is imposed on a competing finite realization.

3 A sharp rank boundary for finite equilibrium realization

For this section the observed stationary path law P is reversible. Let d be its number of active labels. For bounded future cylinder functions, define the reflected form

$$[f, g]_R = \mathbb{E}_P[(\theta f) g], \tag{2}$$

where θf reads the same functional outward into the past. Reflection positivity means $[f, f]_R \geq 0$ for every real cylinder function f. Equivalently, every mirrored matrix formed from a common finite cylinder partition is positive semidefinite. Quotienting by the null space gives a real Hilbert space $\mathcal{H}$ of dimension r, the *reflected rank*. If the supplied CTMC has n states, midpoint conditioning gives $r \leq n$. This rank concerns the entire observed law; it is neither the rank of one motif matrix nor its CP-rank.

Theorem 3.1 (Minimal equilibrium realization at low reflected rank)**.** *Suppose P is the reversible stationary law of instantaneous deterministic labels of a finite irreducible CTMC and is reflection positive, with d active labels and reflected rank $r = \dim \mathcal{H}$. If $r \leq d + 1$, then P has an exact irreducible detailed-balance CTMC realization with r states, the same labels and the same physical clock. No finite CTMC realization of P can have fewer than r states.*

The reflected representation consists of a self-adjoint generator $A \leq 0$, a unit stationary vector v, and orthogonal label projections E_a whose label spaces $\mathcal{H}_a = E_a \mathcal{H}$ decompose $\mathcal{H}$. At rank d, every label space is one-dimensional and positivity of short-time observed transitions makes A a symmetric Metzler matrix. At rank $d + 1$, only one label space can be two-dimensional. The proof rotates that space using the dual of its acute future-event cone. Invariance under its killed semigroup controls the internal rate, while the same dual cone controls every cross-label rate. The resulting symmetric Metzler matrix B has a strictly positive unit null vector w. Thus $Q = \operatorname{diag}(w)^{-1} B \operatorname{diag}(w)$ is the required row generator, with stationary masses w_i^2. Appendix A gives the reconstruction, endpoint rotation and full-cylinder verification.

Low-dimensional positive realization has substantial predecessors. A conditional-rank theorem of Dharmadhikari, whose low-rank corollary credits an announcement by Fox, supplies finite discrete-time Markov realizations when all conditional ranks are at most two [24, Corollary 2.1]. For every sampling interval $\Delta > 0$, his conditional word rank $n_\Delta(a)$ equals $\dim \mathcal{H}_a$ here, as proved in Appendix A.2. Thus the ordinary r-state discrete-time realization count on this slice is already supplied by that corollary. Invariant polyhedral cones are classical realization data [25, 45]. Symmetric scalar positive realization is treated by Grussler and Damm [29]. For binary three-state aggregated models, an explicit rotation to a detailed-balance equivalent appears in the author preprint of Wu and Jia [32], Appendix A; related binary equivalence classes are analyzed by Siekmann [33]. Theorem 3.1 concerns simultaneous preservation of every label projection and every continuous-time cylinder. It requires no renewal assumption. For example, a four-state equilibrium chain with labels (a, a, b, c) can have rank four and entrance-dependent visits to a; an exact example is given at the end of Appendix A.

The binary case with two two-dimensional label spaces admits a further compression when an equilibrium presentation is already available. The space V below is the span of conditional future-cylinder functions; at equilibrium it is isometric to the reflected space $\mathcal{H}$.

Theorem 3.2 (Minimal binary equilibrium compression)**.** *Let Q be a finite irreducible detailed-balance CTMC with two nonempty instantaneous deterministic labels a, b, and let V be the smallest real space of state functions containing $\mathbf{1}$ and invariant under Q and the two label masks $\mathbf{1}_a, \mathbf{1}_b$. If*

$$\dim V = 4, \qquad \dim(\mathbf{1}_a V) = \dim(\mathbf{1}_b V) = 2, \tag{3}$$

then its stationary observed law has an exact irreducible four-state detailed-balance realization, with two states per label and the same physical clock. Four states are minimal among all finite CTMC realizations of that law. The original number of states in either label is unrestricted.

Each label's prediction coordinate can be normalized to have conditional mean zero and variance one. Detailed balance then expresses the cross-label action by a bilinear rate kernel. Conditional destination means make this kernel nonnegative on two crossing strips. Since its value at the origin is positive, one outer corner is positive. Retaining that corner and replacing the opposite coordinates by their negative reciprocals supplies two two-point measures with matching moments and nonnegative cross rates. The retained extrema also determine the signs of both internal rates. Appendix B verifies the generator, full cylinder law, irreducibility and minimality. Dharmadhikari's conditional-rank corollary cited above supplies the ordinary four-state realization count for each sampled binary law; Theorem 3.2 preserves detailed balance and the entire continuous clock in one realization. Its assumption of an existing equilibrium presentation is essential to the stated compression result.

Every equilibrium realization passes all mirrored matrix-CP tests. Theorem 3.1 shows that reflection positivity alone suffices at $r \leq d + 1$. The next result places a limitation exactly one rank increment beyond that general sufficient regime.

Theorem 3.3 (All-window complete positivity without a finite equilibrium lift)**.** *There exists a ten-state irreducible stationary CTMC with rational rates, reciprocal support and six instantaneous deterministic labels whose full observed law is reversible and has reflected rank eight. For every positive window and every finite measurable matched classifier, its mirrored matrix is completely positive, with a common factorization through the ten microscopic coordinates. Nevertheless, no finite stationary detailed-balance CTMC with instantaneous deterministic labels has that same full observed law on the same clock, at any hidden-state count.*

The construction uses five core states sharing one label and five distinct probe labels. Core states carry unit-circle prediction points p_i. Nonnegative affine probe rates cut out their polygon. In any hypothetical equilibrium realization, the same observed law forces every competing core state into that polygon. Its second moment saturates the unit-radius bound, so all these states lie at the original vertices, with the original aggregate masses. Detailed balance would then impose symmetric nonnegative conductances between the vertices. A rational separating identity rules them out. Appendix C supplies the full generator, the common all-window factorization and the argument over every competing state count. The example has one three-dimensional label space and five one-dimensional spaces: $r = d + 2$.

Discrete-time reflection-positive processes without symmetric Markov extensions were already constructed by Arveson [27]. His extension contract allows arbitrary standard Borel state spaces, and his obstruction uses nonfactorability of the reflected half-path operator. Theorem 3.3 is a finite continuous-time construction satisfying the stronger all-window matrix-CP condition. The connection can be made exact: for the observed half-path L^2 space, all these partition matrices are CP if and only if its reflected operator K is factorable as B^*B with B order-preserving. Appendix D proves this comparison using Arveson's closed-cone theorem. For the construction above, B is simply conditional future expectation given the original ten-state midpoint. The continuous-time symmetric-extension setting was subsequently discussed by Mitter [22, pp. 104–105]; the present observed finite-CTMC law is continuous in probability. We retain the explicit standard Borel, ordinary-symmetry and deterministic-observation contract below. The historical formulation alone is not a present-day priority determination. Ordinary symmetry leaves microscopic state values unchanged under reversal; no additional odd-variable involution is allowed.

Corollary 3.4 (No ordinary symmetric same-clock extension)**.** *The observed law in Theorem 3.3 has no realization as an instantaneous deterministic Borel observation of a stationary two-sided ordinary symmetric Markov process on a standard Borel state space, at the same continuous clock.*

The proof extends the polygon rigidity to measurable prediction points. Saturation leaves only ten observable prediction classes. A finite-valued conductance lemma then reproduces the finite observable restriction by a finite detailed-balance generator, contradicting the same rational certificate. The observed law supplies an analytic semigroup on its finite-dimensional observable image; no global microscopic continuity, jump kernel, finite rate, or generator-domain condition on the class indicators is assumed. This conclusion is a continuous-clock extension statement; it does not settle the extension problem for an isolated sampled discrete-time process.

Exact equilibrium nonexistence alone does not establish a positive dissipation infimum or resistance to approximation. For this law, a finite-observation certificate gives both conclusions explicitly. Let $\mathcal{E}_{\mathrm{fin}}$ contain all stationary finite detailed-balance CTMC observation laws on these six labels, with unrestricted hidden dimensions, rates and topology; reducible models are allowed. Enumerate finite strictly increasing rational time tuples containing zero as $(\mathcal{T}_m)_{m\ge1}$ and set

$$D(P,P') = \sum_{m\ge1} 2^{-m}\mathrm{TV}(P_{\mathcal{T}_m}, P'_{\mathcal{T}_m}). \tag{4}$$

Theorem 3.5 (An explicit certificate beyond matrix-CP tests)**.** *For the law P in Theorem 3.3, put $h = 10^{-3}$, $t = 10^{-7}$ and*

$$\mathcal{T} = (0, t, h, h+t, h+2t, 2h+t, 2h+2t), \qquad T = 2(h+t). \tag{5}$$

The rational word statistic $W\colon \{a, b_1, \dots, b_5\}^7 \to [0,1]$ specified in (71)–(88), together with the constants $c = 1/2 + 1/(192B)$ and $\eta = 71/(4800B) > 0$ defined there, satisfies

$$\mathbb{E}_P W \ge c + \eta, \qquad \sup_{P'\in\mathcal{E}_{\mathrm{fin}}} \mathbb{E}_{P'} W \le c. \tag{6}$$

Consequently, if m_0 indexes $\mathcal{T}$ in (4),

$$\inf_{P'\in\mathcal{E}_{\mathrm{fin}}} D(P,P') \ge 2^{-m_0}\eta > 0, \qquad I_\infty(P) := \inf_{Q \text{ exactly compatible}} \sigma(Q) \ge \frac{8\eta^2}{T} > 0. \tag{7}$$

The second infimum ranges over all finite irreducible stationary reciprocal CTMCs with instantaneous deterministic labels producing exactly P on the same clock.

The score tests whether predictions near the polygon vertices can evolve consistently with a prescribed rational finite-lag action. Every mismatch of an unknown competitor with this action enters an observed residual penalty; no model fit or rate bound is assumed. A first-moment polygon estimate controls rounding errors using a weakly monotone dual field, including its zero-coercivity pairs. Appendix E gives the construction and the reversibilization argument for the entropy-production bound. The margin is approximately 1.2398×10^{-26}, so the certificate establishes effective separation without practical statistical power. Thus all matrix-CP deficits can vanish while a finite observed statistic certifies a positive hidden-dissipation floor. The theorem below instead applies when a tested matrix itself fails CP.

4 A population bound for arbitrary finite classifiers

Fix a half-window $\tau > 0$ and one measurable finite classifier Φ. Write $Y_0^{\leftarrow} = Y_0$ and $Y_s^{\leftarrow} = h(X_{(-s)-})$ for $s > 0$, using the usual right-continuous reversal. Deterministic endpoints are almost surely not jump times. Define

$$A = \Phi((Y_s^{\leftarrow})_{0\le s\le\tau}), \qquad B = \Phi((Y_s)_{0\le s\le\tau}), \qquad M_{ij} = \mathbb{P}(A = i, B = j), \quad S = \frac{M + M^{\mathsf{T}}}{2}. \tag{8}$$

The past is read *outward* from the midpoint, using the same rule as the future. The full record has duration 2τ. A common stochastic classifier is also allowed when independent auxiliary randomness is used on the two halves. Arbitrary correlated detector memory is not covered by this observation contract. This distinction matters for entropy inference from last-visited records [36, 37, 42].

For m categories, let $\mathcal{CP}_m = \{\sum_r u_r u_r^{\mathsf{T}} : u_r \ge 0\}$ and $\mathcal{COP}_m = \{W = W^{\mathsf{T}} : u^{\mathsf{T}} W u \ge 0 \ \forall u \ge 0\}$. Pairings sum over *ordered* indices. Set

$$\Delta_{\mathrm{cp}}(S) = 1 - \max\{\mathbf{1}^{\mathsf{T}} G \mathbf{1} : G \in \mathcal{CP}_m,\ 0 \le G \le S\}, \tag{9}$$

$$\mathcal{F}(v) = \sqrt{v}\ \operatorname{artanh}\sqrt{v} = \sum_{k=1}^{\infty} \frac{v^k}{2k-1}, \quad 0 \le v < 1. \tag{10}$$

All inequalities in (9) are entrywise. The retained mass $1 - \Delta_{\mathrm{cp}}(S)$ is the largest completely positive fraction of S; equivalently, Δ_{cp} is the convex resource weight with normalized completely positive matrices as the free set [19]. The dual representation below is this classical weight geometry specialized to categorical pair distributions. General copositive optimization need not be computationally tractable.

Theorem 4.1 (Finite-window dissipation witness)**.** *Under the finite reciprocal CTMC observation contract of Section 2, without requiring observed reversibility, let $\sigma(Q)$ be as in* (1)*, M, S as in* (8)*, and $\Delta_{\mathrm{cp}}, \mathcal{F}$ as in* (9)–(10)*. Then*

$$\boxed{\sigma(Q) \ \ge \ \frac{2}{\tau}\mathcal{F}(\Delta_{\mathrm{cp}}(S)).} \tag{11}$$

In particular, for any symmetric copositive W with $W_{ij} \ge -1$, put $v_W = [-\langle W, M\rangle]_+$. Then $0 \le v_W \le \Delta_{\mathrm{cp}}(S) < 1$ and

$$\sigma(Q) \ge \frac{2}{\tau}\mathcal{F}(v_W) \ge \frac{2v_W}{\tau}. \tag{12}$$

At equilibrium M is completely positive. Conversely, $\Delta_{\mathrm{cp}}(S) = 0$ holds exactly when S is completely positive; this is a statement about one tested matrix, not a characterization of the entire observed process.

Proof of the thermodynamic bound. Condition on $X_0 = x$ and write $a_{xi} = \mathbb{P}(A = i \mid X_0 = x)$ and $b_{xi} = \mathbb{P}(B = i \mid X_0 = x)$. Conditional independence gives $M = \sum_x \pi_x a_x b_x^{\mathsf{T}}$. Define

$$z_{xi} = \sqrt{a_{xi} b_{xi}}, \qquad G = \sum_x \pi_x z_x z_x^{\mathsf{T}}, \qquad H_x = \sum_i \sqrt{a_{xi} b_{xi}}. \tag{13}$$

The arithmetic–geometric mean inequality gives $0 \le G \le S$, and G is completely positive. Its diagonal equals that of S. Thus

$$\Delta_{\mathrm{cp}}(S) \le \delta := 1 - \mathbf{1}^{\mathsf{T}} G \mathbf{1} = \sum_x \pi_x (1 - H_x^2). \tag{14}$$

The elementary relative-entropy/affinity inequality, proved in Appendix F, is

$$J(p,q) := D_{\mathrm{KL}}(p\|q) + D_{\mathrm{KL}}(q\|p) \geq 4\mathcal{F}(1 - H(p,q)^2). \tag{15}$$

Convexity of $\mathcal{F}$ then gives $\sum_x \pi_x J(a_x, b_x) \geq 4\mathcal{F}(\delta)$. The stationary reverse generator is $q^\dagger_{xy} = \pi_y q_{yx}/\pi_x$. Past half-paths conditioned on x have its forward law. The two averaged conditional path divergences, before classification, are each $\tau\sigma(Q)$. Applying data processing in both directions yields

$$4\mathcal{F}(\Delta_{\mathrm{cp}}(S)) \leq 4\mathcal{F}(\delta) \leq \sum_x \pi_x J(a_x, b_x) \leq 2\tau\sigma(Q). \tag{16}$$

The path identities, support issues and coefficient $2/\tau$ are justified in Appendix G. For a copositive $W \geq -1$, $-\langle W, S\rangle \leq \mathbf{1}^{\mathsf{T}}(S - G')\mathbf{1}$ for every feasible G' in (9). This proves (12). At equilibrium $a_x = b_x$, yielding the claimed factorization of M. The remaining cone assertions follow from Proposition 4.2. □

Proposition 4.2 (Optimized witnesses and coarsening)**.** *For any symmetric nonnegative S of total mass one,*

$$\Delta_{\mathrm{cp}}(S) = \sup\{-\langle W, S\rangle : W \in \mathcal{COP}_m,\ W_{ij} \geq -1\}. \tag{17}$$

The primal maximum exists and has $\operatorname{diag} G = \operatorname{diag} S$*. If K is any nonnegative row-stochastic matrix, then*

$$\Delta_{\mathrm{cp}}(K^{\mathsf{T}} S K) \leq \Delta_{\mathrm{cp}}(S). \tag{18}$$

Consequently refining the same classifier on both halves cannot decrease the optimized deficit at fixed τ; independent matched relabeling cannot increase it.

The proof in Appendix H uses only finite-dimensional separation and nonnegative Gram factors. A zero deficit is not a certificate of an equilibrium Markov realization. In particular, compatible factorizations at separate window lengths need not form one Markov semigroup.

For overlapping relabeling channels the contraction can be quantified:

$$\Delta_{\mathrm{cp}}(K^{\mathsf{T}} S K) \leq c(K)\Delta_{\mathrm{cp}}(S), \qquad c(K) = 1 - \min_{i,j} \left(\sum_a \sqrt{K_{ia} K_{ja}} \right)^2. \tag{19}$$

The coefficient is sharp over all symmetric nonnegative probability matrices, as proved in Appendix H. Sharpness over physically compatible CTMC motif matrices is a separate question. Leaving a censoring class isolated can give $c(K) = 1$ even when active classes overlap.

5 Five motif classes and an explicit reversible benchmark

Arrange five active motif classes on the cycle C_5, and reserve class 0 for censoring or nonselection. With A_5 its adjacency matrix, let $W = J_5 - 2A_5$ on the active classes and put zero in its censoring row and column. This is the classical Horn copositive witness. A short proof is given in Appendix I. Define

$$Z = 2\mathbf{1}_{\{A \sim B\}} - \mathbf{1}_{\{A \neq 0,\, B \neq 0\}}, \qquad v = [\mathbb{E}Z]_+. \tag{20}$$

Here adjacency excludes equal classes. The observable score takes values in $\{-1, 0, 1\}$, and

$$\boxed{\sigma \geq \frac{2}{\tau}\mathcal{F}(v) \geq \frac{2v}{\tau}.} \tag{21}$$

At equilibrium, conditional on both classes being active, their adjacency probability cannot exceed $1/2$. This implication is used only when the conditioning event has positive probability; censored records are retained in the unconditional score.

5.1 Calibration against an attained full-law optimum

The inherited construction [10] has five visible states, one common label H for twenty hidden states, and

$$R = I + \frac{3}{5}A_5 + \frac{1}{100}(J_5 - I - A_5), \qquad f(t) = \lambda e^{-\lambda t} R, \qquad \lambda = \frac{111}{5}. \tag{22}$$

Visible stationary masses are $1/10$, total hidden mass is $1/2$, and the symmetric direct visible flows are $F_{ij} = 1/100$ for $i \neq j$. Here R_{ij} is the stationary occurrence rate of an excursion from visible state i through H to visible state j, including returns. The full observed stationary process is reversible. Its exact minimum entropy production among finite reciprocal realizations is

$$I_\infty^{\mathrm{full}} = \frac{3\sqrt{111}}{5} \operatorname{arcosh} \frac{20}{17} \simeq 3.70232333. \tag{23}$$

Appendix J supplies the microscopic channels, stationarity, the inherited sharp lower bound and attainment. This value is used for calibration; it is *not* claimed to be determined by a single motif matrix.

For a half-record starting in H, let Φ return the first visible state reached by time τ. Return 0 if no visible state is reached, or if the midpoint starts in a visible state. For $q = e^{-\lambda\tau}$ and $t_\tau = (1-q)^2/\lambda$, direct stationary calculation gives

$$M = \begin{pmatrix} (1+q^2)/2 & q(1-q)\mathbf{1}^{\mathsf{T}}/10 \\ q(1-q)\mathbf{1}/10 & t_\tau R \end{pmatrix}, \qquad v = \frac{9}{10} t_\tau. \tag{24}$$

The whole six-category matrix is positive definite, not just its active principal block. Its optimized deficit equals this same score: $\Delta_{\mathrm{cp}}(M) = v$, by an explicit primal Gram certificate in Appendix J.

The linear bound is largest at $x = \lambda\tau$ solving $e^x = 1 + 2x$, with $x \simeq 1.25643121$. It gives 0.73307588 in the chosen inverse-time unit. At the same window the nonlinear bound is 0.73820904, approximately 19.9% of (23). The nonlinear optimum is not asserted to occur at the linear optimizer. These decimals are numerical evaluations of proved formulas.

5.2 A finite noise tolerance

Independently on each half, keep an active class with probability $1 - \eta$ and replace it by a uniform class in $\{1, \ldots, 5\}$ with probability η. Class 0 is unchanged. This is the same stochastic map in both directions, so it preserves the validity of the thermodynamic test. The new score is exactly

$$\mathbb{E}Z_\eta = t_\tau \left[\frac{78}{25}(1-\eta)^2 - \frac{111}{50} \right]. \tag{25}$$

It is positive precisely for

$$\eta < \eta_* := 1 - \sqrt{\frac{37}{52}} \simeq 0.15647261. \tag{26}$$

In fact, the whole noisy matrix has $\Delta_{\mathrm{cp}}(M_\eta) = [\mathbb{E}Z_\eta]_+$: it enters the completely positive cone exactly at η_*. Appendix J gives a nonnegative Gram certificate on both sides of this threshold. Complete positivity at one window does not imply equilibrium realizability of the full process. This is a population robustness threshold for *uniform motif relabeling*, not a universal detector-error tolerance. It does not cover asymmetric filters or arbitrary errors in transition times. Both reversibility and positive semidefiniteness of the matrix are preserved by this matched corruption.

The benchmark can also be extended to a reciprocal 50-state copy-switch model in which every observed label is many-to-one. Its full observed law remains reversible and its tested matrix remains positive definite with a strictly positive CP witness for every finite positive switching vector. Appendix K gives the construction and exact score. This is a robustness extension of the non-CP benchmark, separate from the all-window-CP counterexample of Theorem 3.3.

A four-stage driven cycle supplies a second check of the finite-window certificate. With two compound labels, it has a reversible observed law and a quadratic weak-driving signal, whereas the specified persistence mutual-information certificate of Anonymous [15] is quartic on the same data. This binary example tests estimator strength, rather than CP versus PSD. The exact formulas, heterogeneous-rate radius and comparison are in Appendix J.6.

6 Finite-data statements and their assumptions

Population inequalities and empirical confidence bounds are different assertions. Fix Φ, τ and the cycle assignment in advance, or choose them using training records independent of the test records. For n independent stationary snippets, let $\overline{Z}_n$ be the average score. Hoeffding's inequality [11] gives, with probability at least $1 - \alpha$,

$$\sigma \geq \frac{2}{\tau} \mathcal{F}\left([\overline{Z}_n - \sqrt{2\log(1/\alpha)/n}]_+ \right). \tag{27}$$

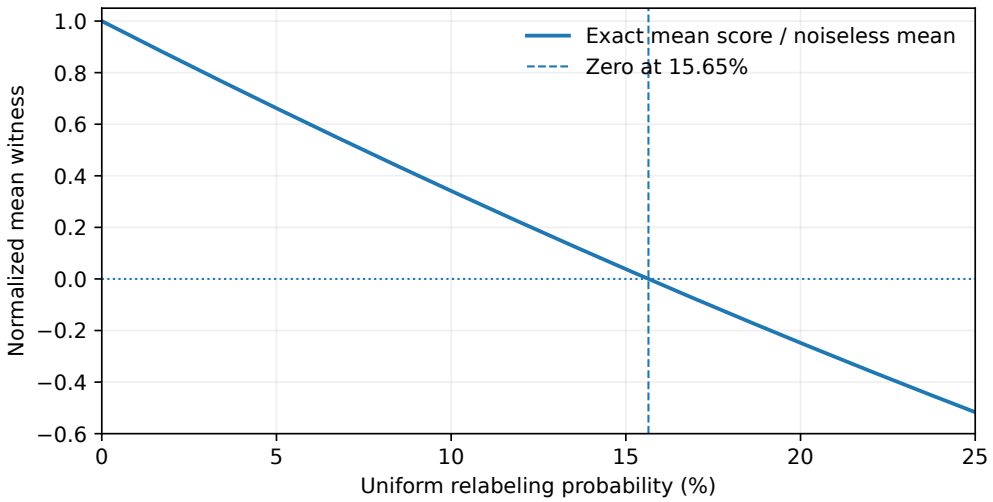


Figure 1: Normalized mean witness under independent uniform relabeling of the active motif categories, at any fixed positive half-window. The exact zero is $1 - \sqrt{37/52}$. Relabeling never creates a violation for an equilibrium process; here it gradually removes the available signal.

The three-valued score also permits a sharper construction using classical exact binomial confidence limits [12]. Let N be the number of nonzero scores and K the number of positive scores. Write $L(k, n; \gamma)$ for the one-sided exact binomial lower endpoint, defined in Appendix L. With probability at least $1 - \alpha$,

$$\sigma \geq \frac{2}{\tau} \mathcal{F}(L(N, n; \alpha/2)\,[2L(K, N; \alpha/2) - 1]_+)\,. \tag{28}$$

An exact equilibrium test follows by conditioning on N: under equilibrium K has law $\mathrm{Bin}(N, p)$ with $p \leq 1/2$. Its upper tail at $p = 1/2$ is a valid conservative p-value. This is a test of the specified equilibrium model class, not a diagnosis of which microscopic assumption fails.

For a single stationary trajectory, the same event-transfer proof applies to nonoverlapping snippets separated by gaps only when an external hidden-state mixing bound is justified. With the bound $b(g)$ in (127), the independent-snippet failure probability increases by at most $(n-1)b(g)$. Appendix L gives this assumption, training and candidate-selection corrections, exact binomial power and contamination bounds. Empirical autocorrelation is not substituted for a hidden-state mixing guarantee.

7 Interpretation and remaining scope

Mirrored observed histories distinguish matrix positivity, equilibrium realization and dissipation. A non-CP tested matrix certifies a uniform positive entropy-production floor for compatible finite reciprocal mechanisms. At reflected rank $r \leq d + 1$, reflection positivity gives a minimal exact finite equilibrium realization. Beyond that regime, even a common CP hierarchy need not admit dynamically compatible equilibrium rates. The rational example makes the rank-gap guarantee sharp in general, without classifying every alphabet/rank pair or label-block geometry.

The realization and thermodynamic conclusions have different quantifiers. Theorem 3.3 excludes every exact finite ordinary detailed-balance CTMC; Corollary 3.4 also excludes arbitrary-state stationary ordinary symmetric Markov extensions. The finite-word separation in Theorem 3.5, rather than exact nonexistence alone, supplies an explicit positive entropy-production floor over all exact compatible finite reciprocal models. Its very small margin does not provide a practical sample-size guarantee. Theorem 4.1 gives a different positive floor when a finite observed matrix fails CP; neither bound is asserted to attain the full-law dissipation optimum. All results retain instantaneous labels, the original clock and even microscopic time-reversal parity. Correlated detector memory and generalized odd-variable reversal require different contracts.

Reproducibility. The appendices give complete proofs and exact rational specifications of the generators, observable spaces, facet and conductance certificates, finite-window formulas and confidence

calculations. Numerical evaluations illustrate analytical results and are not proofs of the all-time or arbitrary-state-count claims. The ancillary file `anc/check_certificate.py` (Python standard library only) recomputes, in exact rational arithmetic and, for the target bounds, outward-rounded rational interval arithmetic, the data of Theorem 3.3 and every numerical input of the certificate in Theorem 3.5, including the target bounds (86).

AI assistance disclosure

OpenAI GPT-5.6 Sol, accessed through ChatGPT and Codex, was used to develop and audit the proofs, exact certificates and numerical experiments, and for drafting. Anthropic Claude Opus 5.5, used through Claude Code, was used for editing, an exact-arithmetic verification of the certificate in Theorem 3.5, the improved constants, and the ancillary checker. The author reviewed and takes responsibility for every statement and proof. AI systems are not authors of this work, and the work has not been peer reviewed.

References

[1] E. Roldán and J. M. R. Parrondo, Estimating dissipation from single stationary trajectories, *Phys. Rev. Lett.* **105**, 150607 (2010), doi:10.1103/PhysRevLett.105.150607.

[2] U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines, *Rep. Prog. Phys.* **75**, 126001 (2012), doi:10.1088/0034-4885/75/12/126001.

[3] D. J. Skinner and J. Dunkel, Improved bounds on entropy production in living systems, *Proc. Natl. Acad. Sci. USA* **118**, e2024300118 (2021), doi:10.1073/pnas.2024300118.

[4] D. J. Skinner and J. Dunkel, Estimating entropy production from waiting time distributions, *Phys. Rev. Lett.* **127**, 198101 (2021), doi:10.1103/PhysRevLett.127.198101.

[5] E. Nitzan, A. Ghosal and G. Bisker, Universal bounds on entropy production inferred from observed statistics, *Phys. Rev. Research* **5**, 043251 (2023), doi:10.1103/PhysRevResearch.5.043251.

[6] J. van der Meer, B. Ertel and U. Seifert, Thermodynamic inference in partially accessible Markov networks: A unifying perspective from transition-based waiting time distributions, *Phys. Rev. X* **12**, 031025 (2022), doi:10.1103/PhysRevX.12.031025.

[7] P. H. Diananda, On non-negative forms in real variables some or all of which are non-negative, *Math. Proc. Cambridge Philos. Soc.* **58**, 17–25 (1962), doi:10.1017/S0305004100036185.

[8] M. Hall, Jr. and M. Newman, Copositive and completely positive quadratic forms, *Math. Proc. Cambridge Philos. Soc.* **59**, 329–339 (1963), doi:10.1017/S0305004100036951.

[9] M. Pfeffer and J. A. Samper, The cone of 5×5 completely positive matrices, *Discrete Comput. Geom.* **71**, 442–466 (2024), doi:10.1007/s00454-023-00620-y.

[10] S. Petkov, *Entropy Production under Hidden-State Reduction: Exact Safety and Statistical Limits*, with Technical Supplement, unpublished manuscript, 4 October 2026. The inherited result is the supplement's exact five-cycle dissipation theorem; Appendix J reproduces a self-contained proof of the comparison value used here.

[11] W. Hoeffding, Probability inequalities for sums of bounded random variables, *J. Amer. Statist. Assoc.* **58**, 13–30 (1963), doi:10.1080/01621459.1963.10500830.

[12] C. J. Clopper and E. S. Pearson, The use of confidence or fiducial limits illustrated in the case of the binomial, *Biometrika* **26**, 404–413 (1934), doi:10.1093/biomet/26.4.404.

[13] T. S. Motzkin and E. G. Straus, Maxima for graphs and a new proof of a theorem of Turán, *Canad. J. Math.* **17**, 533–540 (1965), doi:10.4153/CJM-1965-053-6.

[14] M. Aguilera, S. Ito, and A. Kolchinsky, *Inferring Entropy Production in Many-Body Systems Using Nonequilibrium Maximum Entropy*, Physical Review Letters **136**, 077101 (2026), doi:10.1103/xgkj-dxzh.

[15] Anonymous, *An architecture-independent dissipation floor for alternating-renewal observations*, Zenodo preprint (unrefereed), version 1.1.0-candidate, 3 October 2026, doi:10.5281/zenodo.23118474. The fixed-window Corollary 7 does not require renewal.

[16] J. H. Fritz and U. Seifert, *Waiting-time-based entropy estimators in continuous space without Markovian events*, Physical Review E **114**, 034109 (2026), doi:10.1103/zj2b-w93z.

[17] N. D. Stein, A. Ozdaglar, and P. A. Parrilo, *Exchangeable Equilibria, Part I: Symmetric Bimatrix Games*, arXiv:1307.3586 (2013).

[18] C. Bădescu and R. O'Donnell, *Lower Bounds for Testing Complete Positivity and Quantum Separability*, in *LATIN 2020: Theoretical Informatics*, Lecture Notes in Computer Science **12118**, 375–386 (Springer, 2020), doi:10.1007/978-3-030-61792-9_30.

[19] A. F. Ducuara and P. Skrzypczyk, *Operational Interpretation of Weight-Based Resource Quantifiers in Convex Quantum Resource Theories*, Physical Review Letters **125**, 110401 (2020), doi:10.1103/PhysRevLett.125.110401.

[20] P. Raux, A. M. Maier, and U. Seifert, *Lower bounds on entropy production from dynamical correlation functions*, arXiv:2608.03619 (2026).

[21] J.-R. Chazottes, C. Giardina and F. Redig, Relative entropy and waiting times for continuous-time Markov processes, arXiv:math/0512386v1 (2005).

[22] S. K. Mitter, Markov Random Fields, Stochastic Quantization and Image Analysis, in R. Spigler (ed.), *Applied and Industrial Mathematics*, Mathematics and Its Applications **56**, 101–109 (Kluwer, 1991), doi:10.1007/978-94-009-1908-2_9.

[23] Y. Tu, The nonequilibrium mechanism for ultrasensitivity in a biological switch: Sensing by Maxwell's demons, *Proc. Natl. Acad. Sci. USA* **105**, 11737–11741 (2008), doi:10.1073/pnas.0804641105.

[24] S. W. Dharmadhikari, Sufficient Conditions for a Stationary Process to be a Function of a Finite Markov Chain, *Ann. Math. Statist.* **34**, 1033–1041 (1963), doi:10.1214/aoms/1177704026.

[25] A. Heller, On Stochastic Processes Derived From Markov Chains, *Ann. Math. Statist.* **36**, 1286–1291 (1965), doi:10.1214/aoms/1177700000.

[26] A. Klein, A generalization of Markov processes, *Ann. Probab.* **6**, 128–132 (1978), doi:10.1214/aop/1176995616.

[27] W. Arveson, Markov operators and OS-positive processes, *J. Funct. Anal.* **66**, 173–234 (1986), doi:10.1016/0022-1236(86)90071-6.

[28] P. E. T. Jorgensen, K.-H. Neeb and G. Ólafsson, Reflection Positive Stochastic Processes Indexed by Lie Groups, *SIGMA* **12**, 058 (2016), doi:10.3842/SIGMA.2016.058.

[29] C. Grussler and T. Damm, A symmetry approach for balanced truncation of positive linear systems, in *51st IEEE Conference on Decision and Control*, 4308–4313 (2012), doi:10.1109/CDC.2012.6426442.

[30] M. Wagner and J. Timmer, The Effects of Non-Identifiability on Testing for Detailed Balance in Aggregated Markov Models for Ion-Channel Gating, *Biophys. J.* **79**, 2918–2924 (2000), doi:10.1016/S0006-3495(00)76529-5.

[31] Y. Chen, On the reversibility of the observed process of three-state hidden Markov model, arXiv:1101.5482v1 (2011).

[32] B. Wu and C. Jia, Parameter Inference and Nonequilibrium Identification for Markov Networks Based on Coarse-Grained Observations, *Phys. Rev. Lett.* **134**, 087103 (2025), doi:10.1103/PhysRevLett.134.087103. The explicit binary rotation cited here is in Appendix A of the author preprint arXiv:2406.19586v2.

[33] I. Siekmann, Modelling ion channels with a view towards identifiability, *Bull. Math. Biol.* **88**, 2 (2026), doi:10.1007/s11538-025-01558-3.

[34] J. Gu, Excursion-Resolved Thermodynamic Inference without Observing Reverse Transitions, arXiv:2609.39205v2 (2026).

[35] J. W. Biddle, A method for calculating entropy production without knowing transition rates or dwell times, arXiv:2605.27635 (2026).

[36] D. Hartich and A. Godec, Comment on "Inferring broken detailed balance in the absence of observable currents", *Nat. Commun.* **15**, 8678 (2024), doi:10.1038/s41467-024-52602-0.

[37] G. Bisker, I. A. Martínez, J. M. Horowitz and J. M. R. Parrondo, Reply to: Comment on "Inferring broken detailed balance in the absence of observable currents", *Nat. Commun.* **15**, 8679 (2024), doi:10.1038/s41467-024-52603-z.

[38] M. Fanizza, J. Lumbreras and A. Winter, *Quantum Theory in Finite Dimension Cannot Explain Every General Process with Finite Memory*, *Commun. Math. Phys.* **405**, 50 (2024), doi:10.1007/s00220-023-04913-4.

[39] M. Esposito, Stochastic thermodynamics under coarse graining, *Phys. Rev. E* **85**, 041125 (2012), doi:10.1103/PhysRevE.85.041125.

[40] M. Polettini and M. Esposito, Effective thermodynamics for a marginal observer, *Phys. Rev. Lett.* **119**, 240601 (2017), doi:10.1103/PhysRevLett.119.240601.

[41] J. Ehrich, Tightest bound on hidden entropy production from partially observed dynamics, *J. Stat. Mech.* **2021**, 083214 (2021), arXiv:2105.08803.

[42] I. A. Martínez, G. Bisker, J. M. Horowitz and J. M. R. Parrondo, Inferring broken detailed balance in the absence of observable currents, *Nat. Commun.* **10**, 3542 (2019), doi:10.1038/s41467-019-11051-w.

[43] D. R. Fredkin and J. A. Rice, On aggregated Markov processes, *J. Appl. Probab.* **23**(1), 208–214 (1986), doi:10.2307/3214130.

[44] P. Kienker, Equivalence of aggregated Markov models of ion-channel gating, *Proc. R. Soc. Lond. B* **236**, 269–309 (1989), doi:10.1098/rspb.1989.0024.

[45] M. Fliess, Séries rationnelles positives et processus stochastiques, *Ann. Inst. H. Poincaré Sect. B* **11**, 1–21 (1975), Numdam original.

[46] M. Aznagulov, Thermodynamic Realizability of Hidden Markov Processes: Attainment, Observable Certificates, and the Price of Architecture, arXiv:2609.14623v1 (2026).

[47] P. L. Combettes, Monotone operator theory in convex optimization, *Math. Program.* **170**, 177–206 (2018), author manuscript.

[48] D. A. Levin and Y. Peres, *Markov Chains and Mixing Times*, second edition (2017), with contributions by E. L. Wilmer, author-hosted second edition.

A Reflection reconstruction and the low-rank theorem

We prove Theorem 3.1 directly from the observed law. The Hilbert-space reconstruction is standard reflection-positive structure [26, 28]; the details below also specify the finite-dimensional and label-compatible contract used here.

A.1 The finite reflected representation

Let $\mathcal{C}$ be the real span of bounded future cylinder functions. Reversibility makes (2) symmetric; reflection positivity makes it a semidefinite form. For a supplied n-state CTMC, midpoint conditioning writes every reflected cylinder matrix as a product through $\mathbb{R}^n$. Consequently the quotient $\mathcal{H}$ is finite-dimensional, of dimension $r \le n$. Write $[f]$ for a class and $v = [1]$. Then $\|v\| = 1$.

Time translation $U_t f$ satisfies $[U_t f, g]_R = [f, U_t g]_R$ by stationarity and reversal. This identity shows that a null function remains null after translation. Hence U_t descends to a self-adjoint semigroup T_t on $\mathcal{H}$. Stochastic continuity of the supplied finite CTMC makes its matrix entries continuous in a cylinder basis,

so $T_t = e^{tA}$ with $A = A^{\mathsf{T}}$ and $Av = 0$. For bounded representatives the entries $[f, U_t g]_R$ are uniformly bounded in t. In finite dimension this excludes positive eigenvalues, giving $A \le 0$.

Multiplication by $1_{\{Y_0=a\}}$ is self-adjoint for the reflected form and therefore preserves its null space. It descends to an orthogonal projection E_a. These projections are disjoint and sum to the identity. Set $\mathcal{H}_a = E_a\mathcal{H}$, $v_a = E_a v$ and $p_a = \|v_a\|^2 > 0$. For positive time increments, the cylinder probabilities are

$$P(Y_0 = a_0, Y_{t_1} = a_1, \dots, Y_{t_1+\cdots+t_m} = a_m) = v^{\mathsf{T}} E_{a_0} e^{t_1 A} E_{a_1} \cdots e^{t_m A} E_{a_m} v. \tag{29}$$

This follows by representing the future cylinder as the indicated product applied to v, and pairing it with v. Mixing of the supplied finite irreducible CTMC gives $[f, U_t g]_R \to \mathbb{E}f\mathbb{E}g$. Thus $e^{tA} \to vv^{\mathsf{T}}$ and $\ker A = \operatorname{span}\{v\}$.

Cylinder indicators are dense in ordinary half-path $L^2(P)$. For their differences, Cauchy–Schwarz and reversal give $0 \le [f, f]_R \le \|f\|^2_{L^2(P)}$. Therefore measurable finite-window events have classes in the same closed, finite-dimensional quotient. No additional rank arises when classifiers are allowed to be measurable rather than cylindrical.

A.2 The conditional-rank dictionary at every sampling interval

For $\Delta > 0$, let $n_\Delta(a)$ be the rank of the word matrix $[P(s\,a\,t)]$ on the Δ grid, with arbitrary finite past and future words, including empty words. Reversing the past-word row indexing identifies this matrix with the reflected Gram matrix of the a-restricted future grid-word vectors. Hence its rank is their span dimension.

Let W_Δ be the span of all full future grid-word vectors $E_{a_0} T_\Delta E_{a_1} \cdots T_\Delta E_{a_m} v$. It contains v, is invariant under every E_a, and is invariant under T_Δ: prepending an arbitrary current label and summing over that label gives the shifted word. If $\lambda_1, \dots, \lambda_q$ are the distinct real eigenvalues of A, the numbers $e^{\Delta\lambda_j}$ are distinct. Real polynomial interpolation therefore expresses every T_t as a polynomial in T_Δ. Explicitly, the polynomial

$$p_t(z) = \sum_{j=1}^{q} e^{t\lambda_j} \prod_{\ell \ne j} \frac{z - e^{\Delta\lambda_\ell}}{e^{\Delta\lambda_j} - e^{\Delta\lambda_\ell}} \tag{30}$$

satisfies $T_t = p_t(T_\Delta)$. Thus W_Δ is invariant under all T_t and label masks. These products from v span every future cylinder vector, so $W_\Delta = \mathcal{H}$. Its a-restricted word span is $E_a W_\Delta = \mathcal{H}_a$, proving $n_\Delta(a) = \dim \mathcal{H}_a$ for every $\Delta > 0$. In particular, at $r \le d + 1$, Dharmadhikari's low-conditional-rank corollary already gives an ordinary r-state discrete-time realization of each sampled law. The remaining continuous-time detailed-balance obligation is addressed below, rather than by choosing unrelated sampled representations or matrix logarithms.

A.3 Future-event cones and killed semigroups

For each label let C_a be the closed conical hull of the vectors $x_F = [1_F]$ for future cylinder events contained in $\{Y_0 = a\}$. These vectors span $\mathcal{H}_a$, and

$$\langle x_F, x_G \rangle = P(\theta F \cap G) \ge 0. \tag{31}$$

Thus C_a is acute and pointed. It contains v_a and $v_a - x_F$. Choose an event-vector basis $x_1, \dots, x_k$. For $\sum_i |c_i| \le 1$, each vector $v_a + \sum_i c_i x_i$ is a nonnegative combination of v_a, x_i and $v_a - x_i$: use $v_a + x_i$ for positive coefficients and $v_a - x_i$ for negative coefficients. Hence v_a is interior to C_a. Acuteness and an interior ball imply $\langle v_a, x \rangle \ge \epsilon \|x\|$ on C_a for some $\epsilon > 0$: pair x with $v_a - \epsilon x/\|x\| \in C_a$.

Put $A_a = E_a A E_a|_{\mathcal{H}_a}$. The products $(E_a e^{tA/n} E_a)^n$ preserve C_a, because on an event vector they require label a at the prefix grid times, followed by the shifted event. They converge on $\mathcal{H}_a$ to e^{tA_a}. Closure gives invariance of C_a; self-adjointness gives invariance of its dual C_a^*. For distinct labels, the short-time event vectors give

$$E_a A v_b = \lim_{t \downarrow 0} t^{-1} E_a e^{tA} v_b \in C_a. \tag{32}$$

A.4 An orthogonal positive realization

Since every active label space is nonzero, $r \ge d$. If $r = d$, choose its unit vectors $s_a = v_a/\sqrt{p_a}$. Equation (32) makes every off-diagonal entry of A in this basis nonnegative.

If $r = d + 1$, exactly one label space, say $\mathcal{H}_a$, has dimension two. Choose orthonormal coordinates $s = v_a/\sqrt{p_a}$ and u. Its normalized primal cone slice is

$$C_a \cap \{\langle s, x\rangle = 1\} = \{s + yu : \ell \le y \le h\}, \qquad \ell < 0 < h, \quad 1 + \ell h \ge 0. \tag{33}$$

Compactness follows from the preceding strict directional bound; the endpoint inequality follows from acuteness. The dual slice has endpoints $L = -1/h < 0$ and $U = -1/\ell > 0$, with $-LU \ge 1$. Write A_a in this basis as $\left(\begin{smallmatrix} \alpha & \gamma \\ \gamma & \delta \end{smallmatrix}\right)$. The initial projective derivative under e^{tA_a} is

$$F(y) = \gamma + (\delta - \alpha)y - \gamma y^2. \tag{34}$$

Dual-cone invariance gives $F(U) \le 0$ at its upper endpoint. With $k = \sqrt{1 + U^2}$ choose orthonormal rows

$$j_- = (Us - u)/k, \qquad j_+ = (s + Uu)/k. \tag{35}$$

Their slopes $-1/U$ and U both belong to $[L, U]$, so both are in C_a^*. Their transformed internal entry is

$$\langle j_-, A_a j_+\rangle = \frac{U(\alpha - \delta) + \gamma(U^2 - 1)}{1 + U^2} = -\frac{F(U)}{1 + U^2} \ge 0. \tag{36}$$

Equation (32) makes their entries against every other unit label vector s_b nonnegative. Entries between the other one-dimensional blocks are nonnegative for the same reason. No sign restriction on γ is required.

Let J be this block-orthogonal change of basis, using the unit label vectors on all other blocks. Then $B = JAJ^{\mathsf{T}}$ is symmetric Metzler and $w = Jv > 0$. In the two-dimensional block the coordinates are $\sqrt{p_a}U/k$ and $\sqrt{p_a}/k$. With $D_w = \operatorname{diag}(w)$ define

$$Q = D_w^{-1} B D_w, \qquad \pi_i = w_i^2. \tag{37}$$

Here $Q\mathbf{1} = 0$, its off-diagonals are nonnegative, and $\pi_i q_{ij} = w_i B_{ij} w_j = \pi_j q_{ji}$. Thus Q is a conservative detailed-balance row generator. Its graph is connected: each connected component of a symmetric Metzler matrix with positive null vector contributes a null vector, whereas $\ker B$ is one-dimensional.

The change of basis preserves each label space. If P_a denotes the corresponding diagonal label mask, then $P_a J = J E_a$. Conjugating the semigroups in (37) gives

$$\pi P_{a_0} e^{t_1 Q} P_{a_1} \cdots e^{t_m Q} P_{a_m} \mathbf{1} = v^{\mathsf{T}} E_{a_0} e^{t_1 A} E_{a_1} \cdots e^{t_m A} E_{a_m} v. \tag{38}$$

These are all the original cylinders, on the original clock. Coordinate cylinders determine the finite-alphabet cadlag law. Finally any n-state CTMC realization factors the reflected matrices through $\mathbb{R}^n$, so $n \ge r$. This proves both exact realization and minimality.

A.5 A nonrenewal example at rank $d + 1$

Take four states with uniform stationary masses, labels (a, a, b, c), and symmetric rates $q_{12} = 1$, $q_{13} = 3$, $q_{14} = 1$, $q_{23} = 1$, $q_{24} = 3$, $q_{34} = 2$. The two distinct probe labels give two independent prediction directions; within a, $(1, 1)$ and $(3, 1)$ span its two coordinates. Thus $d = 3$ and $r = 4$. The killed a-block is $\left(\begin{smallmatrix} -5 & 1 \\ 1 & -5 \end{smallmatrix}\right)$, whose next-exit-to-$b$ probabilities are $(2/3, 1/3)$. Entry from b gives hidden weights $(3/4, 1/4)$, whereas entry from c gives $(1/4, 3/4)$. The probability of leaving a next for b is consequently $7/12$ or $5/12$, respectively. Entrance-label dependence is compatible with the theorem; binary alternating renewal is not assumed.

B Joint compression of the binary rank-four profile

We prove Theorem 3.2. The generator acts on column functions. At detailed balance, conditional expectation of a future cylinder f given X_0 satisfies $[f, g]_R = \langle v_f, v_g\rangle_\pi$ by midpoint conditioning. Their span contains $\mathbf{1}$ and is invariant under label masks and e^{tQ}. Finite dimension permits differentiation at zero, so it is Q-invariant. Conversely, the smallest space with these invariances contains every future cylinder vector. Thus that span is V, and V is isometric to $\mathcal{H}$; in particular $r = 4$.

Features and the complete cross action. Put $\alpha = \pi(a) > 0$, $\beta = \pi(b) > 0$ and use conditional stationary measures $\mu_i = \pi_i/\alpha$ on a and $\nu_j = \pi_j/\beta$ on b. Choose features x, y supported on their respective labels with conditional mean zero and second moment one. Each $f \in V$ is uniquely $(c_a + ux, c_b + vy)$. Write their extrema as $L_x < 0 < R_x$, $L_y < 0 < R_y$. Since $\mathbb{E}_\mu[(R_x - X)(X - L_x)] = -1 - L_x R_x \geq 0$, and likewise for Y,

$$\begin{aligned} -L_x R_x \geq 1, \qquad -L_y R_y \geq 1, \\ I_x = [-1/R_x, -1/L_x] \subset [L_x, R_x], \qquad I_y = [-1/R_y, -1/L_y] \subset [L_y, R_y]. \end{aligned} \tag{39}$$

Invariance of V gives the affine total cross rate and destination feature moment

$$\begin{aligned} k_a(x_i) &= \sum_{j \in b} q_{ij} = \kappa_0 + \kappa_x x_i, \\ m_a(x_i) &= \sum_{j \in b} q_{ij} y_j = \kappa_y + \kappa_{xy} x_i, \\ H(x, y) &= \kappa_0 + \kappa_x x + \kappa_y y + \kappa_{xy} xy. \end{aligned} \tag{40}$$

For $C_{ij} = \pi_i q_{ij} = \pi_j q_{ji}$, the four sums $\sum_{ij} C_{ij}(1, x_i, y_j, x_i y_j)$ equal $\alpha(\kappa_0, \kappa_x, \kappa_y, \kappa_{xy})$. The reverse total rate and destination moment are affine in y by the same invariance. Taking their ν inner products with $1, y$ therefore determines them pointwise:

$$k_b(y) = \frac{\alpha}{\beta}(\kappa_0 + \kappa_y y), \qquad m_b(y) = \frac{\alpha}{\beta}(\kappa_x + \kappa_{xy} y). \tag{41}$$

Irreducibility and the nonempty partition give a cross edge, so $\kappa_0 > 0$. All features are dimensionless; the four coefficients, drifts and generator entries have inverse-time units.

Joint selection of the two replacement pairs. Nonnegative original rates give $L_y k_a(x_i) \leq m_a(x_i) \leq R_y k_a(x_i)$. The two differences are affine in x, hence nonnegative on the whole interval between its attained extrema. Consequently

$$H(x, -1/R_y) = \frac{R_y k_a(x) - m_a(x)}{R_y} \geq 0, \qquad H(x, -1/L_y) = \frac{m_a(x) - L_y k_a(x)}{-L_y} \geq 0. \tag{42}$$

Affinity gives $H \geq 0$ on $[L_x, R_x] \times I_y$. Applying the reverse destination-mean bounds in (41) gives $H \geq 0$ on $I_x \times [L_y, R_y]$. No pointwise escape rate is divided out; zero cross rows are allowed.

Bilinear interpolation expresses $H(0, 0) = \kappa_0 > 0$ as a strictly positive convex combination of its four outer-corner values. Choose an outer corner (x_*, y_*) with positive H and replace each feature law by the two points $\{x_*, -1/x_*\}$ and $\{y_*, -1/y_*\}$. The reciprocal points are endpoints of I_x, I_y. Thus the other three new product corners lie in the nonnegative strips. Naming the points $x_+ > 0 > x_-$ and $y_+ > 0 > y_-$, give them weights

$$\begin{aligned} x_+ x_- = -1, \qquad \mu'_+ = \frac{1}{1 + x_+^2}, \quad \mu'_- = \frac{x_+^2}{1 + x_+^2}, \\ y_+ y_- = -1, \qquad \nu'_+ = \frac{1}{1 + y_+^2}, \quad \nu'_- = \frac{y_+^2}{1 + y_+^2}. \end{aligned} \tag{43}$$

These positive measures match mass, mean and second moment. The retained corner must be selected jointly: independent choices of endpoints would not establish its cross-rate sign.

Internal rates and generator intertwining. Let $T_a = Q_{aa} + \mathrm{diag}(k_a(x_i))$ be the conservative internal generator on a. It is reversible for μ even if its graph is disconnected. Since $(Qx)|_a = h_a(x)$ is affine,

$$T_a x = D_a(x), \qquad D_a(x) = h_a(x) + k_a(x) x, \qquad \mathbb{E}_\mu D_a = 0. \tag{44}$$

This drift has degree at most two, so the replacement measure also has $\mathbb{E}_{\mu'} D_a = 0$. At a retained original maximum the CTMC maximum principle gives $D_a(x_*) \leq 0$; at a retained minimum it gives $D_a(x_*) \geq 0$. Positive weights and zero mean drift therefore give $D_a(x_+) \leq 0 \leq D_a(x_-)$ in either case. The rates

$$q'_{a_+, a_-} = -\frac{D_a(x_+)}{x_+ - x_-}, \qquad q'_{a_-, a_+} = \frac{D_a(x_-)}{x_+ - x_-} \tag{45}$$

are nonnegative, realize the drift and satisfy detailed balance for μ'. The identical construction with $D_b(y) = h_b(y) + k_b(y)y$, where $(Qy)|_b = h_b(y)$, gives both internal b rates. Zero drifts remove an internal edge without altering the construction.

Set $\pi'_{a_e} = \alpha\mu'_e$, $\pi'_{b_f} = \beta\nu'_f$ for $e, f \in \{+,-\}$, and set the cross rates to

$$q'_{a_e,b_f} = \nu'_f H(x_e, y_f), \qquad q'_{b_f,a_e} = \frac{\alpha}{\beta}\mu'_e H(x_e, y_f). \tag{46}$$

They are nonnegative and have symmetric conductances $\alpha\mu'_e\nu'_f H(x_e, y_f)$. Choosing diagonals as minus outgoing sums defines a conservative detailed-balance generator Q'. The first two moments of ν' give $\sum_f q'_{a_e,b_f} = k_a(x_e)$ and $\sum_f q'_{a_e,b_f} y_f = m_a(x_e)$; the moments of μ' give both reverse actions in (41).

Let U evaluate $f = (c_a + ux, c_b + vy)$ at the four new points. At a_e, the action on Uf is

$$\begin{aligned}(Q'Uf)_{a_e} &= uD_a(x_e) + (c_b - c_a - ux_e)k_a(x_e) + vm_a(x_e)\\ &= uh_a(x_e) + (c_b - c_a)k_a(x_e) + vm_a(x_e) = (UQf)_{a_e}.\end{aligned}$$

At b_f the same calculation gives $vh_b(y_f) + (c_a - c_b)k_b(y_f) + um_b(y_f) = (UQf)_{b_f}$. Thus every row action, including its escape rate, is preserved:

$$Q'U = U(Q|_V). \tag{47}$$

Full law, clock and minimality. The two distinct points in each label make U a bijection onto all four-state functions. It preserves $\mathbf{1}$ and both label masks. Moment matching also gives, for $g = (\widetilde{c}_a + \widetilde{u}x, \widetilde{c}_b + \widetilde{v}y)$,

$$\langle f, g\rangle_\pi = \alpha(c_a\widetilde{c}_a + u\widetilde{u}) + \beta(c_b\widetilde{c}_b + v\widetilde{v}) = \langle Uf, Ug\rangle_{\pi'}. \tag{48}$$

Exponentiating (47) at the same real time and inserting the label masks transports every chronological cylinder function. Pairing with $\mathbf{1}$ through (48) preserves its stationary probability. The finite-alphabet cadlag laws are therefore equal on the original physical clock.

Irreducibility of Q gives $\ker(Q|_V) = \operatorname{span}\{\mathbf{1}\}$, so the bijective intertwining gives the same kernel for Q'. The detailed-balance Dirichlet form $-\langle f, Q'f\rangle_{\pi'} = \frac{1}{2}\sum_{i,j}\pi'_i q'_{ij}(f_i - f_j)^2$ has one kernel dimension per connected component. Hence Q' is irreducible, including cases with missing internal edges or zero cross corners. The reflected rank is four, and midpoint conditioning bounds it by the state count of any finite CTMC presentation. Four is therefore minimal.

C A common CP hierarchy with no finite equilibrium realization

We prove Theorem 3.3 by an explicit rational construction. All displayed numerical rates use one fixed inverse-time unit.

C.1 The core regression and its separating identity

Index five points by $i = 0, \dots, 4$ and put

$$\begin{gathered}p_0 = (1,0), \quad p_1 = (0,1), \quad p_2 = (-1,0),\\ p_3 = (-3/5, -4/5), \quad p_4 = (3/5, -4/5), \qquad \omega = (5,8,5,5,5)/28.\end{gathered} \tag{49}$$

They have weighted mean zero, covariance $\Sigma = \operatorname{diag}(17/35, 18/35)$, and unit norm. Define the directed stationary flows by the following table; unlisted entries are zero:

(i,j)	F_{ij}	(i,j)	F_{ij}	(i,j)	F_{ij}
$(0,1)$	$11/126$	$(0,3)$	$5/126$	$(0,4)$	$5/72$
$(1,3)$	$25/63$	$(1,4)$	$25/63$	$(2,1)$	$11/126$
$(2,3)$	$5/72$	$(2,4)$	$5/126$	$(3,1)$	$13/42$
$(3,2)$	$11/56$	$(4,0)$	$11/56$	$(4,1)$	$13/42$

Direct rational calculation (reproduced by the ancillary checker `anc/check_certificate.py`) gives $\sum_j F_{ij} = \sum_j F_{ji}$ and $\sum_j F_{ij}(p_j - p_i) = \omega_i \operatorname{diag}(-1,-5)p_i$. For $i \neq j$ set

$$R_{ij} = F_{ij}/\omega_i + \omega_j/10, \qquad R_{ii} = -\sum_{j\neq i} R_{ij}. \tag{50}$$

The refresh term preserves ω and makes every off-diagonal positive. Thus R is stationary, irreducible and reciprocal, and its coordinate drift is Lp_i, where $L = \mathrm{diag}(-\alpha, -\beta)$, $\alpha = 11/10$ and $\beta = 51/10$.

No symmetric nonnegative conductances $C_{ij} = C_{ji} \geq 0$ can satisfy

$$\sum_{j \neq i} C_{ij}(p_j - p_i) = \omega_i L p_i. \tag{51}$$

Indeed choose $y_0 = y_1 = y_4 = 0$, $y_2 = (-35/3, 35/3)$ and $y_3 = (-21, 7)$. For unordered pairs let $d_{ij} = (y_i - y_j) \cdot (p_j - p_i)$. All $d_{ij} \leq 0$; their only nonzero values are

$$d_{02} = -70/3, \quad d_{03} = d_{24} = -28, \quad d_{34} = -126/5. \tag{52}$$

But $\sum_i \omega_i y_i \cdot L p_i = \beta - 13\alpha/3 = 1/3$. Multiplying (51) by y_i and summing would give $1/3 = \sum_{i<j} C_{ij} d_{ij} \leq 0$. This is an exact separating certificate.

C.2 Facet probes and the ten-state generator

Let $K = \mathrm{conv}\{p_0, \ldots, p_4\}$. Its facet functions are

$$\begin{aligned} g_1(x,y) &= 2(1 - x - y), & g_2(x,y) &= 2(1 + x - y), \\ g_3(x,y) &= 2 + 2x + y, & g_4(x,y) &= 8/5 + 2y, \\ g_5(x,y) &= 2 - 2x + y. \end{aligned} \tag{53}$$

Their simultaneous nonnegativity describes exactly K: for $0 \leq y \leq 1$ it gives $y - 1 \leq x \leq 1 - y$; for $-4/5 \leq y \leq 0$ it gives $-1 - y/2 \leq x \leq 1 + y/2$. Write $g_\ell = b_\ell + m_\ell \cdot p$, with $b = (2, 2, 2, 8/5, 2)$. The normals sum to zero and span the plane, while $\sum_\ell g_\ell = \lambda = 48/5$.

Use five core states a_i and five probe states b_ℓ. With $G_{i\ell} = g_\ell(p_i)$ and $D_\omega = \mathrm{diag}(\omega)$, $D_b = \mathrm{diag}(b)$, set

$$Q_0 = \begin{pmatrix} R - \lambda I_5 & G \\ 5G^{\mathsf{T}} D_\omega & -5D_b \end{pmatrix}, \qquad \pi_0 = (\omega/2, 1/10, 1/10, 1/10, 1/10, 1/10). \tag{54}$$

The core cross-row sums are λ; the probe cross-row sums are $5b_\ell$, by the zero weighted mean of the points. Thus $Q_0 \mathbf{1} = 0$. Its cross stationary flows match pairwise, $(\omega_i/2) g_\ell(p_i) = (1/10) 5 \omega_i g_\ell(p_i)$. Together with $\omega R = 0$, this proves $\pi_0 Q_0 = 0$. Each probe is connected reciprocally to three core states, and the core graph is complete. Hence Q_0 is irreducible and reciprocal. Its core flux difference across (a_0, a_1) is $F_{01}/2 = 11/252$, so it is not detailed balanced. Give all core states label a and each probe its distinct label b_ℓ. Let P be this actual stationary observed path law.

C.3 Self-adjoint observed dynamics and the common CP factor

Let $\mathcal{W}$ consist of functions $f(a_i) = c + u \cdot p_i$ and $f(b_\ell) = z_\ell$. Its dimension is eight; it contains 1 and is preserved by every label mask. The generator action is

$$\begin{aligned} c' &= -\lambda c + \sum_\ell b_\ell z_\ell, \\ u' &= (L - \lambda I_2) u + \sum_\ell m_\ell z_\ell, \\ z'_\ell &= 5 b_\ell c + 5 m_\ell^{\mathsf{T}} \Sigma u - 5 b_\ell z_\ell. \end{aligned} \tag{55}$$

The stationary inner product is

$$\langle f, \widetilde{f} \rangle_{\pi_0} = \tfrac{1}{2} c \widetilde{c} + \tfrac{1}{2} u^{\mathsf{T}} \Sigma \widetilde{u} + \tfrac{1}{10} \sum_\ell z_\ell \widetilde{z}_\ell. \tag{56}$$

Equations (55)–(56) show that $Q_0|_{\mathcal{W}}$ is self-adjoint in this actual anisotropic metric. Each label mask is self-adjoint there too.

Every conditional future-cylinder vector belongs to $\mathcal{W}$. Conversely, the current probe indicators supply five directions, and the limits $t^{-1} E_a e^{tQ_0} 1_{b_\ell} \to 1_a g_\ell$ supply its three-dimensional core block. Their span is closed in finite dimension, so the complete future-vector span is $\mathcal{W}$. Reversing a chronological cylinder product takes its adjoint on $\mathcal{W}$ and preserves its scalar expectation between copies of 1. This proves

full observed reversibility. For future cylinder functionals F, G, let $f_F(s) = \mathbb{E}[F \mid X_0 = s]$. The same restricted-adjoint identity gives

$$\mathbb{E}_P[(\theta F)G] = \langle f_F, f_G \rangle_{\pi_0}. \tag{57}$$

It does not identify the actual microscopic conditional past vector with f_F; Q_0 is not globally self-adjoint. Cylinder density, conditional-expectation contraction and finite-dimensional closure extend (57) to bounded measurable half-path functionals. For every finite classifier its conditional cell probabilities $f_i(s)$ are nonnegative, and therefore

$$M_{ij} = \sum_{s=1}^{10} (\pi_0)_s f_i(s) f_j(s). \tag{58}$$

The same ten coordinates factor every window and every classifier. The positive-definite metric on the full eight-dimensional span also proves that the reflected rank is eight.

C.4 Rigidity of every competing equilibrium realization

Suppose a finite stationary detailed-balance row generator $\widehat{Q}$, stationary law $\widehat{\pi}$ and deterministic labels give exactly P on the same clock. Retain only positive-mass states; detailed balance forbids transitions from those states into zero-mass states. Its future-event vectors $\widehat{f}_F$ have the same reflected Gram kernel as (57). Thus $f_F \mapsto \widehat{f}_F$ extends to a well-defined isometry $U : \mathcal{W} \to L^2(\widehat{\pi})$. Current-label restrictions and translated future events give

$$U1 = 1, \quad UE_\ell = \widehat{E}_\ell U, \quad Ue^{tQ_0}|_{\mathcal{W}} = e^{t\widehat{Q}}U, \quad UQ_0|_{\mathcal{W}} = \widehat{Q}U. \tag{59}$$

The last identity follows by finite-matrix differentiation at zero. No state-count bound or product-preserving property of U is assumed.

Let x, y be the core-supported coordinate functions. At a competing state s with label a put $\xi_s = ((Ux)(s), (Uy)(s))$. Applying (59) to each probe indicator gives

$$g_\ell(\xi_s) = (\widehat{Q}1_{b_\ell})(s) = \sum_{t:\widehat{h}(t)=b_\ell} \widehat{q}_{st} \geq 0. \tag{60}$$

Hence every $\xi_s \in K$, even with arbitrary probe multiplicities. Isometry fixes the conditional core weights $\mu_s = 2\widehat{\pi}_s$ and moments:

$$\sum_{s:a} \mu_s = 1, \qquad \sum_{s:a} \mu_s \xi_s = 0, \qquad \sum_{s:a} \mu_s \xi_s \xi_s^{\mathsf{T}} = \Sigma. \tag{61}$$

In particular $\sum \mu_s |\xi_s|^2 = 1$. The polygon lies in the unit disk, and only its five vertices have norm one. For a convex combination of distinct unit points, $|\sum_i t_i p_i|^2 = 1 - \sum_{i<j} t_i t_j |p_i - p_j|^2 < 1$. All $\mu_s > 0$, so radius saturation forces each ξ_s to be one of the five vertices.

Let S_i be the competing core states at vertex p_i, and let $\widehat{\omega}_i = \sum_{s \in S_i} \mu_s$. The xy moment gives $\widehat{\omega}_3 = \widehat{\omega}_4$; the x mean gives $\widehat{\omega}_0 = \widehat{\omega}_2$; the y mean gives $\widehat{\omega}_1 = (8/5)\widehat{\omega}_3$. Normalization and the x^2 moment then give $\widehat{\omega} = \omega$. These are inner-product identities; they do not apply U to quadratic products or assume that it preserves multiplication.

The facet sum fixes total exit to all probe labels at each core state to λ. The original coordinate drift is $(L - \lambda I_2)p_i$; since the coordinates vanish off the core, cancelling that loss in (59) gives

$$\sum_{t:a} \widehat{q}_{st}(\xi_t - p_i) = Lp_i, \qquad s \in S_i. \tag{62}$$

Define the aggregated conditional stationary conductances

$$C_{ij} = 2 \sum_{s \in S_i, t \in S_j} \widehat{\pi}_s \widehat{q}_{st}, \qquad i \neq j. \tag{63}$$

They are symmetric and nonnegative by detailed balance. Multiplying the drift identity by $2\widehat{\pi}_s$ and summing over S_i gives precisely (51), contradicting its rational separating certificate. Transitions within a vertex class contribute zero; vertex grouping need not be a Markov lumping. The argument excludes every finite detailed-balance realization of the complete same-clock law, and completes the proof.

D Half-path factorability and symmetric extensions

In this appendix $W = \mathcal{W}$ is the observable space of Appendix C. The extension corollary proves Corollary 3.4.

Lemma D.1 (Half-path factorability)**.** *Let Y be a stationary reversible finite-alphabet càdlàg process. On $\mathcal{L} = L^2(\mathcal{F}_+, \mathbb{P})$, let $\mathsf{K} = E_+\theta|_{\mathcal{L}}$ be its reflected half-path operator. Every matched finite-window partition matrix is classically completely positive if and only if $\mathsf{K} = B^*B$ for an order-preserving bounded operator $B : \mathcal{L} \to L^2(\nu)$.*

Proof. Let P_n be conditional expectation onto increasing finite coordinate fields whose union generates $\mathcal{F}_+$. Their atoms A_i of positive probability p_i give the positive orthonormal basis $e_i = \mathbf{1}_{A_i}/\sqrt{p_i}$. In this basis the matrix of $P_n\mathsf{K}P_n$ is

$$C_{ij} = \frac{\mathbb{P}(\theta A_i \cap A_j)}{\sqrt{p_i p_j}}.$$

Complete positivity is preserved by this positive diagonal rescaling. Writing $C = Z^{\mathsf{T}}Z$ with $Z \geq 0$ gives $P_n\mathsf{K}P_n = \sum_\ell h_\ell \otimes h_\ell$, where $h_\ell = \sum_i Z_{\ell i}e_i \geq 0$. The compressions converge strongly to K and have norm at most one, hence also converge in the σ-weak topology. Arveson's closed-cone characterization of factorability then applies [27, Theorem 1.2.7].

Conversely, if $\mathsf{K} = B^*B$, put $z_i = B\mathbf{1}_{A_i} \geq 0$. The tested matrix is $M_{ij} = \int z_i z_j \, d\nu$. Normalize by $w = \sum_i z_i$: on $\{w > 0\}$, the vector $(z_i/w)_i$ lies in the probability simplex, and $w^2 d\nu$ is a probability measure because $\langle \mathbf{1}, \mathsf{K}\mathbf{1}\rangle = 1$. Thus M belongs to the compact convex hull of $\{uu^{\mathsf{T}} : u \geq 0,\ \sum_i u_i = 1\}$. A finite-dimensional convex decomposition supplies a finite nonnegative Gram factor. □

Lemma D.2 (Finite-valued reversible reduction)**.** *Let $(T_t)_{t\geq 0}$ be a self-adjoint Markov contraction semigroup on $L^2(\mu)$, with μ a probability measure. Let V be a finite-dimensional real invariant subspace containing $\mathbf{1}$ and invariant under every deterministic label multiplier. Assume continuity of $T_t|_V$ at zero, and write $T_t|_V = e^{tA_V}$. Suppose that its functions are constant on a finite measurable partition. Then the observed law has an exact finite detailed-balance CTMC realization at the same clock.*

Proof. Remove null classes, and merge classes with identical joint values of a basis $h = (h_1, \ldots, h_k)$ of V. Write these distinct value vectors as v_i and their positive masses as μ_i. The finite-dimensional restriction gives $(T_t h - h)/t \to A_V h$ in L^2 for every $h \in V$. For $i \neq j$, set

$$c_{ij}(t) = \frac{\langle \mathbf{1}_{S_i}, T_t \mathbf{1}_{S_j}\rangle}{t}.$$

These quantities are symmetric and nonnegative. Since $T_t\mathbf{1} = \mathbf{1}$, an exact expansion gives

$$E_t := \sum_{\ell=1}^{k} \frac{\langle h_\ell, (I - T_t)h_\ell\rangle}{t} = \frac{1}{2}\sum_{i\neq j} c_{ij}(t)|v_i - v_j|^2 \longrightarrow -\sum_\ell \langle h_\ell, A_V h_\ell\rangle.$$

If there is only one class, the law is constant and its one-state realization is immediate. Otherwise the minimum separation between the v_i is positive. The boundedness of E_t therefore bounds every $c_{ij}(t)$. A common subsequence $t_n \downarrow 0$ yields finite, symmetric, nonnegative limits c_{ij}. Moreover,

$$\sum_{j\neq i} c_{ij}(t)(v_j - v_i) = \int_{S_i} \frac{T_t h - h}{t}\, d\mu \longrightarrow \int_{S_i} A_V h \, d\mu.$$

Define $q_{ij} = c_{ij}/\mu_i$ for $i \neq j$ and choose the diagonal for zero row sums. Then Q is a finite conservative row generator and $\mu_i q_{ij} = \mu_j q_{ji}$. Since $A_V h \in V$ is class-constant, the last identity says $Qh = A_V h$ on each class. Consequently e^{tQ} and T_t agree on V. The labels are constant on the classes, and label-mask invariance keeps all chronological cylinder products starting from $\mathbf{1}$ in V. Their stationary probabilities therefore agree in both realizations. This proves equality of the full observed law.

The partition indicators were used only as L^2 test functions; no generator-domain condition on them, hidden finite jump rate, or strong lumpability was required. □

Corollary D.3 (Exclusion of arbitrary-state equilibrium extensions)**.** *The rational ten-state, six-label construction has no exact realization as an instantaneous deterministic Borel observation of a stationary two-sided ordinary symmetric Markov process on any standard Borel state space at the same continuous clock.*

Proof. Suppose such a realization exists. Its conditional future vectors and those of the original model have the same reflected Gram kernel. Their correspondence therefore gives an isometry U from the original eight-dimensional observable space W into $L^2(\mu)$. Its transition operators T_t are self-adjoint Markov contractions. The isometry preserves $\mathbf{1}$, intertwines the label masks, and satisfies

$$T_t U = U e^{tQ_0}|_W.$$

Thus $V = U(W)$ is an invariant finite-dimensional space with the explicit analytic restriction $T_t|_V = e^{tA_V}$, where $A_V = U(Q_0|_W)U^{-1}$. On the core label a, set $\xi = (Ux, Uy)$. For each probe label b_ℓ, positivity of the transition semigroup gives

$$\mathbf{1}_a A_V \mathbf{1}_{b_\ell} = \lim_{t\downarrow 0} \frac{\mathbf{1}_a T_t \mathbf{1}_{b_\ell}}{t} = \mathbf{1}_a g_\ell(\xi) \geq 0 \qquad \text{in } L^2.$$

Hence ξ lies almost surely in the polygon defined by the five probe inequalities. Isometry fixes its conditional mean to zero and its conditional covariance to $\mathrm{diag}(17/35, 18/35)$, so $\mathbb{E}[|\xi|^2 \mid a] = 1$. The polygon is contained in the unit disk and meets its boundary only at the five specified vertices. Saturation therefore forces ξ to take those vertex values almost surely; the moments fix their positive conditional weights to $(5, 8, 5, 5, 5)/28$. Off the core label, $Ux = Uy = 0$, and each probe is distinguished by its label indicator. Consequently every function in $U(W)$ is constant on the resulting ten-class measurable partition. Lemma D.2 supplies an exact finite detailed-balance CTMC realization, contradicting the rational symmetric-conductance obstruction of the construction. □

Remark D.4. The construction has the explicit order-preserving factor $Bf(s) = \mathbb{E}[f(Y_{[0,\infty)}) \mid X_0 = s]$ into $L^2(\pi_0)$, and $\mathsf{K} = B^*B$. Its obstruction is therefore distinct from the nonfactorability mechanism used by Arveson. The corollary does not by itself imply a positive entropy-production infimum; the separate closure argument in Appendix E establishes that stronger conclusion for this law.

E An explicit finite-observation separation certificate

We prove Theorem 3.5 for the unchanged generator (54). A rational polynomial in its known generator defines a finite-lag prediction action. Mismatch with that action is measured by observed residuals; it is never assumed away. No competing generator is differentiated, reconstructed by a logarithm or subject to a rate bound.

E.1 Rational features and a finite-lag observable action

Fix

$$\begin{gathered} h = 10^{-3}, \qquad t = 10^{-7}, \qquad \rho = 0, \qquad N = 12, \\ \beta_t = \frac{25}{2352t} + \frac{25}{32}, \qquad a_t = \frac{275}{9408t} + \frac{225}{256}, \\ C_\delta = 63843, \qquad C_G = \left(\frac{23}{h} + 36\right) a_t + \left(12 + \frac{12}{h}\right)\beta_t, \qquad C_R = \frac{491}{48h} + \frac{12960}{h^2}. \end{gathered} \tag{64}$$

For a future observed record, set $e = 1_{\{Y_0 = a\}}$, $e_\ell = 1_{\{Y_0 = b_\ell\}}$ and

$$\begin{gathered} F_\ell = \frac{1_{\{Y_0 = a, Y_t = b_\ell\}}}{t}, \qquad F_x = \frac{F_2 - F_1}{4}, \qquad F_y = e - \frac{F_1 + F_2}{4}, \\ F_{r,\ell} = F_\ell - b_\ell e - m_{\ell,x} F_x - m_{\ell,y} F_y \quad (\ell = 3, 4, 5). \end{gathered} \tag{65}$$

Here b_ℓ, m_ℓ are the fixed facet coefficients in (53), not random probe indicators.

Let $\phi = (e, x_0, y_0, e_1, \ldots, e_5)$ be the ideal target basis, where x_0, y_0 take the original vertex coordinates on core states and vanish on probes. Define C by $Q_0\phi_j = \sum_i C_{ij}\phi_i$. Equation (55) specifies it explicitly:

$$C = \begin{pmatrix} -\lambda & 0 & 0 & b^{\mathsf{T}} \\ 0 & -(\lambda+\alpha) & 0 & m_x^{\mathsf{T}} \\ 0 & 0 & -(\lambda+\beta) & m_y^{\mathsf{T}} \\ 5b & 5\Sigma_{xx} m_x & 5\Sigma_{yy} m_y & -5\,\mathrm{diag}(b) \end{pmatrix}, \qquad A_h = \sum_{n=0}^{N} \frac{(hC)^n}{n!}. \tag{66}$$

Here $b = (b_\ell)_{\ell=1}^5$, $m_x = (m_{\ell,x})_{\ell=1}^5$ and $m_y = (m_{\ell,y})_{\ell=1}^5$ are column vectors. The first three basis indices are e, x, y, followed by the probes. For the corresponding observed basis $\mathbf{F} = (e, F_x, F_y, e_1, \dots, e_5)$, define

$$F_{Aj} = \sum_i (A_h)_{ij} F_i, \qquad F_{Rj} = \mathrm{shift}_h F_j - F_{Aj} \quad (j = e, x, y), \qquad F_e = e. \tag{67}$$

Here the shift reads the same functional starting at time h. These are finite-action residuals, not generator residuals. On a core coordinate $\xi = (x, y)$ let $A_j(\xi) = (A_h)_{ej} + (A_h)_{xj} x + (A_h)_{yj} y$ and put

$$H_h(\xi) = (A_x(\xi), A_y(\xi)) - \xi A_e(\xi) = c_h + B_h \xi - \xi(k_h \cdot \xi). \tag{68}$$

Thus B_h is the core coordinate coefficient block minus $(A_h)_{ee} I$ and $k_h = ((A_h)_{xe}, (A_h)_{ye})$. For $w_i = y_i$ set $z_i = w_i \cdot H_h(p_i)$. The rational polynomial $q_h = d_0 + d_1 x + d_2 y + d_3 x^2 + d_4 xy$ is specified by

$$\begin{aligned} d_0 &= \frac{25}{72}(z_3 + z_4) + \frac{5}{9} z_1 - \frac{1}{8}(z_0 + z_2), \qquad d_1 = \frac{z_0 - z_2}{2}, \\ d_2 &= z_1 - d_0, \qquad d_3 = \frac{z_0 + z_2}{2} - d_0, \qquad d_4 = \frac{5}{4} d_1 + \frac{25}{24}(z_3 - z_4). \end{aligned} \tag{69}$$

This is the inverse vertex-evaluation rule, so $q_h(p_i) = z_i$. Its constant term is core restricted. With reflection θ read outward from the midpoint, define

$$\begin{aligned} G_3 &= F_{r,3} + \frac{147}{5} te, \qquad G_5 = F_{r,5} + \frac{147}{5} te, \\ G_4 &= F_{r,4} + \frac{842}{25} te - 2t F_y, \qquad G_{\mathrm{obs}} = \sum_{\ell=3}^5 (\theta G_\ell) G_\ell. \end{aligned} \tag{70}$$

$$\begin{aligned} q_{\mathrm{obs}} &= d_0 e + d_1 F_x + d_2 F_y + d_3 (\theta F_x) F_x + d_4 (\theta F_x) F_y, \\ \delta_{\mathrm{obs}} &= e - (\theta F_x) F_x - (\theta F_y) F_y, \qquad R_{\mathrm{obs}} = \sum_{j=e,x,y} (\theta F_{Rj}) F_{Rj}, \\ L_{\mathrm{obs}} &= \frac{q_{\mathrm{obs}}}{h} - C_\delta \delta_{\mathrm{obs}} - C_G G_{\mathrm{obs}} - C_R R_{\mathrm{obs}}. \end{aligned} \tag{71}$$

Only times $-h - t, -h, -t, 0, t, h, h + t$ are used. Translation by $h + t$ gives the seven-time tuple in Theorem 3.5. Every coefficient is rational.

E.2 The universal equilibrium inequality

Here $\mathbb{E}_{\mathrm{core}}$ denotes an unnormalized core expectation. Take any stationary finite detailed-balance competitor, with no rate, dimension or topology restriction. Write $\xi = (x, y)$ for its conditional coordinate features at the hidden midpoint, zero off the core, and f_ℓ, r_ℓ for its conditional probe and facet residual features. Let T_h be its genuine semigroup and $\bar{\phi} = \mathbb{E}[\mathbf{F} \mid X_0]$. Retain the actual residuals

$$R_j = T_h \bar{\phi}_j - \sum_i (A_h)_{ij} \bar{\phi}_i \quad (j = e, x, y). \tag{72}$$

As before, $g_1(\xi) = f_1 \geq 0$, $g_2(\xi) = f_2 \geq 0$, and $g_\ell(\xi) = f_\ell - r_\ell \geq -r_\ell$ for $\ell = 3, 4, 5$. Set $\eta_f = \max(0, r_3, r_4, r_5)$ and $a = 1 + 5\eta_f/8$. Every facet intercept is at least $8/5$, so $z = \xi/a \in K$. Choose a nearest vertex $v(\xi)$ with a fixed tie rule, and set $v = 0$ off the core. Since all vertices have unit norm, this maximizes $z \cdot p_i$. A barycentric representation of z gives $\max_i z \cdot p_i \geq |z|^2$. Consequently

$$|\xi - v|^2 \leq 1 - |\xi|^2 + \frac{5}{4}\eta_f + \frac{25}{32}\eta_f^2. \tag{73}$$

Indeed the left side is at most $1 + a^2|z|^2 - 2a|z|^2 = 1 - |\xi|^2 + 2a(a-1)|z|^2$, and $|z| \leq 1$.

A first-moment estimate is sharper here. For any distinct unit vertices with minimum distance d_* and any z in their convex hull, take independent barycentric vertices P_1, P_2 with mean z. Then

$$1 - |z|^2 = \tfrac{1}{2}\mathbb{E}|P_1 - P_2|^2 \geq \tfrac{d_*}{2}\mathbb{E}|P_1 - P_2| \geq \tfrac{d_*}{2}\mathbb{E}|P_1 - z| \geq \tfrac{d_*}{2}\,\mathrm{dist}(z, V).$$

The middle inequality is conditional Jensen; the coefficient $2/d_*$ is attained by two vertices at their midpoint. Here $d_*^2 = 4/5$, so choose $\kappa = 9/4 \geq \sqrt{5}$ before rescaling z. It follows that

$$|\xi - v| \leq \kappa(1 - |\xi|^2) + \frac{55}{16}\eta_f + \frac{225}{256}\eta_f^2.$$

Indeed $|\xi - v| \leq (a-1)|z| + \kappa(1-|z|^2)$ and $1 - |z|^2 = 1 - |\xi|^2 + (a^2-1)|z|^2$. Only the combined right side is required to be nonnegative.

Let $\gamma_\ell = \mathbb{E}[G_\ell \mid X_0]$. On the core these are $r_\ell + c_\ell$, where $c_3 = c_5 = 147t/5$ and $c_4 = t(842/25 - 2y)$. Nonnegativity of f_1, f_2 gives $y = 1 - (f_1 + f_2)/4 \leq 1$, so $c_4 \geq 792t/25 > c_{\min} := 147t/5$. For every real r and $c \geq c_{\min} > 0$,

$$r_+ \leq \frac{(r+c)^2}{4c_{\min}}, \qquad r_+^2 \leq (r+c)^2.$$

These inequalities follow from $(r-c)^2 \geq 0$ when $r > 0$; both left sides vanish when $r \leq 0$. Thus, with

$$\delta = \mathbb{E}e - \mathbb{E}|\xi|^2, \quad A_G = \mathbb{E}\sum_{\ell=3}^{5}\gamma_\ell^2, \quad D^2 = \mathbb{E}|\xi - v|^2, \quad U = \mathbb{E}|\xi - v|, \quad \mathcal{R} = \mathbb{E}\sum_{j=e,x,y} R_j^2, \tag{74}$$

averaging (73) gives

$$D^2 \leq S := \delta + \beta_t A_G, \qquad U \leq \frac{9}{4}\delta + a_t A_G. \tag{75}$$

Here $S \geq 0$ even when $\delta < 0$. Detailed balance and conditional independence identify δ, A_G and $\mathcal{R}$, and hence S, with their reflected observed expressions; D^2 remains a realization-dependent rounding error. The R_j remain the competitor's actual finite-action residuals in (72).

We record two stability estimates. Strong monotonicity and its regularization are classical [47, author manuscript, pp. 1–2]. The certificate below instead uses the first moment and only weak monotonicity; no global field extension is required.

Lemma E.1 (Strongly monotone dual-field stability)**.** *Let $V = \{p_1, \dots, p_N\} \subset \mathbb{R}^d$ consist of distinct vertices, $N \geq 2$. Let $w_1, \dots, w_N \in \mathbb{R}^d$ satisfy*

$$A_{ij} := (w_i - w_j)\cdot(p_i - p_j) > 0 \qquad (i \neq j), \qquad K := \max_{i<j}\frac{|w_i - w_j|^2}{4A_{ij}}.$$

Suppose (e_0, ξ_0) and (e_h, ξ_h) have an exchange-symmetric joint law, $e \in \{0,1\}$, and $\xi = 0$ on $e = 0$. On $e = 1$, choose $v(\xi) \in V$ using one fixed measurable rule; set $v = 0$ otherwise. Write $u = \xi - v$ and $D^2 = \mathbb{E}|u|^2 < \infty$. The field is tagged by the sector: $w(e,v) = w_i$ if $e = 1, v = p_i$, and zero if $e = 0$. Then

$$\mathbb{E}[w(e_0, v_0)\cdot(\xi_h - \xi_0 e_h)] \leq 2KD^2. \tag{76}$$

If the pair consists of functions of a stationary finite reversible continuous-time Markov chain at times $0, h$, the stronger bound is

$$\mathbb{E}[w(e_0, v_0)\cdot(\xi_h - \xi_0 e_h)] \leq KD^2. \tag{77}$$

Proof. The integrand vanishes outside the core-core sector $e_0 = e_h = 1$. Exchange symmetry writes its expectation as

$$\frac{1}{2}\mathbb{E}_{\text{core,core}}[(w(v_0) - w(v_h))\cdot(\xi_h - \xi_0)].$$

For a distinct rounded pair p_i, p_j, its integrand is at most

$$-\frac{A_{ij}}{2} + \frac{|w_i - w_j|}{2}|u_h - u_0| \leq \frac{|w_i - w_j|^2}{8A_{ij}}|u_h - u_0|^2 \leq \frac{K}{2}|u_h - u_0|^2.$$

The equal-vertex term is zero. Identical marginals give $\mathbb{E}|u_h - u_0|^2 \leq 4D^2$, proving (76). For a reversible continuous-time semigroup, $T_h = T_{h/2}^2$ with $T_{h/2}$ self-adjoint, so T_h is positive semidefinite on $L^2(\pi)$ [48, Lemma 12.2 and Section 20.1]. Thus $\langle f, T_h f\rangle_\pi \geq 0$. Hence

$$\mathbb{E}|u_h - u_0|^2 = 2D^2 - 2\sum_{j=1}^{d}\langle u_j, T_h u_j\rangle_\pi \leq 2D^2.$$

Restricting this nonnegative squared increment to the core-core sector only decreases its expectation. This proves (77), including reducible models. □

For the same exchange-symmetric pair, nonnegative A_{ij} suffice for the first-moment bound

$$W_h := \mathbb{E}[w(e_0, v_0) \cdot (\xi_h - \xi_0 e_h)] \le M\mathbb{E}|u|, \qquad M = \max_{i,j} |w_i - w_j|. \tag{78}$$

Indeed symmetrization gives the integrand $-A_{ij}/2 - (w_i - w_j) \cdot (u_0 - u_h)/2$. Discard its nonpositive first term and use $\mathbb{E}|u_h - u_0| \le 2\mathbb{E}|u|$. This estimate requires neither strict coercivity nor continuous-time semigroup positivity.

Corollary E.2 (An observed finite-lag transition penalty)**.** *In the continuous-time setting of Lemma E.1, suppose $|p_i| = 1$ and $A_{ij} \ge \rho|p_i - p_j|^2$ for some $\rho > 0$. Let $S \ge D^2$ and set*

$$\Psi_h = \langle \xi, (I - T_h)\xi \rangle_\pi, \qquad p_h = \mathbb{P}(e_0 = 1, e_h = 0), \qquad \Lambda_h(S) = [[\sqrt{\Psi_h} - \sqrt{S}]_+^2 - p_h]_+.$$

For $0 \le \vartheta < 1$, the work in (77) satisfies

$$\mathbb{E}[w(e_0, v_0) \cdot (\xi_h - \xi_0 e_h)] \le \frac{KS}{1 - \vartheta} - \vartheta\rho\Lambda_h(S). \tag{79}$$

For every $b > 0$ it also satisfies the affine bound

$$\mathbb{E}[w(e_0, v_0) \cdot (\xi_h - \xi_0 e_h)] \le -\vartheta\rho \left(\frac{\Psi_h}{1 + b} - p_h \right) + \left(\frac{K}{1 - \vartheta} + \frac{\vartheta\rho}{b} \right) S. \tag{80}$$

For conditional features of the observed probes used here, Ψ_h, p_h and the certified S are determined by the existing seven-time law.

Proof. Retaining a fraction ϑ of the negative term in the completed square of Lemma E.1 gives

$$W_h \le -\frac{\vartheta}{2} G_h + \frac{KD^2}{1 - \vartheta}, \qquad G_h = \mathbb{E}_{\mathrm{core},\mathrm{core}} A_{v_0, v_h}.$$

Put $E_v = \langle v, (I - T_h)v \rangle_\pi$. The unit vertex norms and stationarity give

$$2(E_v - p_h) = \mathbb{E}_{\mathrm{core},\mathrm{core}} |v_h - v_0|^2 \ge 0, \qquad G_h \ge 2\rho(E_v - p_h).$$

The seminorm induced by $I - T_h$ obeys the triangle inequality. Since $0 \le I - T_h \le I$, the error $u = \xi - v$ has energy at most $D^2 \le S$. Thus $\sqrt{E_v} \ge [\sqrt{\Psi_h} - \sqrt{S}]_+$, proving (79). Alternatively, Young's inequality gives $\Psi_h \le (1 + b)E_v + (1 + 1/b)S$ and proves (80). Finally, reflection and midpoint conditioning identify

$$\Psi_h = \sum_{j=x,y} \mathbb{E}[(\theta F_j)F_j - (\theta F_j)(\mathrm{shift}_h F_j)], \qquad p_h = \mathbb{P}(Y_0 = a, Y_h \ne a).$$

These expressions use a subset of the existing observation times. The work W_h remains an internal rounded-feature term; a strengthened bounded score still requires the interpolation and residual estimates. □

For the fixed weak field $w_i = y_i$, all $A_{ij} \ge 0$ and $M^2 \le 490 < 23^2$. Write $H_h(\xi) = c_h + B_h\xi - \xi(k_h \cdot \xi)$ and $M_q = \left(\begin{smallmatrix} d_3 & d_4/2 \\ d_4/2 & 0 \end{smallmatrix} \right)$. For $\xi = p_i + u$ define

$$\ell_i = \nabla q_h(p_i) - B_h^{\mathsf{T}} w_i + (w_i \cdot p_i)k_h + (k_h \cdot p_i)w_i, \qquad M_i = M_q + \frac{w_i k_h^{\mathsf{T}} + k_h w_i^{\mathsf{T}}}{2}.$$

Interpolation gives the exact cancellation

$$q_h(\xi) - w_i \cdot H_h(\xi) = \ell_i \cdot u + u^{\mathsf{T}} M_i u.$$

The displayed rational action and interpolant yield

$$\begin{gathered} \max_i |\ell_i|^2 < (36h)^2, \qquad \max_i \lambda_{\max}(M_i) < 12h, \\ \max_i |w_i|^2 < 491, \qquad \max_i \big(|w_i|^2 + (w_i \cdot p_i)^2 \big) < 540, \\ \sum_i |(A_h)_{ij}| \le 2 \ (j = e, x, y), \qquad \sum_{r=0}^{4} |d_r| \le 60h. \end{gathered} \tag{81}$$

The eigenvalue bound follows from rational Gershgorin upper bounds. Expanding the residual pairing at p_i bounds it by $\sqrt{540}\sqrt{\mathcal{R}} + \sqrt{491}D\sqrt{\mathcal{R}}$. Use $\sqrt{491}D\sqrt{\mathcal{R}} \le 12D^2 + (491/48)\mathcal{R}$. The residual identity is

$$T_h\xi - \xi T_h e = H_h(\xi) + R_{(x,y)} - \xi R_e \quad \text{on the core,} \tag{82}$$

Combining it with (78) gives

$$\mathbb{E}_{\text{core}}\frac{q_h(\xi)}{h} \le \left(\frac{23}{h} + 36\right)U + \left(12 + \frac{12}{h}\right)D^2 + \frac{\sqrt{540}}{h}\sqrt{\mathcal{R}} + \frac{491}{48h}\mathcal{R}. \tag{83}$$

Substitute both bounds in (75) with their separate δ and A_G coefficients; δ need not be nonnegative. The remaining completed square is

$$\frac{\sqrt{540}}{h}\sqrt{\mathcal{R}} \le \frac{1}{96} + \frac{12960}{h^2}\mathcal{R}.$$

The coefficients in (64) prove

$$\mathbb{E}_{P'}L_{\text{obs}} \le \frac{1}{96} \qquad (P' \in \mathcal{E}_{\text{fin}}). \tag{84}$$

There is no square-root rounding budget or inverse-coercivity factor.

E.3 Fixed-target evaluation and bounded normalization

The target generator satisfies $\|Q_0\|_\infty < 26$. For $s = t, h$ put $P_s = \sum_{n=0}^{12}(sQ_0)^n/n!$ and

$$u_s = \frac{2(26s)^{13}}{13!}, \qquad \|e^{sQ_0} - P_s\|_\infty \le u_s. \tag{85}$$

The exponential-series bound uses $26s < 1/2$ and concerns only the fixed target. Build the polynomial conditional coordinates $x_i^{\text{p}}, y_i^{\text{p}}$ from P_t by the same probe formula as the observed features. Their coordinate errors are at most $e_t = u_t/(2t)$ and their magnitudes are below two. Let $\delta^{\text{p}}, q^{\text{p}}, A_G^{\text{p}}$ denote the corresponding rational stationary sums. Then

$$\begin{aligned}
\delta &\le \delta^{\text{p}} + \sum_{i:a}\pi_i\{2e_t(|x_i^{\text{p}}| + |y_i^{\text{p}}|) + 2e_t^2\},\\
\mathbb{E}_P(q_{\text{obs}}/h) &\ge q^{\text{p}}/h - \frac{1}{h}\sum_{i:a}\pi_i\big\{(|d_1| + |d_2|)e_t + |d_3|(2|x_i^{\text{p}}|e_t + e_t^2)\\
&\qquad + |d_4|[(|x_i^{\text{p}}| + |y_i^{\text{p}}|)e_t + e_t^2]\big\}.
\end{aligned}$$

The errors in the shifted conditional facets are at most $(5e_t, (4+2t)e_t, 5e_t)$. If $\bar\phi^{\text{p}}$ is the polynomial conditional basis, set $R^{\text{p}} = P_h\bar\phi^{\text{p}} - \bar\phi^{\text{p}}A_h$ in its first three columns. Sup-norm contraction and (81) give $|R_j - R_j^{\text{p}}| \le 3e_t + 2u_h$. Bounding each squared component by its rational interval endpoint, the fixed parameters give

$$\begin{aligned}
\mathbb{E}_P(q_{\text{obs}}/h) > 0.1134506396, \qquad \delta < 1.137856\times10^{-6},\\
A_G < 5.51524\times10^{-24}, \qquad \mathcal{R} < 8.21447\times10^{-17}.
\end{aligned} \tag{86}$$

All four bounds are outward-rounded rational endpoints. The ancillary file `anc/check_certificate.py` reproduces them independently: it encloses e^{sQ_0}, $s \in \{t, h\}$, by its exact degree-16 Taylor polynomial plus an entrywise remainder interval, propagates outward-rounded rational intervals, and evaluates the four moments both through (57) and directly with the stationary reverse chain of Q_0 for the reflected half. It also recomputes (81), the constants in (64), and L_{obs} on all 6^7 words, confirming $|L_{\text{obs}}| \le B$. The target restricted-adjoint Gram identity (57) identifies these conditional quadratic moments with their reflected observable expectations; microscopic symmetry of Q_0 is not assumed. Substitution yields

$$\mathbb{E}_P L_{\text{obs}} > \frac{1}{25}. \tag{87}$$

This target calculation is an interval enclosure of the complete finite-lag action, rather than extrapolation of small-t values.

The normalization holds on every word over the six labels. The first three observed features have possible values

$$V_t = \left\{(0,0,0), (1,0,1), \left(1, -\frac{1}{4t}, 1 - \frac{1}{4t}\right), \left(1, \frac{1}{4t}, 1 - \frac{1}{4t}\right)\right\}.$$

Every entry of the complete observed basis has magnitude at most $1/(4t)$, and the diameter of V_t is at most $1/(2t)$. Since $\|C\|_1 = 288/5$, put $r = h\|C\|_1 = 36/625$; then

$$\|A_h - I\|_1 < \frac{r}{1-r} = \frac{36}{589} < \frac{1}{10}.$$

Write $F_R = (\mathrm{shift}_h \mathbf{F}_{e,x,y} - \mathbf{F}_{e,x,y}) - ((A_h - I)^{\mathsf{T}}\mathbf{F})_{e,x,y}$. The first difference has norm at most $1/(2t)$, and each entry of the second is at most $\|A_h - I\|_1/(4t) < 1/(40t)$. Thus $\|F_R\| \le 11/(20t)$ and, by Cauchy–Schwarz, $|R_{\mathrm{obs}}| \le 121/(400t^2)$. The rescaled unshifted facet residual vector is $z_j - tc$, where $c = (3, 18/5, 3)$ and

$$z_j \in \{(0,0,0), (3/4, 1/2, -1/4), (-1/4, 1/2, 3/4), (1,0,0), (0,1,0), (0,0,1)\}.$$

Hence $\|F_r\| \le (1+6t)/t$. The shifts contribute at most $93t + 1/2$ in Euclidean norm, using $|F_y| \le 1/(4t)$. Since $t \le 10^{-6}$, this gives $\|G\| \le (1+7t)/t$ and $|G_{\mathrm{obs}}| \le (1+7t)^2/t^2$. Also $|q_{\mathrm{obs}}/h| \le 60/t^2$ and $|\delta_{\mathrm{obs}}| \le 3/t^2$. These are pointwise word bounds, independent of any generator. Define the rational constants

$$B = \frac{60 + 3C_\delta + (1+7t)^2 C_G + \frac{121}{400} C_R}{t^2}, \qquad W = \frac{L_{\mathrm{obs}} + B}{2B},$$
$$c = \frac{1}{2} + \frac{1}{192B}, \qquad \eta = \frac{71}{4800B}. \tag{88}$$

They satisfy $0 \le W \le 1$. The raw gap $1/25 - 1/96 = 71/2400$ proves the claimed separation, with $\eta \simeq 1.2398 \times 10^{-26}$. These constants still do not establish practical statistical power.

E.4 Cylinder separation and entropy production

Let m_0 index the displayed seven-time tuple, with the current value of t, in the enumeration of (4). Changing the probe lag also changes this tuple; no comparison of its enumeration weight with that of another tuple is required. A $[0,1]$ score gap is bounded by supremum-event total variation, so

$$\mathrm{TV}(P_{\mathcal{T}}, P'_{\mathcal{T}}) \ge \eta, \qquad D(P, P') \ge 2^{-m_0}\eta \quad (P' \in \mathcal{E}_{\mathrm{fin}}). \tag{89}$$

The metric concerns rational coordinate cylinders, rather than uniform path-window total variation.

E.5 Reversibilization and the explicit dissipation floor

For any exact compatible finite irreducible reciprocal realization Q, π, h, let $Q^\dagger$ be its stationary reverse and set $\overline{Q} = (Q + Q^\dagger)/2$. Stationarity gives $q^\dagger_{ii} = q_{ii}$, so the two generators have the same escape rate at every state. The comparison $\overline{Q}$ is irreducible and detailed balanced at π, with unchanged labels and clock. For a supported unordered pair write $a = \pi_i q_{ij} > 0$, $b = \pi_j q_{ji} > 0$. The stationary path likelihoods have equal initial and holding-time factors, giving the exact finite-interval identity

$$D_{\mathrm{KL}}(\mathbb{P}_{Q,[0,T]} \| \mathbb{P}_{\overline{Q},[0,T]}) = T \sum_{i<j} \left[a \log \frac{2a}{a+b} + b \log \frac{2b}{a+b} \right]. \tag{90}$$

This is the usual finite-state path-relative-entropy calculation [21], with the holding terms cancelling. For $s = a + b$ and $z = (a-b)/s$, the scalar bracket is $\frac{s}{2}[(1+z)\log(1+z) + (1-z)\log(1-z)]$. It is at most $(a-b)\log(a/b)/4$. Indeed the difference after multiplying by $4/s$ is an even function vanishing at zero, with derivative for $0 \le z < 1$

$$\frac{2z}{1-z^2} - \log \frac{1+z}{1-z} \ge 0, \tag{91}$$

because the logarithm is $2\int_0^z (1-u^2)^{-1} du$. Consequently

$$D_{\mathrm{KL}}(\mathbb{P}_{Q,[0,T]} \| \mathbb{P}_{\overline{Q},[0,T]}) \le T\sigma(Q)/4. \tag{92}$$

Pinsker and data processing give, for any observed time tuple of span T,

$$\mathrm{TV}(P_{\mathcal{T}}, P_{\overline{Q},\mathcal{T}}) \le \min\bigl(1, \sqrt{T\sigma(Q)/8}\bigr).$$

The seven-time statistic has span $T = 2(h+t)$, so its separation from the equilibrium comparison law gives $\eta \le \sqrt{T\sigma(Q)/8}$ and hence $\sigma(Q) \ge 8\eta^2/T$. This is approximately 6.1475×10^{-49} in the fixed inverse-time unit. The compatible class is nonempty because it contains Q_0. The positive floor follows from the explicit finite-word separator, rather than from exact equilibrium nonexistence alone.

E.6 Transfer to smaller lags from observed prediction residuals

The full finite prediction action also controls how far conditional probe coordinates lie outside the polygon. The comparison below uses its fixed target metric, without requiring the competing realization to have the same Gram matrix or a bounded generator.

Lemma E.3 (Fractional transfer of a finite prediction action)**.** *Let T be a positive self-adjoint contraction on a real Hilbert space $\mathcal{H}$, and let A be a symmetric $m \times m$ matrix with spectrum in $[a_*, 1]$, where $a_* > 0$. For every linear map $U : \mathbb{R}^m \to \mathcal{H}$ and $0 < \alpha \le 1$,*

$$\|T^\alpha U - UA^\alpha\|_{\mathrm{HS}} \le a_*^{\alpha-1} \|TU - UA\|_{\mathrm{HS}}. \tag{93}$$

The coefficient is sharp. Neither invariance of the range of U nor an isometry assumption on U is required.

Proof. Choose an orthonormal eigenbasis v_j of A, with eigenvalues a_j. For $a \in [a_*, 1]$, define

$$g_a(x) = \frac{x^\alpha - a^\alpha}{x - a} \quad (x \ne a), \qquad g_a(a) = \alpha a^{\alpha-1}.$$

Concavity gives $0 \le g_a(x) \le a^{\alpha-1}$ for $0 \le x \le 1$: the largest secant slope at fixed a is its value at $x = 0$. Spectral calculus therefore gives

$$(T^\alpha - a_j^\alpha I)Uv_j = g_{a_j}(T)(T - a_j I)Uv_j, \qquad \|g_{a_j}(T)\| \le a_*^{\alpha-1}.$$

Summing the squared column bounds proves (93), including $\alpha = 1$. Taking $T = 0$ and $A = a_* I$ gives equality. The same coefficient is approached on the centered coordinate of a reversible two-state chain with equal stationary masses and jump rate r, whose time-h eigenvalue is e^{-2rh}. Thus unrestricted competing rates prevent a uniform additional factor proportional to α as $\alpha \downarrow 0$. This sharpness example concerns the operator inequality, rather than compatibility with the fixed target law. □

To apply the lemma, retain $h = 10^{-3}$ and the known C, A_h from (66). Its reference Gram matrix and symmetric conjugates are

$$G = \mathrm{diag}(1/2, 17/70, 18/70, 1/10, \dots, 1/10), \qquad B_0 = G^{1/2} C G^{-1/2}, \qquad A = G^{1/2} A_h G^{-1/2}.$$

The restricted-adjoint identity makes B_0 symmetric. Stationarity of the target gives $\langle f, Q_0 f\rangle_\pi \le 0$, so its spectrum lies in $[-26, 0]$ by $\|Q_0\|_\infty < 26$. Consequently A has spectrum in $[a_*, 1]$, where $a_* = 487/500$: for $0 \le x \le 26h$, the alternating polynomial $P_{12}(-x)$ lies between $1 - x$ and 1. Its remainder gives

$$\|A - e^{hB_0}\| \le u_h, \qquad u_h = \frac{2(26h)^{13}}{13!}.$$

These are bounds on the known target action alone.

Take any finite stationary detailed-balance competitor, allowing arbitrary hidden state count, rates, topology and reducibility. Let $\mathbf{F} = (e, F_x, F_y, e_1, \dots, e_5)$ be bounded observed functionals of the times $0, t$, where e, e_ℓ are the actual label indicators at time zero and F_x, F_y vanish off the core label. The raw features in (65), or specified bounded affine corrections with the same support, are permitted. Define $\bar\phi_j = \mathbb{E}[F_j \mid X_0]$, and let $\Phi : \mathbb{R}^8 \to L^2(\pi)$ have columns $\bar\phi_j$. Set

$$\begin{aligned} U = \Phi G^{-1/2}, \qquad R_j = T_h \bar\phi_j - \sum_i (A_h)_{ij} \bar\phi_i \quad (j = e, x, y, e_1, \dots, e_5), \\ \mathcal{R}_G = \sum_j \frac{\|R_j\|_\pi^2}{G_{jj}}, \qquad \mathcal{M}_G = \sum_j \frac{\|\bar\phi_j\|_\pi^2}{G_{jj}}. \end{aligned} \tag{94}$$

The norms are thus $\|T_h U - UA\|_{\mathrm{HS}}^2 = \mathcal{R}_G$ and $\|U\|_{\mathrm{HS}}^2 = \mathcal{M}_G$. All eight columns are used; the competitor's actual Gram matrix need not equal G.

Proposition E.4 (Observed control of every facet violation)**.** *For every $0 < s \le h$, put*

$$Z_s = \frac{1}{\sqrt{10}} \left[a_*^{s/h-1} \frac{\sqrt{\mathcal{R}_G}}{s} + \left(a_*^{s/h-1} \frac{u_h}{h} + 338 s \right) \sqrt{\mathcal{M}_G} \right]. \tag{95}$$

Write $\xi = (\bar{\phi}_x, \bar{\phi}_y)$ and $g_\ell(\xi) = b_\ell + m_\ell \cdot \xi$ on the core. Set $\eta_f = \max(0, -g_1(\xi), \dots, -g_5(\xi))$ on the core and zero elsewhere. Then

$$\mathbb{E}\eta_f^2 \le Z_s^2, \qquad \mathbb{E}\eta_f \le Z_s. \tag{96}$$

For the nearest vertex $v(\xi)$ with the earlier tie rule, zero off the core, and the signed defect $\delta = \mathbb{E}e - \mathbb{E}|\xi|^2$, it follows that

$$\begin{aligned} \mathbb{E}|\xi - v| &\le \frac{9}{4}\delta + \frac{55}{16}Z_s + \frac{225}{256}Z_s^2, \\ \mathbb{E}|\xi - v|^2 &\le \delta + \frac{5}{4}Z_s + \frac{25}{32}Z_s^2. \end{aligned} \tag{97}$$

The inputs $\mathcal{R}_G, \mathcal{M}_G, \delta$ are determined by the same seven-time observation law. The parameter s is an analytical lag and introduces no further observation time.

Proof. For the reversible continuous-time semigroup, $T_s = T_h^{s/h}$ and T_h is a positive self-adjoint contraction. Apply Lemma E.3 with $\alpha = s/h$. The commuting matrices A and e^{hB_0} have eigenvalues at least a_*. The scalar derivative $\alpha x^{\alpha-1}$ therefore bounds their power difference by $\alpha a_*^{\alpha-1} u_h$. Also $\|e^{sB_0} - I - sB_0\| \le 338 s^2$ by the integral Taylor remainder and $B_0 \le 0$. Hence

$$\|T_s U - U(I + sB_0)\|_{\mathrm{HS}} \le a_*^{s/h-1}\sqrt{\mathcal{R}_G} + \left(a_*^{s/h-1}\frac{u_h s}{h} + 338 s^2\right)\sqrt{\mathcal{M}_G}.$$

On the core, the ideal probe column of $\Phi(I + sC)$ is $s g_\ell(\xi)$, whereas the genuine transition probability $T_s e_\ell$ is nonnegative. Thus its error divided by s bounds $[-g_\ell(\xi)]_+$. The five probe entries of G are all $1/10$. Their unweighted squared column errors sum to at most one tenth of the displayed Hilbert–Schmidt square; this gives (96), with Cauchy–Schwarz for the first moment.

Every facet intercept is at least $8/5$, so $z = \xi/(1 + 5\eta_f/8)$ lies in the same polygon. Its vertices have unit norm and minimum distance squared $4/5$. The first-moment geometry used above, with the rational coefficient $9/4$ chosen before radial rescaling, and the barycentric squared-distance bound give pointwise on the core

$$\begin{aligned} |\xi - v| &\le \frac{9}{4}(1 - |\xi|^2) + \frac{55}{16}\eta_f + \frac{225}{256}\eta_f^2, \\ |\xi - v|^2 &\le 1 - |\xi|^2 + \frac{5}{4}\eta_f + \frac{25}{32}\eta_f^2. \end{aligned}$$

Averaging and using (96) proves (97). The complete right sides are nonnegative; δ itself need not be.

Finally, put $F_{Rj} = \mathrm{shift}_h F_j - \sum_i (A_h)_{ij} F_i$ for all eight columns. Detailed balance and midpoint conditional independence give

$$\mathcal{R}_G = \mathbb{E}\sum_j \frac{(\theta F_{Rj})F_{Rj}}{G_{jj}}, \qquad \mathcal{M}_G = \mathbb{E}\sum_j \frac{(\theta F_j)F_j}{G_{jj}}.$$

The analogous reflected coordinate products give δ. The future times $0, t, h, h+t$ and their reflections are exactly $-h-t, -h, -t, 0, t, h, h+t$. The energies $\mathcal{R}_G, \mathcal{M}_G$ are nonnegative on competitors, rather than wordwise nonnegative statistics; the defect δ may be negative. □

This estimate permits all five facets to be tested after a specified coordinate correction. It supplies no improved normalized separation margin by itself: its coefficients, target enclosure and arbitrary-word range must still be propagated through a bounded score. In particular, the preceding numerical certificate is retained.

F The affinity inequality and its exact scope

Let p, q be probability vectors with common support. Put $H = \sum_i \sqrt{p_i q_i}$ and use weights $r_i = (p_i + q_i)/2$. For $u_i = 2\sqrt{p_i q_i}/(p_i + q_i)$ define

$$h(u) = \sqrt{1 - u^2}\ \operatorname{artanh}\sqrt{1 - u^2}, \qquad 0 < u \le 1. \tag{98}$$

Then $\sum_i r_i u_i = H$ and $J(p, q) = 4\sum_i r_i h(u_i)$. For $\theta = \sqrt{1 - u^2}$,

$$h''(u) = \frac{\theta/(1 - \theta^2) - \operatorname{artanh}\theta}{\theta^3} > 0, \tag{99}$$

since the numerator vanishes at zero and has derivative $2\theta^2/(1-\theta^2)^2 > 0$. The endpoint at $u = 1$ is handled by continuity. Jensen gives $J \geq 4h(H) = 4\mathcal{F}(1-H^2)$. The positive power series in (10) proves monotonicity and convexity of $\mathcal{F}$. For $p = ((1+r)/2, (1-r)/2)$ and q with its entries exchanged, $1 - H^2 = r^2$ and equality holds. Thus this *probability-pair* inequality is sharp. This does not establish sharpness of the Markov-process inference bound, which includes further data-processing and cone relaxations. If supports differ, the divergence is infinite and the inequality follows by a limiting argument.

G Path reversal, support and the factor of two

Stationarity gives

$$\sum_{y \neq x} q^\dagger_{xy} = \frac{1}{\pi_x} \sum_{y \neq x} \pi_y q_{yx} = -q_{xx}. \tag{100}$$

Thus Q and $Q^\dagger$ have the same escape rates. Their finite-time conditional path laws P_x and $\mathsf{P}^\dagger_x$, both started at x, have the same support by reciprocity. The log Radon–Nikodym derivative contains only jump terms,

$$\log \frac{d\mathsf{P}_x}{d\mathsf{P}^\dagger_x} = \sum_{\text{jumps } u \to v} \log \frac{q_{uv}}{q^\dagger_{uv}}. \tag{101}$$

All log ratios on supported edges are finite; finite-state jump counts have finite expectation on bounded intervals. Averaging over $x \sim \pi$ therefore gives

$$\sum_x \pi_x D_{\mathrm{KL}}(\mathsf{P}_x \| \mathsf{P}^\dagger_x) = \tau \sum_{u \neq v} \pi_u q_{uv} \log \frac{\pi_u q_{uv}}{\pi_v q_{vu}} = \tau\sigma(Q), \tag{102}$$

$$\sum_x \pi_x D_{\mathrm{KL}}(\mathsf{P}^\dagger_x \| \mathsf{P}_x) = \tau\sigma(Q^\dagger) = \tau\sigma(Q). \tag{103}$$

The two outward half-paths of a stationary two-sided Markov process are conditionally independent given X_0; the past half has generator $Q^\dagger$. Applying the same observation/classification channel to the two conditional laws gives a_x, b_x and the final inequality in (16). An event has zero probability under one conditional law exactly when it does under the other, so $H_x > 0$. There are finitely many midpoint states, hence $\delta < 1$.

The total measured interval is 2τ, whereas each one-sided conditional divergence contributes $\tau\sigma$. This explains $2/\tau$ in (11); substituting the full window length for τ without adjusting the coefficient would change the claim.

Proposition G.1 (Strictness for finite dissipative chains)**.** *For every finite chain and finite positive window in Theorem 4.1, $\sigma(Q) > 0$ implies $\sigma(Q) > 2\mathcal{F}(\Delta_{\mathrm{cp}}(S))/\tau$. Equality holds exactly at equilibrium.*

Proof. The affinity argument in Appendix F applies to mutually absolutely continuous probability measures as well as finite vectors. For $R = (P + P')/2$ and $U = 2\sqrt{pp'}/(p + p')$, it gives $J(P, P') = 4\mathbb{E}_R h(U) \geq 4h(\mathbb{E}_R U)$, with equality only if U is constant. Each conditional path pair $\mathsf{P}_x, \mathsf{P}^\dagger_x$ has a common no-jump atom of equal positive mass $e^{q_{xx}\tau}$; there $U = 1$. Hence equality forces identical conditional path laws. If $\sigma > 0$, an edge has nonzero stationary affinity, and the event of exactly that jump gives a different likelihood ratio with positive probability for its initial midpoint. At least one conditional affinity inequality is therefore strict. A common classification channel increases affinity by Cauchy–Schwarz. Writing H^{path}_x for the path affinity yields

$$2\tau\sigma > 4 \sum_x \pi_x \mathcal{F}(1 - (H^{\mathrm{path}}_x)^2) \geq 4\mathcal{F}\left(\sum_x \pi_x (1 - (H^{\mathrm{path}}_x)^2) \right) \geq 4\mathcal{F}(\Delta_{\mathrm{cp}}(S)).$$

At equilibrium the identical-half Gram representation gives zero deficit. □

This is pointwise strictness. It supplies no uniform positive gap over varying generators and does not rule out asymptotic sharpness.

H Complete positivity: optimization, duality and contraction

For completeness, $\mathcal{CP}_m$ is closed. Normalize each nonzero generator uu^{T} by its trace. These normalized generators form a compact set, and their convex hull is compact in the finite-dimensional space of symmetric matrices. A convergent sequence of completely positive matrices has bounded traces, and thus has a limit in the same cone. Equivalently, finite-dimensional Carathéodory reduction supplies uniformly bounded numbers of normalized factors.

For fixed $S \geq 0$, the feasible set $\{G \in \mathcal{CP}_m : 0 \leq G \leq S\}$ is nonempty and compact. A maximizing G must have $G_{ii} = S_{ii}$, since a positive diagonal deficit can be added as a nonnegative coordinate Gram factor. Its total mass cannot exceed one. If the maximum is one, the nonnegative matrix $S - G$ has zero total mass and hence vanishes. This proves $\Delta_{\mathrm{cp}}(S) = 0 \iff S \in \mathcal{CP}_m$.

Weak duality follows from

$$-\langle W, S\rangle = -\langle W, G\rangle - \langle W, S - G\rangle \leq \mathbf{1}^{\mathsf{T}}(S - G)\mathbf{1}. \tag{104}$$

For the reverse inequality take $0 \leq t < \Delta_{\mathrm{cp}}(S)$ and consider the closed convex set

$$\mathcal{C}_t = \{G + E : G \in \mathcal{CP}_m,\ E = E^{\mathsf{T}} \geq 0, \mathbf{1}^{\mathsf{T}}E\mathbf{1} \leq t\}. \tag{105}$$

It is closed because it is the sum of a closed cone and a compact set. Since $S \notin \mathcal{C}_t$, strict separation gives a symmetric W with $\langle W, S\rangle < \inf_{C \in \mathcal{C}_t}\langle W, C\rangle$. The infimum is finite only if W is copositive. Its value is $t\min(0, \min_{ij} W_{ij})$. Since $S \geq 0$, separation is impossible if $W \geq 0$ entrywise. Normalize the negative minimum entry to -1; then $-\langle W, S\rangle > t$. Letting $t \uparrow \Delta_{\mathrm{cp}}(S)$ proves (17). If $\Delta_{\mathrm{cp}}(S) = 0$, $W = 0$ and weak duality suffice. No strict-positivity assumption on S is used.

Finally, if G is a maximizing minorant and $K \geq 0$ has row sums one, then $K^{\mathsf{T}}GK$ is completely positive, is dominated entrywise by $K^{\mathsf{T}}SK$, and has the same total mass as G. It is a feasible minorant for the coarsened matrix and proves (18).

Proposition H.1 (Sharp contraction under matched relabeling)**.** *Let $S = S^{\mathsf{T}} \geq 0$ have total mass one, and let $K \geq 0$ be row-stochastic. Define the row affinities and channel coefficient*

$$b_{ij}(K) = \sum_a \sqrt{K_{ia}K_{ja}}, \qquad c(K) = 1 - \min_{i,j} b_{ij}(K)^2.$$

Then

$$\Delta_{\mathrm{cp}}(K^{\mathsf{T}}SK) \leq c(K)\Delta_{\mathrm{cp}}(S). \tag{106}$$

For at least two input categories this is the exact best multiplier over all symmetric nonnegative probability matrices. In particular, pairwise overlap of all channel-row supports implies strict contraction whenever $\Delta_{\mathrm{cp}}(S) > 0$.

Proof. For probability vectors p, q, let $A = (pq^{\mathsf{T}} + qp^{\mathsf{T}})/2$ and $v_a = \sqrt{p_a q_a}$. AM–GM gives $vv^{\mathsf{T}} \leq A$ entrywise. Every completely positive minorant $H \leq A$ is positive semidefinite, and hence

$$\mathbf{1}^{\mathsf{T}}H\mathbf{1} \leq \left(\sum_a \sqrt{H_{aa}}\right)^2 \leq \left(\sum_a \sqrt{p_a q_a}\right)^2.$$

The rank-one minorant attains this bound.

Now choose an optimal minorant $G \leq S$, and put $E = S - G$. Diagonal saturation gives $E_{ii} = 0$. For rows i, j of K, define $v_a^{ij} = \sqrt{K_{ia}K_{ja}}$. The completely positive matrix

$$B = \sum_{i<j} 2E_{ij}v^{ij}(v^{ij})^{\mathsf{T}}$$

is dominated entrywise by $K^{\mathsf{T}}EK$, by AM–GM. Thus $K^{\mathsf{T}}GK + B$ is a feasible output minorant, and

$$\Delta_{\mathrm{cp}}(K^{\mathsf{T}}SK) \leq \sum_{i<j} 2E_{ij}(1 - b_{ij}(K)^2) \leq c(K)\Delta_{\mathrm{cp}}(S).$$

Taking $S = (e_i e_j^{\mathsf{T}} + e_j e_i^{\mathsf{T}})/2$ for a pair of rows with minimum affinity proves sharpness by the pair calculation. □

The sharpness domain here is the class of all symmetric nonnegative probability matrices. The pure off-diagonal extremizers are not motif matrices of the matched finite reciprocal CTMC contract; sharpness within that smaller physical class is not asserted.

I The five-cycle witness

Let $u \geq 0$ and normalize $\sum_i u_i = 1$. The function $\sum_{\{i,j\} \in E(C_5)} u_i u_j$ attains a maximum on the simplex. If two nonadjacent coordinates are positive, move their combined mass to the one with the larger weighted neighbor sum. There is no product between these coordinates, so the objective does not decrease and the support shrinks. Repetition leaves a clique. A clique in C_5 has at most two vertices, on which the objective is at most $1/4$. This is the classical Motzkin–Straus merging argument [13]. Thus $u^{\mathsf{T}}(J_5 - 2A_5)u \geq 0$. Appending a zero row and column preserves copositivity. Its negative pairing with M is exactly (20); the score includes both orientations of every cycle edge.

J Self-contained benchmark realization and calibration

J.1 Twenty positive reciprocal channels

Use $d = 1$, $a = 3/5$, $b = 1/100$, $c = d/2 + b = 51/100$ and indices modulo five. Put

$$v_i = e_{i-1} + 2e_i + e_{i+1}, \qquad u_i = e_i + e_{i+1}, \qquad G_0 = b \sum_i v_i v_i^{\mathsf{T}} + \frac{d - 6b}{2} \sum_i u_i u_i^{\mathsf{T}}. \tag{107}$$

The diagonal, cycle-edge and nonedge entries of G_0 are respectively d, c, b. For every summand $\omega w w^{\mathsf{T}}$ assign a binary mark k_j on its support: a star center has mark one and its leaves zero; one endpoint of an edge has mark one and the other zero. Let

$$r = \frac{a + \sqrt{a^2 - c^2}}{c}, \qquad N_{jk}^{\pm} = \frac{\omega}{2} w_j w_k r^{\pm(k_j - k_k)}. \tag{108}$$

Each matrix is nonnegative rank one with identical row and column support. Their sum over the ten summands and two signs is exactly R, because diagonal/nonedge entries are unchanged and edge entries are multiplied by $(r + r^{-1})/2 = a/c$.

For a channel N, let $s = \sum_{ij} N_{ij}$. Introduce one hidden state of mass s/λ, entrance flows $N\mathbf{1}$ and exit flows $\mathbf{1}^{\mathsf{T}}N$. Its exit rate is λ, and its excursion occurrence matrix is $(N\mathbf{1})(\mathbf{1}^{\mathsf{T}}N)/s = N$. Visible masses and direct flows are as in (22). Balance holds at each hidden state and, since R and the direct flows are symmetric, at every visible state. Since $\sum_{ij} R_{ij} = 111/10$ and $\lambda = 111/5$, hidden mass is $1/2$. Boundary supports coincide and the direct visible graph is connected, proving reciprocity and irreducibility.

J.2 Why the comparison value is an exact full-law optimum

If an observed visible label contains several microscopic states, the boundary indices in the matrix construction below first range over those microscopic states. After forming the completely positive boundary matrix, sum its rows and columns by the observed labels. This nonnegative congruence preserves complete positivity, and Cauchy–Schwarz bounds each aggregated entry by the corresponding observed excursion rate. Grouping the endpoint path measures by those same labels gives the affinity lower bound below. Thus no visible-singleton assumption is used.

The following argument restates the inherited five-cycle theorem [10], not a new optimum. For an arbitrary compatible finite reciprocal chain, write stationary flows $n_{xy} = \pi_x q_{xy}$, hidden outflows $d_h = \sum_{y \neq h} n_{hy}$, and

$$Z_{ih} = \sqrt{n_{ih} n_{hi}}, \qquad J_{hh} = d_h, \quad J_{hk} = -\sqrt{n_{hk} n_{kh}} \ (h \neq k). \tag{109}$$

Positivity needs stationarity. For any real u the arithmetic–geometric mean inequality gives

$$u^{\mathsf{T}} J u \geq \sum_h \left[d_h - \frac{1}{2} \sum_{k \in H \setminus \{h\}} (n_{hk} + n_{kh}) \right] u_h^2 \tag{110}$$

$$= \frac{1}{2} \sum_h \left(\sum_{i \in V} n_{hi} + \sum_{i \in V} n_{ih} \right) u_h^2. \tag{111}$$

The remainder is a sum over unordered hidden pairs of $\frac{1}{2}(\sqrt{n_{hk}} u_h - \sqrt{n_{kh}} u_k)^2 + \frac{1}{2}(\sqrt{n_{kh}} u_h - \sqrt{n_{hk}} u_k)^2$. Every connected hidden component touches a visible state, so equality forces $u = 0$ by propagation

along reciprocal edges. Thus J is positive definite and has nonpositive off-diagonal entries. Its inverse square root is entrywise nonnegative, by $J^{-1/2} = \pi^{-1/2} \int_0^\infty t^{-1/2} e^{-Jt}\, dt$. Consequently $\overline{G} = ZJ^{-1}Z^{\mathsf{T}}$ is completely positive.

The Neumann/path expansion of J^{-1} shows that $\overline{G}_{ij}$ is the sum of geometric means of the stationary rates of a microscopic excursion and its reverse. A path $i, h_1, \dots, h_k, j$ has occurrence rate $n_{ih_1} n_{h_1 h_2} \cdots n_{h_k j}/(d_{h_1} \cdots d_{h_k})$; its geometric mean with the reverse has exactly the matrix-product factors above. Holding-time factors agree after reversal and integrate to one. Cauchy–Schwarz gives $0 \le \overline{G}_{ij} \le R_{ij}$ for the symmetric target.

The entropy assigned to excursions between two distinct endpoints with equal occurrence rates a is at least $4ah(\overline{G}_{ij}/a)$, by the affinity argument in Appendix F applied to their two equal-mass path-rate measures. Summing the path log ratios charges each hidden or boundary edge its stationary entropy contribution; finite mean jump counts and finite edge log ratios justify the summation. Copositivity of the Horn matrix gives

$$2 \sum_{\{i,j\} \in E(C_5)} \overline{G}_{ij} \le \sum_i \overline{G}_{ii} + 2 \sum_{\substack{i<j \\ \{i,j\} \notin E(C_5)}} \overline{G}_{ij} \le 5d + 10b = 10c. \tag{112}$$

Convexity and monotone decrease of h therefore imply $\sigma \ge 20ah(c/a)$; other path classes and direct visible edges contribute nonnegative terms. For the channels (108), every cycle-edge path has absolute log ratio $2 \log r$, and all other endpoint classes have ratio one. They attain this lower bound. Substitution gives (23).

J.3 Observed reversibility and the motif matrix

Visible holding-time factors depend only on the visible state occupied. A complete visible-to-visible segment has either a symmetric direct stationary-flow factor or a symmetric excursion factor $f_{ij}(t)$. The stationary density of a concatenated record is their product divided by the visible stationary masses at internal joins, times its visible holding-time factors. Reversing the record preserves these factors. Almost surely, a finite observation window can be enclosed by visible visits on both sides. Integrating the reversible enclosing records over their unobserved portions proves reversibility of the stationary label law on the finite window, including windows beginning in H.

For a hidden midpoint within an excursion from i to j, let s, u be the backward and forward distances to its endpoints. Its stationary density in (s, u) is $f_{ij}(s+u) = \lambda e^{-\lambda(s+u)} R_{ij}$. Integration over $0 \le s, u \le \tau$ yields the active block $t_\tau R$. Integrating one distance beyond τ and the other below it gives $q(1-q)/10$ for each active–censored entry. Two censored halves have mass $q^2/2$ at a hidden midpoint and mass $1/2$ at a visible midpoint. This proves (24) with total mass one.

The eigenvalues of R are

$$\frac{111}{50}, \qquad \frac{139 - 59\sqrt{5}}{200} \text{ (twice)}, \qquad \frac{139 + 59\sqrt{5}}{200} \text{ (twice)}. \tag{113}$$

The smaller eigenvalue is positive and exceeds $7/200$. The nonconstant active modes of M have these latter eigenvalues multiplied by t_τ. On the span of the censoring coordinate and the normalized constant active vector, the matrix is

$$\begin{pmatrix} (1+q^2)/2 & \sqrt{5}q(1-q)/10 \\ \sqrt{5}q(1-q)/10 & (1-q)^2/10 \end{pmatrix}, \tag{114}$$

whose determinant is $(1-q)^2/20 > 0$. Thus M is positive definite for every finite $\tau > 0$.

To certify the exact deletion deficit, factor the active minorant $t_\tau G_0 = UU^{\mathsf{T}}$ using the nonnegative factors above. Its row sum is $t_\tau \rho_0$ with $\rho_0 = 51/25$. Write $c_0 = q(1-q)/10$, and append to each column of U the top coordinate $c_0 \mathbf{1}^{\mathsf{T}} U_{\cdot r}/(t_\tau \rho_0)$. The resulting Gram matrix has the desired cross entries and top diagonal $5c_0^2/(t_\tau \rho_0) = (37/68)q^2$. The remaining top diagonal is $1/2 - 3q^2/68 > 0$, supplied by a coordinate factor. This gives a completely positive $G \le M$ agreeing with M except on the cycle edges of the active block. Its deleted mass is $10(a-c)t_\tau = (9/10)t_\tau = v$. The Horn dual value is also v, proving $\Delta_{\mathrm{cp}}(M) = v$.

J.4 Window choice and motif noise

The linear bound equals $\frac{9}{5}(1 - e^{-x})^2/x$ for $x = \lambda\tau$. Its derivative has the sign of $1 + 2x - e^x$. That expression starts at zero with positive derivative, is strictly concave and eventually negative, so it has exactly one positive zero. This establishes the stated linear optimizer.

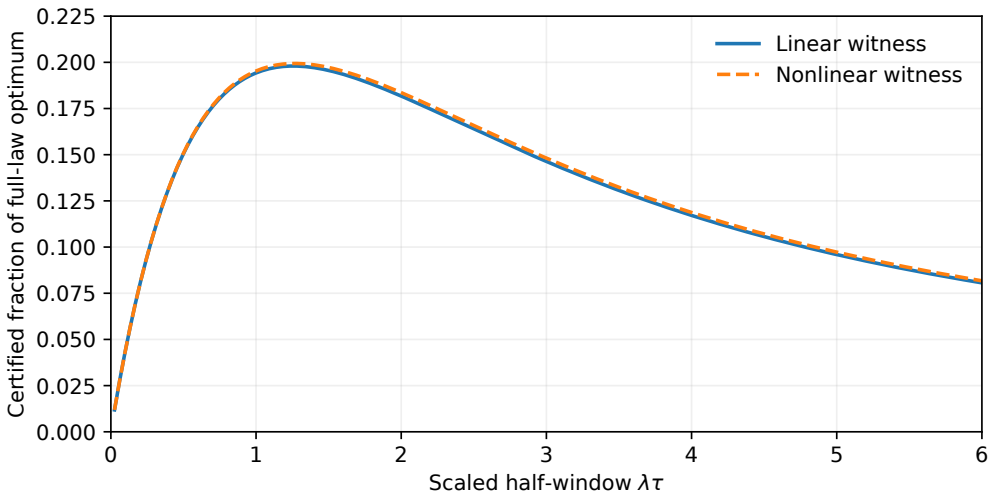


Figure 2: Finite-window certificates on the exact benchmark. The vertical scale is the certified bound divided by the inherited full-law optimum (23). The observed path-reversal divergence is zero. The nonlinear correction improves the linear certificate but does not make the finite-window bound tight. Curves evaluate analytical expressions; they are not simulation-based guarantees.

Let $C_\eta = (1-\eta)I + \eta J_5/5$. Since $R\mathbf{1} = (111/50)\mathbf{1}$,

$$C_\eta^{\mathsf{T}} R C_\eta = (1-\eta)^2 R + (2\eta - \eta^2)\frac{111}{250} J_5. \tag{115}$$

Using $\langle W, R\rangle = -9/10$ and $\langle W, J_5\rangle = 5$ gives (25) and (26). Extending C_η by the identity on category zero realizes this as a nonnegative stochastic congruence of the full motif matrix.

Proposition J.1 (Exact deletion under uniform motif noise)**.** *For the six-category benchmark matrix in* (24)*, apply the independent matched active-label channel* $C_\eta = (1-\eta)I + \eta J_5/5$*, leaving category zero unchanged. For every* $\tau > 0$ *and* $0 \le \eta \le 1$*,*

$$\Delta_{\mathrm{cp}}(M_\eta) = t_\tau[\frac{78}{25}(1-\eta)^2 - \frac{111}{50}]_+. \tag{116}$$

Thus M_η *is completely positive exactly when* $\eta \ge 1 - \sqrt{37/52}$*. Below this threshold the Horn witness is optimal for the deletion problem.*

Proof. Set $z = (1-\eta)^2$, $z_* = 37/52$ and $\alpha_0 = 111/250$. The active matrix is $R_z = zR + (1-z)\alpha_0 J_5$, with diagonal, edge and nonedge entries

$$d_z = z + (1-z)\alpha_0, \qquad a_z = \frac{3}{5}z + (1-z)\alpha_0, \qquad b_z = \frac{z}{100} + (1-z)\alpha_0.$$

Its row sum is $111/50$. The censoring cross entries remain $c_0 = q(1-q)/10$ and the censoring diagonal remains $(1+q^2)/2$.

If $z \ge z_*$, the star and edge vectors from Appendix J give the completely positive minorant

$$G_z = b_z \sum_i v_i v_i^{\mathsf{T}} + \frac{d_z - 6b_z}{2} \sum_i u_i u_i^{\mathsf{T}}.$$

Indeed $d_z - 6b_z \ge 37/1300$, and its diagonal and nonedges match R_z, while its edges are $d_z/2 + b_z \le a_z$. Its row sum $\rho_z = 2d_z + 4b_z$ is at least $51/25$. Complete a nonnegative Gram factorization of $t_\tau G_z$ with the censoring coordinate as in the preceding certificate. The generated top diagonal is

$$\frac{5c_0^2}{t_\tau \rho_z} = \frac{111q^2}{100\rho_z} \le \frac{37q^2}{68},$$

so the remaining coordinate mass is positive. This full minorant deletes only active cycle-edge mass, totaling $10t_\tau(a_z - d_z/2 - b_z) = t_\tau(78z/25 - 111/50)$. The Horn dual has the same value, proving equality.

At $z = z_*$ no mass is deleted, so R_{z_*} is completely positive. If $z \le z_*$, the identity

$$R_z = \frac{z}{z_*} R_{z_*} + \left(1 - \frac{z}{z_*}\right) \alpha_0 J_5$$

proves complete positivity of R_z. Its row sum 111/50 gives a full cross completion whose generated top diagonal is $q^2/2$; adding the remaining coordinate mass 1/2 proves $M_\eta \in \mathcal{CP}_6$. □

J.5 A kinetic baseline and weak-drive comparison

Consider a four-stage catalytic cycle with clockwise rate $u > 0$ and counterclockwise rate $w > 0$, in inverse-time units. The stages may represent free enzyme, ATP-bound enzyme, the post-hydrolysis ADP–phosphate complex, and ADP-bound enzyme. This homogeneous model is an idealized kinetic subfamily, not a fit to an enzyme. Its cycle affinity is $\mathcal{A} = 4\log(u/w) = \beta\Delta\mu$, where $\Delta\mu = \mu_{\mathrm{ATP}} - \mu_{\mathrm{ADP}} - \mu_{\mathrm{Pi}}$ is chemical work and $\beta = (k_{\mathrm{B}}T)^{-1}$ is inverse thermal energy; suitable state free energies distribute the chemical work consistently among the four elementary reactions [2]. Observe only $h = (0, 0, 1, 1)$, so each label aggregates two even microscopic states. With $a = u + w$ and $b = u - w$, stationarity is uniform and $\sigma = b\log(u/w)$. The reflection $j \mapsto 1 - j \pmod 4$ preserves h and conjugates the generator to its stationary reverse. Thus the entire observed law is reversible.

Apply the same classifier to each half-window: class 0 means no visible label switch before time τ, and class 1 means at least one switch. Put $c = \sqrt{uw}$ and

$$\begin{aligned} A_\tau &= e^{-a\tau}\left[\cosh(c\tau) + \frac{u}{c}\sinh(c\tau)\right], & B_\tau &= e^{-a\tau}\left[\cosh(c\tau) + \frac{w}{c}\sinh(c\tau)\right], \\ s_\tau &= \frac{A_\tau + B_\tau}{2}, & d_\tau &= \frac{(A_\tau - B_\tau)^2}{4}. \end{aligned} \tag{117}$$

The matched-classifier matrix and its exact deletion deficit are

$$S_\tau = \begin{pmatrix} A_\tau B_\tau & s_\tau - A_\tau B_\tau \\ s_\tau - A_\tau B_\tau & 1 - 2s_\tau + A_\tau B_\tau \end{pmatrix}, \qquad \Delta_{\mathrm{cp}}(S_\tau) = 2\left[(S_\tau)_{01} - \sqrt{(S_\tau)_{00}(S_\tau)_{11}}\right]. \tag{118}$$

Its determinant is $-d_\tau < 0$ whenever $u \ne w$ and $\tau > 0$. Consequently the finite-window certificate is positive at every such window, although observed path irreversibility and correlation-asymmetry estimators vanish [20].

Binary persistence obstructions already underlie the recent fixed-window certificate of Anonymous [15]. At the same window and conditional on either visible block, its coordinates are $p = 1/2$, $a_{\mathrm{EP}} = s_\tau$, and $b_{\mathrm{EP}} = A_\tau B_\tau$. Here $I(s, r)$ is the mutual information of the Bernoulli pair with matrix $\left(\begin{smallmatrix} r & s-r \\ s-r & 1-2s+r \end{smallmatrix}\right)$. Its one-block certificate is $I(s_\tau, A_\tau B_\tau)/\tau$. At fixed a, τ and $b \to 0$, writing $s_0 = e^{-a\tau/2}$ and $\kappa = e^{-2a\tau}\sinh^2(a\tau/2)/a^2$, our certificate is

$$\frac{2}{\tau}\mathcal{F}(\Delta_{\mathrm{cp}}(S_\tau)) = \frac{2\kappa b^2}{\tau s_0(1 - s_0)} + O(b^4), \qquad \frac{I(s_\tau, A_\tau B_\tau)}{\tau} = \frac{\kappa^2 b^4}{2\tau s_0^2(1 - s_0)^2} + O(b^6). \tag{119}$$

Adding the two equal block certificates changes only the latter constant. The quadratic order matches $\sigma = 2b^2/a + O(b^4)$. This is a comparison with that specific certificate: dwell-time variance inference also detects this model [4]. Since $\mathcal{CP}_2$ equals the doubly nonnegative cone, this baseline does not illustrate the distinction between complete positivity and positive semidefiniteness.

The signal also persists in a heterogeneous-rate neighborhood. For the reference generator Q_0, put $m = \max(s_\tau, 1 - s_\tau)$ and $z = (1 - s_\tau, -s_\tau)/m$. The fixed PSD witness $W_0 = zz^{\mathsf{T}}$ has score $v_0 = d_\tau/m^2$. For any stationary irreducible reciprocal generator Q on the same four states, with the same labels and classifier,

$$-\langle W_0, S(Q)\rangle \ge v_0 - \varepsilon(\sqrt{3}/a + 4\tau), \qquad \varepsilon = \|Q - Q_0\|_\infty, \tag{120}$$

where the norm is the maximum absolute row sum, including the diagonal. Thus $\varepsilon < v_0/(\sqrt{3}/a + 4\tau)$ ensures positive certification without homogeneous rates. This sufficient radius does not preserve the reference law's exact reversibility; Appendix J.6 proves the bound.

J.6 Calculation for the catalytic-cycle baseline

The row generator is $q_{j,j+1} = u$, $q_{j,j-1} = w$ and $q_{jj} = -(u + w)$, with indices modulo four. Uniform stationarity gives the entropy formula in Section J.5. For a chemical interpretation, assign reservoir work

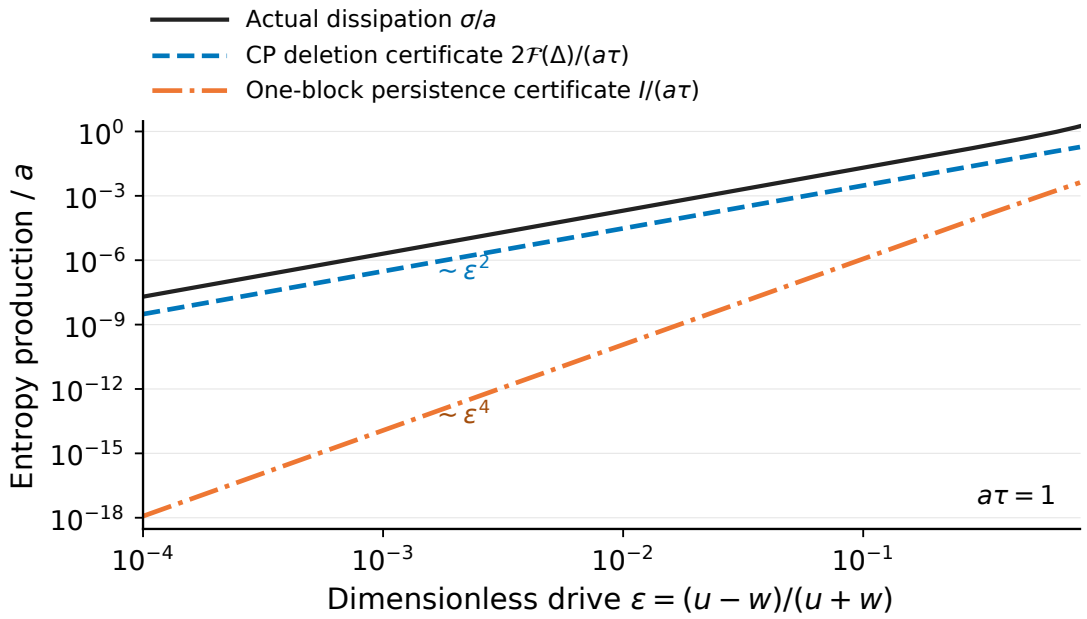


Figure 3: Weak driving in the four-stage cycle, with $a\tau = 1$ and $\varepsilon = (u - w)/(u + w)$. Both certificates use the same stationary persistence pair and total record duration 2τ. The blue curve is $2\mathcal{F}(\Delta_{\mathrm{cp}}(S_\tau))/\tau$; the orange curve is the specified prior one-block persistence certificate $I(s_\tau, A_\tau B_\tau)/\tau$ [15]. The former has quadratic order, the latter quartic order, as shown by (119). Rates are divided by a; curves evaluate analytical formulas, not simulations. The comparison does not rank other waiting-time estimators.

$(\mu_{\mathrm{ATP}}, 0, -\mu_{\mathrm{Pi}}, -\mu_{\mathrm{ADP}})$ to the four clockwise reactions and choose $G_{j+1} - G_j = \mathrm{work}_j - \Delta\mu/4$. The increments sum to zero, and local detailed balance gives $\log(u/w) = \beta\Delta\mu/4$ on every edge. This specifies a consistent operating point; varying the affinity while retaining homogeneous rates is not asserted to describe titration of one fixed enzyme.

Let R exchange 0 with 1 and 2 with 3. Then $RQR = Q^{\mathsf{T}}$, $R\pi = \pi$ and $h \circ R = h$. Applying this conjugacy to stationary path probabilities proves full observed-law reversibility. This reflection is not declared to be physical reversal of a microscopic conformation. Within either visible block the killed generator is

$$T = \begin{pmatrix} -a & u \\ w & -a \end{pmatrix}, \qquad e^{T\tau}\mathbf{1} = (A_\tau, B_\tau)^{\mathsf{T}}.$$

The reversed chain interchanges A_τ, B_τ. Conditional independence of the two half-paths at the midpoint therefore gives (118); all four probabilities are strictly positive. The determinant reduces to $-(A_\tau - B_\tau)^2/4$.

For any positive-entry 2×2 matrix $S = \left(\begin{smallmatrix} p & q \\ q & r \end{smallmatrix}\right)$ with $q > \sqrt{pr}$, every feasible completely positive minorant has $G_{00} \le p$, $G_{11} \le r$ and $G_{01} \le \sqrt{pr}$. The rank-one matrix $\left(\begin{smallmatrix} p & \sqrt{pr} \\ \sqrt{pr} & r \end{smallmatrix}\right)$ attains these bounds. Its retained mass gives $\Delta_{\mathrm{cp}}(S) = 2(q - \sqrt{pr})$. The positive-semidefinite dual witness with diagonal entries $\sqrt{r/p}, \sqrt{p/r}$ and off-diagonal entries -1 attains the same value.

At fixed a, write $u = (a + b)/2$ and $w = (a - b)/2$. Expansion of (117) gives $s_\tau = s_0 + O(b^2)$ and $d_\tau = \kappa b^2 + O(b^4)$. In terms of $s = s_\tau$ and $d = d_\tau$, the four probabilities are $(s^2 - d, s(1 - s) + d, s(1 - s) + d, (1 - s)^2 - d)$. Consequently

$$\Delta_{\mathrm{cp}} = \frac{d}{s(1 - s)} + O(d^2), \qquad I(s, s^2 - d) = \frac{d^2}{2s^2(1 - s)^2} + O(d^3).$$

Using $\mathcal{F}(v) = v + O(v^2)$ proves (119). The dwell entrance distribution is $(u/a, w/a)$, so $\mathbb{E}T = 2/a$, $\mathbb{E}T^2 = 6/(a^2 - uw)$, and $\mathrm{CV}^2(T) = 1 - b^2/[2(a^2 - uw)] < 1$ for $b \neq 0$. Thus the established equilibrium variance restriction detects the same family; no dominance over optimized waiting-time inference is claimed.

To prove (120), let $\Pi_0 = \mathbf{1}\pi_0$ and $Z_0 = \int_0^\infty (e^{tQ_0} - \Pi_0)\, dt$. Fourier diagonalization of the homogeneous cycle gives

$$\|e_j^{\mathsf{T}} e^{tQ_0} - \pi_0\|_1 \le \sqrt{2e^{-2at} + e^{-4at}} \le \sqrt{3}e^{-at}, \qquad \|Z_0\|_\infty \le \sqrt{3}/a.$$

For $E = Q - Q_0$, stationarity and $Q_0 Z_0 = \Pi_0 - I$ imply $\pi - \pi_0 = \pi E Z_0$, so $\|\pi - \pi_0\|_1 \le \sqrt{3}\varepsilon/a$. Couple transitions by their common rates until the first mismatch. Its hazard is at most $\sum_{k \ne j} |E_{jk}| \le \varepsilon$. On a duration-$2\tau$ record the path-law ℓ^1 distance is at most $\|\pi - \pi_0\|_1 + 4\tau\varepsilon$. Projection to the motif pair cannot increase it. The entries of W_0 have absolute value at most one, and $-\langle W_0, S(Q_0)\rangle = d_\tau/m^2$; this proves (120). Here ε has inverse-time units and $\sqrt{3}/a + 4\tau$ has time units.

K Proof of the copy-switch extension

K.1 Construction and statement

The following benchmark extension removes every singleton observed label while preserving a non-CP finite-window certificate. Let Q, π, h be the 25-state construction of Appendix J and let $D_\mu = \mathrm{diag}(\mu_x)$ with $0 < \mu_x \le \bar{\mu}$. Define a 50-state chain

$$\widetilde{Q}_\mu = \begin{pmatrix} Q - D_\mu & D_\mu \\ D_\mu & Q^\dagger - D_\mu \end{pmatrix}, \qquad \widetilde{\pi} = \tfrac{1}{2}(\pi, \pi), \qquad \widetilde{h}(x,+) = \widetilde{h}(x,-) = h(x). \tag{121}$$

Every microscopic state is still assigned even physical time-reversal parity.

Theorem K.1 (Reversible coarse records with no singleton label)**.** *The chain* (121) *is finite, irreducible and reciprocal. Its observed stationary process is exactly reversible and $\sigma(\widetilde{Q}_\mu) = \sigma(Q)$. None of its six observed labels is a singleton. The projection $(x, \pm) \mapsto x$ is not strongly lumpable when $Q \ne Q^\dagger$.*

Let M_μ be its motif matrix for the same first-visible-hit rule, with output zero at a visible midpoint. Put

$$\beta_h(\tau) = \frac{\lambda}{\lambda + 2\mu_h}(1 - e^{-(\lambda + 2\mu_h)\tau}), \qquad h \in H. \tag{122}$$

For every finite positive switching vector and every $\tau > 0$, the complete matrix M_μ is positive definite and the witness is exactly

$$v_\mu = \mathbb{E} Z_\mu = \frac{3}{37} \sum_{h \in H} \pi_h \beta_h(\tau)^2 > 0. \tag{123}$$

Thus $\sigma(\widetilde{Q}_\mu) \ge 2\mathcal{F}(v_\mu)/\tau > 0$ without a small-switching assumption. If the hidden switching rates satisfy $\mu_h \le \bar{\mu}$, then $v_\mu \ge 3\beta_{\bar{\mu}}(\tau)^2/74$, where $\beta_{\bar{\mu}}$ is (122) *with μ_h replaced by $\bar{\mu}$. At $\tau = 5/111$ and $\bar{\mu} = 10^{-3}$ the resulting uniform linear certificate is* 0.71918335149 *in the chosen inverse-time unit.*

The verification below is constructive. Switching the two copies conjugates the generator to its stationary reverse while preserving the observed labels. It is not a physical time-reversal operation in the all-even model. The nonlumpability assertion rules out interpreting the example merely as an independent, thermodynamically irrelevant duplication of each base state. It does not claim that every conceivable observable regenerative event has been eliminated.

K.2 Verification

Both Q and $Q^\dagger$ preserve π, and each switch has equal stationary flow in both directions, so (121) preserves $\widetilde{\pi}$ and has reciprocal support. Positive switches join two irreducible layers. Within-layer entropy contributions are $\sigma(Q)/2$ and $\sigma(Q^\dagger)/2$; switches add zero. This proves entropy preservation.

Let $S_\pm$ be the permutation interchanging the copies. Direct calculation gives

$$\widetilde{Q}_\mu^\dagger = S_\pm \widetilde{Q}_\mu S_\pm, \qquad \widetilde{h} \circ S_\pm = \widetilde{h}, \qquad \widetilde{\pi} \circ S_\pm = \widetilde{\pi}. \tag{124}$$

Therefore the reversed observed path law equals the original one. The microscopic physical parity remains the identity; interpreting $S_\pm$ as a physical odd-variable reversal would be a different model class. The rates from $(x,+)$ and $(x,-)$ to the block $\{(y,+),(y,-)\}$ are q_{xy} and $q_{xy}^\dagger$ for $y \ne x$. These differ somewhere because $\sigma(Q) > 0$. Hence the projection onto the base microscopic states is not strongly lumpable.

For the twenty-channel benchmark, every hidden state h has exit rate λ and no base hidden-to-hidden transitions. Write s_h for its total stationary entrance (equivalently exit) flow, so its base stationary mass is $\pi_h = s_h/\lambda$. Let p_h and q_h be its forward and reversed visible-exit distributions, respectively, and put

$$c_h = \frac{p_h + q_h}{2}, \qquad d_h = \frac{p_h - q_h}{2}, \qquad \alpha_\tau = 1 - e^{-\lambda\tau}, \qquad \beta_h(\tau) = \frac{\lambda}{\lambda + 2\mu_h}(1 - e^{-(\lambda + 2\mu_h)\tau}).$$

The classifier returns zero at a visible midpoint. At a hidden midpoint it reports the first visible label hit within τ, returning zero if none is hit. For this classifier, all switching rates at visible base states are irrelevant to the motif matrix. For every finite nonnegative hidden switching vector,

$$M_\mu - M_0 = 0 \oplus \sum_{h \in H} \pi_h (\alpha_\tau^2 - \beta_h(\tau)^2) d_h d_h^{\mathsf{T}} \succeq 0, \tag{125}$$

$$\mathbb{E} Z_\mu = \frac{3}{37} \sum_{h \in H} \pi_h \beta_h(\tau)^2 > 0. \tag{126}$$

Thus the complete six-category motif matrix remains positive definite, and the Horn witness remains strictly positive, for every finite choice of switching rates and every $\tau > 0$. Positive switching rates still ensure irreducibility of the full lifted chain. If $\mu_h \le \bar{\mu}$ for all hidden states, then

$$\mathbb{E} Z_\mu \ge \frac{3}{74} \left[\frac{\lambda}{\lambda + 2\bar{\mu}} (1 - e^{-(\lambda + 2\bar{\mu})\tau}) \right]^2 .$$

At $\tau = 1/\lambda = 5/111$ and $\bar{\mu} = 10^{-3}$, the resulting uniform linear certificate is

$$\sigma(\widetilde{Q}_\mu) \ge \frac{9}{5} \left[\frac{\lambda}{\lambda + 2\bar{\mu}} (1 - e^{-(1 + 2\bar{\mu}/\lambda)}) \right]^2 \simeq 0.71918335149.$$

The decimal evaluates a proved expression. It does not give the exact dissipation inferred from this motif matrix.

Proof. Until its first visible hit, the two copies of a base hidden state have killed generator

$$\begin{pmatrix} -\lambda - \mu_h & \mu_h \\ \mu_h & -\lambda - \mu_h \end{pmatrix}.$$

Its symmetric and antisymmetric modes decay at rates λ and $\lambda + 2\mu_h$. Integration of the visible exit probabilities gives the active-class conditional future vectors $\alpha_\tau c_h + \beta_h d_h$ and $\alpha_\tau c_h - \beta_h d_h$. The conditional past interchanges them. Both censoring probabilities are $e^{-\lambda\tau}$. Conditional independence of the two halves and stationary copy weights $\pi_h/2$ therefore give the active contribution

$$\pi_h \left(\alpha_\tau^2 c_h c_h^{\mathsf{T}} - \beta_h^2 d_h d_h^{\mathsf{T}} \right).$$

All entries involving category zero remain unchanged. At zero switching $\beta_h = \alpha_\tau$, proving (125). Since $0 < \beta_h \le \alpha_\tau$, the difference is positive semidefinite.

For $r = (20 + \sqrt{111})/17$, write $u = (r-1)/(r+1)$; then $u^2 = 3/37$. Up to cyclic position and the tilt sign, star channels have $c_h = (1/4, 1/2, 1/4)$ and $d_h = u(1/4, -1/2, 1/4)$; edge channels have $c_h = (1/2, 1/2)$ and $d_h = u(-1/2, 1/2)$. For the Horn matrix $W = J_5 - 2A_5$, direct multiplication gives $c_h^{\mathsf{T}} W c_h = 0$ and $d_h^{\mathsf{T}} W d_h = 3/37$ in both cases. Pairing the preceding active contribution with $-W$ proves (126). Finally $\sum_h \pi_h = 1/2$, and $\beta_h = \int_0^\tau \lambda e^{-(\lambda + 2\mu_h)t}\, dt$ decreases with μ_h, yielding the uniform bound. Positive definiteness follows from that of M_0 and (125). □

L Confidence, dependence and contamination

L.1 Detection power and dependence assumptions

For the noiseless benchmark at the linear-optimal window, $N \sim \mathrm{Bin}(n, s)$ with $s = (111/10)t_\tau$, and $K \mid N \sim \mathrm{Bin}(N, 20/37)$. The finite-sum power calculation (129) below evaluates the test without trajectory simulation. It gives the independent-snippet detection scales plotted in Figure 4; these are not single-trajectory acquisition guarantees. At $n = 10{,}000$, the level-0.05 test has power 0.992671, whereas the probability that the 95% bound in (28) is positive is 0.983216. The latter allocates level $\alpha/2$ to the conditional-sign endpoint. At the same sample size, relabeling rates 0.10 and 0.15 reduce test power to 0.395692 and 0.065699, respectively. Population tolerance therefore does not imply a fixed finite-data cost.

For a single stationary trajectory, suppose an externally justified bound

$$\sup_x \| P_g(x, \cdot) - \pi \|_{\mathrm{TV}} \le b(g) \tag{127}$$

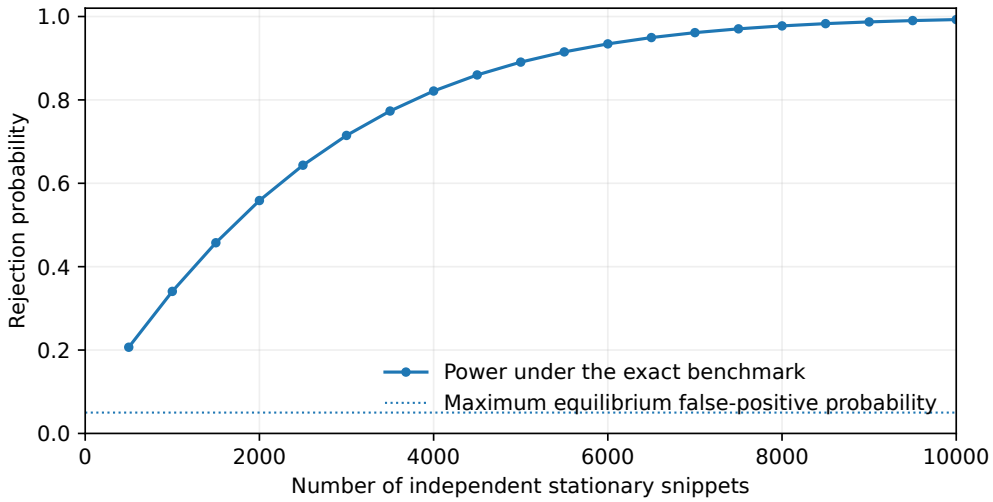


Figure 4: Power of the exact conditional-binomial equilibrium test at nominal level 0.05, for independent stationary snippets of the noiseless benchmark. Values are numerical evaluations of the finite binomial sum in (129), not Monte Carlo results. Dependence, nonstationary starts and learned selection rules would change the observation contract.

is available. Use n nonoverlapping snippets of length 2τ, separated by gaps of length g. The product-law coupling proved below implies that (27) has failure probability at most $\alpha + (n-1)b(g)$. Thus a target failure probability α_0 is obtained by using $\alpha = \alpha_0 - (n-1)b(g) > 0$ in the radius. This is conditional on (127); an empirical autocorrelation time is not silently substituted for it. The total deterministic observation span is $2n\tau + (n-1)g$.

For a finite list of K candidates evaluated on independent test snippets, a union bound replaces α by α/K; candidate scores may reuse the same independent test records. Training on the same trajectory adds a first-test separation budget $b(g_0)$, as proved below. Arbitrary adaptive choice of classifiers, windows or witness matrices from the test data is not justified by the fixed-rule bound.

L.2 Coverage and coupling proofs

For independent stationary snippets, Hoeffding's bound for variables in $[-1,1]$ is $\mathbb{P}(\overline{Z}_n - \mathbb{E}Z \geq r) \leq e^{-nr^2/2}$. Choosing $r = \sqrt{2\log(1/\alpha)/n}$ and applying the increasing function $\mathcal{F}$ proves (27). Under equilibrium $\mathbb{E}Z \leq 0$, equivalently $p_+ \leq p_-$ for the probabilities of $+1$ and -1. Conditional on N nonzero outcomes, their signs are independent Bernoulli with success parameter $p_+/(p_+ + p_-) \leq 1/2$. If $N = 0$ report p-value one. Otherwise the upper binomial tail at one half is conservative by stochastic monotonicity.

Define $L(0, n; \gamma) = 0$ and $L(0, 0; \gamma) = 0$. For $1 \leq k \leq n$, let $L(k, n; \gamma)$ be the unique $\ell \in (0,1)$ solving

$$\sum_{j=k}^{n} \binom{n}{j} \ell^j (1-\ell)^{n-j} = \gamma. \tag{128}$$

Monotonicity of the binomial upper tail gives $\mathbb{P}_p\{L(K, n; \gamma) > p\} \leq \gamma$ for $K \sim \mathrm{Bin}(n, p)$. Let $s = p_+ + p_-$ and, for $s > 0$, $p = p_+/s$. Then $N \sim \mathrm{Bin}(n, s)$ and $K \mid N \sim \mathrm{Bin}(N, p)$. Applying the lower-endpoint guarantee to s, and conditionally to p, with error $\alpha/2$ each, gives simultaneous lower bounds with probability at least $1 - \alpha$. Since $[\mathbb{E}Z]_+ = s[2p-1]_+$, their product proves (28). If $s = 0$, all scores vanish and the reported lower bound is zero. These are classical exact-binomial limits; the thermodynamic consequence uses Theorem 4.1.

For explicit power, let $c_\alpha(N)$ be the smallest integer k with $\mathbb{P}\{\mathrm{Bin}(N, 1/2) \geq k\} \leq \alpha$, or $N + 1$ if no such integer exists. Set $c_\alpha(0) = 1$. Under an alternative with active probability s and conditional positive probability p, the rejection probability is exactly

$$\mathrm{Power}_n(s, p) = \sum_{N=0}^{n} \binom{n}{N} s^N (1-s)^{n-N} \mathbb{P}\{\mathrm{Bin}(N, p) \geq c_\alpha(N)\}. \tag{129}$$

In the benchmark $p = 6/(111/10) = 20/37$. The finite sum determines the reported numerical power values; those evaluations are distinct from analytical coverage.

For the gapped-trajectory version, the first snippet has the stationary law. Conditional on the entire past up to an earlier snippet's end, the hidden state at the next snippet's start has law $P_g(x, \cdot)$, or a mixture thereof if only observations are conditioned on. Assumption (127) bounds its distance from π by $b(g)$. Continuing the path for 2τ and taking its score cannot increase this distance. Iterated maximal coupling, or a telescoping bound on conditional probability kernels, places the joint law of n scores within $(n-1)b(g)$ of the independent stationary product law. Any failure event is therefore bounded by its product-law probability plus this amount. The same event-transfer argument applies to the conditional-binomial test and (28). If the classifier and window are chosen from training data through time T, condition on that training history and put a gap g_0 before the first test snippet. The first starting law is then within $b(g_0)$ of stationarity. Iterating the same conditional-kernel comparison gives total failure probability at most $\alpha + b(g_0) + (n-1)b(g)$. Durations and gaps are fixed after training; test-data stopping or unrestricted selection is not covered. No assertion is made for overlapping snippets without an additional argument.

An external hidden-state mixing bound can be very conservative. For a constant copy-switch rate μ, projection onto the copy coordinate implies $b(g) \geq e^{-2\mu g}/2$. A dependence budget γ for $n-1$ gaps then requires $g \geq \log((n-1)/(2\gamma))/(2\mu)$ by this sufficient proof method. This is not a lower bound on all observable sampling procedures. In general an independent arbitrarily slow equilibrium hidden factor can leave the entire observed law unchanged, so the motif law cannot determine such a uniform hidden-state mixing budget.

There is also a deterministic robustness statement, distinct from matched relabeling. Suppose the distribution of a measured snippet differs in total variation by at most ε from the ideal snippet distribution. For any fixed score in $[-1, 1]$, $\mathbb{E}Z_{\text{ideal}} \geq \mathbb{E}Z_{\text{measured}} - 2\varepsilon$. Thus a population dissipation bound is obtained from $[\mathbb{E}Z_{\text{measured}} - 2\varepsilon]_+$. Independent measured snippets permit the additional Hoeffding subtraction. This requires a justified error bound; it does not infer one from the data. The explicit relabeling result in Section 5 is stronger when its matched-channel assumptions hold, since it needs no subtraction to remain valid.

M Open-case discussion: binary rank-$(3,2)$ realization, a subclass and an exact criterion

This appendix gives supporting results beyond Theorem 3.2. The universal rank-$(3,2)$ compression question remains open. Its normalization conventions are stated locally.

M.1 Model and exact open question

Let Q be a finite irreducible row CTMC generator, acting on column functions, with positive stationary law π and ordinary detailed balance. Two nonempty labels A, B are instantaneous deterministic functions of the state. The chronological future-function space V is the smallest real linear space containing $\mathbf{1}$ and invariant under Q and the two label masks. The question is whether

$$\dim V = 5, \qquad \dim(\mathbf{1}_A V) = 3, \quad \dim(\mathbf{1}_B V) = 2$$

always permits an irreducible five-state detailed-balance realization with three A states and two B states, preserving every stationary observed cylinder at every real time. No bound is imposed on the original hidden size, rates or topology. The physical time unit is fixed throughout.

The earlier $(2,2)$ result uses simultaneous reciprocal endpoint replacement of two one-dimensional feature intervals. A three-dimensional label space contains two nonconstant feature directions, so its geometry is different. The results below retain a constructive route and the possibility of an exact counterexample to universal compression.

M.2 A within-hull obstruction in the actual model

Lemma M.1. *The uniform law on $(1,1),(1,-1),(-1,1),(-1,-1)$ cannot be replaced by a law on at most three points of $[-1,1]^2$ preserving mass, mean and the full second-moment matrix.*

Proof. The two second moments are one. Since $1-X^2$ and $1-Y^2$ are nonnegative on the square with mean zero, both vanish almost surely. The support must therefore lie on the four corners. Mass, the two zero means and $\mathbb{E}[XY]=0$ uniquely force mass $1/4$ at every corner. All four are needed. □

This refutes only a hull-retaining moment replacement. New realization coordinates need not remain inside that hull.

In the six-state physical basis use

$$Q_0 = \begin{pmatrix} -7 & 1/2 & 1/2 & 0 & 4 & 2 \\ 1/2 & -6 & 0 & 3/2 & 4 & 0 \\ 1/2 & 0 & -5 & 5/2 & 0 & 2 \\ 0 & 3/2 & 5/2 & -4 & 0 & 0 \\ 4 & 4 & 0 & 0 & -9 & 1 \\ 2 & 0 & 2 & 0 & 1 & -5 \end{pmatrix}. \tag{130}$$

The first four states have label A, and the last two label B. The matrix is symmetric and conservative, with a connected positive-rate graph. Hence it is an irreducible reciprocal DB generator with $\pi_i = 1/6$. Entries have inverse-time units in the declared time unit.

On A put $X=(1,1,-1,-1)$, $Y=(1,-1,1,-1)$ and $Z=XY=(1,-1,-1,1)$. Extend Z by zero on B, writing the result z. Then $Q_0 z = -8z$. Its orthogonal complement W is reducing for Q_0 and both label masks, contains $\mathbf{1}$, and has label dimensions $(3,2)$. It is the complete chronological space: the B part of $Q_0\mathbf{1}_B$ is $(-8,-4)$, so both B coordinates are cyclic. Applying Q_0 to them gives $2(1+X)$ and $1+Y$ on A. Along with $\mathbf{1}_A$ these span all three A directions. Thus the complete rank is exactly five. Conditionally on A, the original stationary feature law is the square law of the lemma. The hull obstruction therefore occurs in a genuine input to the stated CTMC problem.

M.3 An exact five-state realization outside the old hull

Let P have its first three columns supported on A, with first four rows

$$P_A = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & -1 & -1 \end{pmatrix},$$

and let its last two columns be the original two B coordinate vectors. Then $P^TP = I_5$, $Q_0P = PB_0$, and

$$B_0 = \begin{pmatrix} -3 & -2 & -1 & 4 & 2 \\ -2 & -5 & 0 & 4 & 0 \\ -1 & 0 & -6 & 0 & 2 \\ 4 & 4 & 0 & -9 & 1 \\ 2 & 0 & 2 & 1 & -5 \end{pmatrix}, \qquad v = P^T\mathbf{1} = (2,0,0,1,1)^T.$$

Define the rational block-orthogonal matrix $O = \mathrm{diag}(O_A, O_B)$ by

$$A_0 = \begin{pmatrix} 0 & -\zeta & y \\ \zeta & 0 & -x \\ -y & x & 0 \end{pmatrix}, \quad O_A = (I - A_0)^{-1}(I + A_0),$$

$$O_B = \frac{1}{1+w^2}\begin{pmatrix} 1-w^2 & -2w \\ 2w & 1-w^2 \end{pmatrix}, \qquad (x, y, \zeta, w) = (-5/12, 1/5, -3/8, -1/12).$$

The skew symmetry of A_0 makes $I - A_0$ invertible and O_A orthogonal. Exact rational multiplication gives

$$h = O^Tv = \left(\frac{28598}{19501}, \frac{16800}{19501}, \frac{20520}{19501}, \frac{119}{145}, \frac{167}{145}\right)^T > 0.$$

Every off-diagonal of $H_0 = O^TB_0O$ is positive, with exact minimum $2756/2827645$. Every entry is determined by the displayed small rational inputs. Put $D = \mathrm{diag}(h)$ and

$$Q_0' = D^{-1}H_0D, \qquad \pi_i' = h_i^2/6.$$

Since $H_0h = 0$ and $\|h\|^2 = 6$, this is a conservative irreducible row generator with stationary mass π' and ordinary detailed balance. The first three states have label A and the last two B.

For the full-law claim, define $T = D^{-1}O^T$ on reduced coordinates. It satisfies

$$Q_0'T = TB_0, \quad Tv = \mathbf{1}, \quad TE_A = E_A'T, \quad T^T\,\mathrm{diag}(\pi')T = I_5/6. \tag{131}$$

Original functions Pc, Pd have stationary pairing $c^Td/6$. Equation (131) transports matrix exponentials and every alternating label-mask/semigroup word, including its stationary endpoint pairing. Therefore every stationary observed cylinder at arbitrary real time increments is identical. Equality of these finite-alphabet cadlag path laws follows, with no clock rescaling.

The transformed A features are

$$X' = \left(-\frac{13200}{14299}, \frac{1493}{1200}, \frac{164}{171}\right), \qquad Y' = \left(-\frac{1260}{14299}, -\frac{59}{35}, \frac{13349}{10260}\right).$$

They leave $[-1,1]^2$, as the moment obstruction requires. The original five-dimensional stationary reflected cylinder pairing is nondegenerate on V. Conditioning at its midpoint factors this pairing through any finite hidden realization, so every such realization has at least five states. The constructed realization is minimal.

M.4 A relative open family with one fixed compression map

Set $K = \mathbf{1}\mathbf{1}^T - 6I_6$ and $Q_* = Q_0 + K$. The following radius has inverse-time units in the same time unit as (130).

Theorem M.2. *Every real symmetric matrix Q satisfying*

$$Q\mathbf{1} = 0, \qquad Qz \in \mathrm{span}(z), \qquad \|Q - Q_*\|_2 < 1/4$$

is an irreducible positive DB generator with uniform stationary mass, complete chronological rank five and label dimensions $(3,2)$. The same fixed map $T = D^{-1}O^T$ yields an exact irreducible five-state DB realization with three A states, two B states and the original clock.

Proof. Write $E = Q - Q_*$. Each off-diagonal of Q_* is at least one, and $|E_{ij}| \le \|E\|_2 < 1/4$. Thus Q has positive off-diagonals; symmetry and zero row sums prove the original generator assertions. Since z is an eigenvector, $W = z^\perp$ is reducing and $QP = PB$, where $B = P^TQP$ and $Bv = 0$. Also

$$P^TKP = vv^T - 6I_5.$$

For $H = O^TBO$ the centre is

$$H_* = H_0 + hh^T - 6I_5,$$

whose off-diagonals exceed $(\min_i h_i)^2 = (119/145)^2 > 1/4$. Orthogonal compression gives $\|H - H_*\|_2 \le \|E\|_2 < 1/4$, so all off-diagonals of H are positive and $Hh = 0$. Define $Q' = D^{-1}HD$ and $\pi'_i = h_i^2/6$. The identities (131), with B, Q' in place of B_0, Q'_0, prove full-law transport.

To check complete cyclic rank, let $C = Q_{A,B}$. At the centre its column sums are 12 and 8. Their perturbation vector is $E_{A,B}^T\mathbf{1}_A$, of norm below $2\|E\|_2 < 1/2$. Their difference therefore changes by less than $\sqrt{2}/2 < 4$. The B part of $Q\mathbf{1}_B$ is nonconstant, so both B coordinates are cyclic. In the orthonormal A basis P_A, the central cross block is

$$P_A^TC_* = \begin{pmatrix} 6 & 4 \\ 4 & 0 \\ 0 & 2 \end{pmatrix}.$$

Its bottom two rows are $\mathrm{diag}(4, 2)$. Their perturbation has operator norm below $1/4$, so their least singular value is at least $2 - 1/4 > 0$. Applying Q to the two cyclic B coordinates and masking A gives two independent directions modulo $\mathbf{1}_A$. Hence all of W is cyclic. The complete rank is five, and the preceding rank lower bound proves minimality of the five-state realization. □

This neighbourhood is open relative to the stated linear constraints. It is not open among all six-state generators: perturbations revealing the invisible direction generally increase chronological rank. Nor does it cover all original hidden sizes, stationary laws or feature geometries in the open universal question.

M.5 A complete compact test for minimal reversible realization

The following criterion ranges over every minimal DB realization, rather than a chosen feature hull. It is an exact finite feasibility problem; it does not establish that the problem is always feasible.

Choose a stationary-inner-product orthonormal basis of the complete space V, separated by labels, and write B for the symmetric matrix of Q and v for the coordinates of $\mathbf{1}$. Normalize $\|v\| = 1$. Let $E = \mathrm{diag}(I_3, 0_2)$ be the A mask. The vector v is cyclic under B, E: their words applied to v span $\mathbb{R}^5$.

Theorem M.3. *An exact five-state ordinary-DB realization exists if and only if there is a block-orthogonal matrix $O \in O(3) \times O(2)$ such that*

$$h = O^Tv \ge 0, \qquad H = O^TBO, \qquad H_{ij} \ge 0 \quad (i \ne j).$$

Every feasible matrix automatically has $h_i > 0$ and a connected positive-off-diagonal graph. It yields the irreducible generator $D^{-1}HD$ and stationary masses h_i^2, where $D = \mathrm{diag}(h)$. Thus feasibility is a closed polynomial test on a compact set; a separate strict-positivity condition or prescribed stationary law is unnecessary.

Proof. Suppose first that such O exists. Since $Bv = 0$, $Hh = 0$. If $h_i = 0$, then $0 = (Hh)_i = \sum_{j \ne i} H_{ij}h_j$ forces every edge from i to the positive support of h to vanish. The zero support is therefore invariant under H and the diagonal mask E and orthogonal to all words applied to h. Cyclicity rules out any zero support. If the graph had multiple components, each component restriction of h would be a distinct nonzero kernel vector of H. But B has a simple zero eigenvalue: it is a restriction of an irreducible reversible generator containing its constant vector. Hence the graph is connected. The stated generator, masks, constant and stationary pairing transport exactly as in (131), now with unit normalization.

Conversely, consider any five-state DB realization of the same law. For a future observed cylinder F, write its conditional probability as a function g_F of the current hidden state, including the current label where specified. Ordinary reversibility identifies $\langle g_F, g_G\rangle_\pi$ with the probability of the corresponding time-reflected past cylinder and future cylinder joined at the present. Thus equality of stationary observed laws makes the map sending every original g_F to its new counterpart a well-defined isometry on their linear span. That span is V: finite-dimensional invariance under semigroups implies invariance under their derivatives Q, and conversely the semigroups are functions of Q. The masks are included in the

words. Nondegeneracy makes the isometry injective; dimension five makes it onto the entire new five-state function space. Cylinder extension intertwines the masks and semigroups, hence their generators and constants.

Pass to weighted physical coordinates in the new realization. Its generator becomes a symmetric Metzler matrix, its constant becomes $\sqrt{\pi'} > 0$, and the preceding isometry is orthogonal. It respects the label subspaces, which consequently have dimensions three and two. Its matrix is a block O of the stated kind, proving necessity. □

There is also an exact strict-margin form of the counterexample task. For each block-orthogonal O define

$$\Phi(O) = \sum_i [\min(0, (O^T v)_i)]^2 + \sum_{i<j} [\min(0, (O^T B O)_{ij})]^2.$$

The compact domain guarantees an attained minimum. The target law has no five-state DB realization exactly when $\min_O \Phi(O) > 0$. Rates in this formula are numerical entries in the fixed declared physical time unit; the two types of terms are only a feasibility penalty, not a physical cost. An optimizer failing to find zero is not a positive-margin certificate. Conversely a rigorous global lower bound would exclude all candidate minimal DB realizations, not merely a chosen hull or rotation.

For any positive minimum, sufficiently small perturbations of B, v retain infeasibility by uniform continuity on the compact orthogonal domain, provided they remain valid cyclic reversible inputs with the same profile. No such positive-margin example is constructed here. The rational rotation already exhibited supplies a feasible point for the six-state fixture and its relative family. The universal question is still open.

M.6 Eliminating the two-dimensional rotation exactly

The complete criterion can be tested with one fewer rotation parameter. This is an analytic elimination, not a global infeasibility certificate. Fix $O_A \in O(3)$ and write

$$A = O_A^T B_{AA} O_A, \quad C = O_A^T B_{AB}, \quad S = B_{BB}, \quad h_A = O_A^T v_A.$$

Any feasible choice must have $h_A > 0$ and $A_{ij} \geq 0$ off diagonal. Both are independent of O_B. Since a reflection of two coordinate axes differs from a rotation by a column permutation, and that permutation preserves all feasibility inequalities, it suffices to take

$$O_B(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

The vector v_B is nonzero. Give it angle α, with a fixed choice of angular lift. A nonzero row c_j of C in a feasible representation must have $c_j \cdot v_B > 0$, because its rotated coordinates are nonnegative and its dot product with $O_B^T v_B > 0$ is then positive. If this fails, the fixed O_A is infeasible. Otherwise give that row its unique lifted angle $\phi_j \in (\alpha - \pi/2, \alpha + \pi/2)$. Zero rows impose no angular condition.

The ground and cross-block inequalities are now exactly

$$\theta \in I := (\alpha - \pi/2, \alpha) \cap \bigcap_{j : c_j \neq 0} [\phi_j - \pi/2, \phi_j]. \tag{132}$$

Indeed, multiplication of a row by $O_B(\theta)$ subtracts θ from its angle; both its coordinates are nonnegative precisely when the resulting angle lies in $[0, \pi/2]$. The strict ground interval ensures both stationary coordinates are positive. These lifts lose no feasible rotation modulo 2π.

There is only one remaining B-off-diagonal constraint:

$$(O_B^T S O_B)_{12} = S_{12}\cos(2\theta) + \tfrac{1}{2}(S_{22} - S_{11})\sin(2\theta) \geq 0. \tag{133}$$

For the fixed O_A the complete criterion is therefore equivalent to the stated A conditions, the positive row dot products, and existence of a point of I satisfying (133). The allowed set of the last harmonic is a union of explicit closed half-period arcs (or the whole circle if both coefficients vanish). Intersecting those arcs with I is a finite exact angular test. Open ground endpoints must be retained: a zero attained only at an excluded endpoint is insufficient. Algebraic sign or interval certification is needed for exact numerical inputs; floating-point angle comparisons are not certificates.

This reduces the unresolved search to the three-dimensional O_A choice, with the entire O_B choice eliminated rather than sampled. It excludes a chosen O_A when its interval test fails, but does not exclude all O_A, construct a valid globally infeasible input, or prove universal feasibility. Both constructive and obstruction routes remain.

A rational quadratic version. Angles are unnecessary for exact computation. Write $v_B = (v_1, v_2)^T$, $\ell^2 = v_1^2 + v_2^2 > 0$, and $R_\alpha = \ell^{-1} \left(\begin{smallmatrix} v_1 & -v_2 \\ v_2 & v_1 \end{smallmatrix}\right)$. Put $S' = R_\alpha^T S R_\alpha$. Its entries require only division by ℓ^2, with no square-root coefficient. For each nonzero row c_j put $x_j = c_j \cdot v_B$, $y_j = v_1 c_{j2} - v_2 c_{j1}$; first require $x_j > 0$. With $t = \tan(\theta - \alpha)$ the complete remaining test is exactly

$$t < 0, \qquad x_j + y_j t \geq 0, \qquad y_j - x_j t \geq 0 \quad \text{for every row,}$$

$$S'_{12}(1 - t^2) + (S'_{22} - S'_{11})t \geq 0.$$

The first inequalities cut out an interval, and the last is a quadratic inequality. They follow by writing the relative rotation as $(1 + t^2)^{-1/2} \left(\begin{smallmatrix} 1 & -t \\ t & 1 \end{smallmatrix}\right)$; the omitted denominators are positive. Conversely every such t defines a rotation with strictly positive ground coordinates and all required rates nonnegative. Zero rows contribute only tautological inequalities. For rational B, v, O_A this is an exact univariate problem with rational coefficients, including strict-endpoint handling. It replaces the entire two-dimensional label-basis choice by linear inequalities and one quadratic, while leaving the global O_A question unresolved.