When Pairwise Counterfactual Bounds Are Exact: Conflict Certificates and Temporal Compression

Anonymous authors
Paper under double-blind review

Abstract

Pairwise counterfactual bounds can combine incompatible response mechanisms. We characterize when a supplied query attains this relaxation and construct a compact attained representation of the complete sharp function for a structured temporal class. For rational mechanism polytopes, occupied Bellman residuals and a backward no-leakage recursion give an exactness criterion and recover a rational witness, with polynomial bit cost at fixed coordinate dimension. For a structured three-state, binary-successor, same-action model with independent untied time laws and nonnegative rational one-defect terminal costs, original-domain repair, control of nonmetric histories and first-entry counting bound the entire representation by $(392n + 81)490^r$, where r counts exceptional layers. The guarantee transfers to standard exact full-cone pruning. Same-support complete-atom transport gives certified enclosures, with error scaled by terminal range and entrance mass. Coverage and chronological attainment are checked separately. Fixed three-state instances require $\Omega(r)$ attained values. A public token-circulation protocol supplies native $r = 0$ dynamics for a shared-response completion query. Synthetic comparisons measure block/generic cold cost ratio 0.6746 and changes in optimal route designs; supplied-query batches favor diagnosis or direct search. A growing-state construction establishes NP-hard evaluation and an exponential attained-family boundary.

1 Introduction

Marginal outcome or transition distributions do not specify how counterfactual worlds share responses. Optimizing each pair separately can demand incompatible complete mechanisms. Yet incompatibility outside a query's occupied pairs need not affect that query. We ask: *when is the pairwise bound sharp, and when can the entire sharp lower-bound function be represented compactly as the entrance varies?*

Optimization over shared response functions is established in counterfactual bounding (Balke & Pearl, 1994; Zhang et al., 2022); automated polynomial optimization also gives sharp bounds and certified outer approximations (Duarte et al., 2024). Dynamic counterfactual models retain shared noise or response uncertainty (Oberst & Sontag, 2019; Haugh & Singal, 2023). The incompatibility of simultaneous pairwise optima is itself known: generalized Gumbel couplers and pairwise multi-marginal transport study this obstruction under different coupling objectives (Lorberbom et al., 2021; Li & Anantharam, 2019). We study query-specific exactness and attained temporal representation within this shared-response framework.

For a fair three-state example, entrance masses a, b, c on $01, 12, 20$ give gap $\frac{1}{2}\min\{a, b, c\}$. Removing an occupied requirement eliminates the gap; changing positive weights changes its size without changing exactness. The distinction between diagnosing equality and optimizing a strict gap is central.

Two outputs. For query diagnosis, Theorem 2 decides equality by rational LPs and support recursion, with polynomial bit cost at fixed continuation dimension. For representation, Theorem 4 constructs finitely many attained vectors whose lower envelope is sharp for every $d \geq 0$, polynomial

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

in horizon at fixed exceptional count r. Original-domain repair, nonmetric-history control and first-entry counting also bound generic full-cone pruning. The growing-state boundary appears in Section 5. Same-support transport gives bounded-cost enclosures around compressible references (Corollary 19). We measure complete-function cost, changed route decisions, and the choice between diagnosis, direct search and compilation.

2 Shared mechanisms: general diagnosis and a structural subclass

General exactness interface. At time t, E_t indexes m_t continuation coordinates. A mechanism parameter lies in a nonempty compact rational polytope Q_t given by linear constraints. Its kernel $K_t(q) \in \mathbb{R}^{m_t \times m_{t+1}}$ is rational affine in q, nonnegative and substochastic on Q_t. A nonnegative rational terminal vector g and entrance $d \geq 0$ define

$$J(w) = d^T K_0(q_0) \cdots K_{n-1}(q_{n-1}) g, \qquad L = \min_{w \in \prod_t Q_t} J(w). \tag{1}$$

One mechanism parameter determines the kernel for the entire mass distribution at each time. Choices are offline, not conditional on an observed random pair. The product domain specifies independent time noises and untied mechanism laws. In a structural causal model (SCM), a joint law over complete deterministic response functions (a *complete response law*) specifies simultaneous potential responses (Balke & Pearl, 1994; Zhang et al., 2022). State–action rows share their response variables inside Q_t; equal marginals do not identify different variables. A sequence of feasible mechanism parameters is called a *word*; in the SCM specialization, it is a response-law sequence.

Set $\phi_n = g$ and compute the coordinatewise relaxation by LP:

$$\phi_t(e) = \min_{q \in Q_t} [K_t(q)\phi_{t+1}]_e, \qquad L_{\mathrm{rel}} = d^T \phi_0 \leq L. \tag{2}$$

We call the relaxation *tight* when its value equals the sharp minimum; the *exactness criterion* decides this equality, not arithmetic precision. For a suffix entrance x, write $L_t(x)$ for its sharp minimum and $G_t(x) = L_t(x) - x^T \phi_t$. Additional restrictions inside Q_t are retained by (2); an unconstrained coupling set must not replace its projection. Compactness gives attainment. The general theorem charges the full linear input representation, not a supplied list of vertices. Figure 1 summarizes this interface.

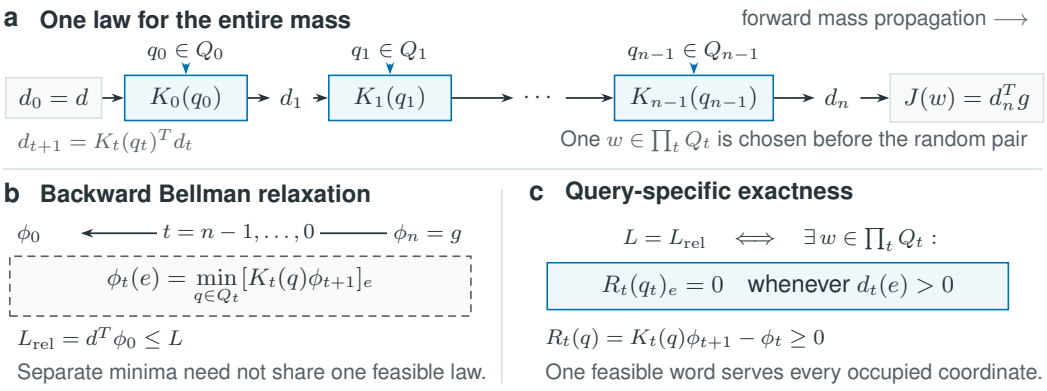


Figure 1: Shared mechanisms and query exactness. **a**, One law acts on the entire mass. **b**, Coordinatewise minima relax joint feasibility. **c**, Zero residual on every occupied coordinate characterizes exactness; backward support viability tests it (Theorem 2).

Three-state subclass for temporal compression. Here $i \in \{0, 1, 2\}$ has two sorted successors $s_{t,i}(0), s_{t,i}(1)$ and rational second-successor probability $p_{t,i} \in (0, 1)$. A simultaneous response atom

$z \in \{0,1\}^3$ defines $f_{t,z}(i) = s_{t,i}(z_i)$ and

$$Q_t = \{q \in \mathbb{R}^8_+ : \textstyle\sum_z q_z = 1,\ \sum_z z_i q_z = p_{t,i}\ (i = 0,1,2)\}, \tag{3}$$

$$[K_t(q)v]_{ij} = \sum_z q_z v_{f_{t,z}(i), f_{t,z}(j)}. \tag{4}$$

Both query worlds use the same scheduled core action, with independent time noises and untied response laws. A known selector Z and deterministic initial routes supply $d_{ij} = \sum_{z:r_1(z)=i, r_2(z)=j} \Pr(Z = z)$. The full rational table includes a probability-one factual wrapper and horizon $T = n+2$. Event queries use $g_{ij} = \mathbf{1}\{h(i) - h(j) \geq \eta\}$ with $h \in [0,1]^3$, $\eta > 0$; we also allow a supplied nonnegative rational zero-diagonal cost g. Its objective is expected cost. Write B^* for the entire model and terminal encoding (Appendix A.1). The core query is unconditioned.

For $g_{ij} = \mathbf{1}\{h(i) - h(j) \geq \eta\}$, equal states stay equal, so $E = (01, 02, 10, 12, 20, 21)$ suffices without renormalizing off-diagonal mass. Separate affinity gives an attaining word of original vertices $\mathcal{V}_t$; at most 70 candidate four-column bases enumerate each menu. In this unrestricted marginal polytope, every coupling of two nominal rows extends to a complete response law. Thus (2) is the coordinatewise pairwise-coupling relaxation (hereafter, the *pairwise relaxation*), with backup

$$(T_t^p v)_{ij} = \min_{\pi \in \Pi(P_t(i,\cdot), P_t(j,\cdot))} \sum_{u,v'} \pi(u,v') v_{uv'}\ (i \neq j), \quad (T_t^p v)_{ii} = 0, \tag{5}$$

with $L_{\mathrm{pair}} = L_{\mathrm{rel}}$. Larger-state same-action inputs retain diagonal absorption, but not the six-coordinate reduction or compression theorem. Different-action inputs may require diagonal coordinates as well; the general diagnosis interface admits them.

Why three states. With the same action and nonnegative zero-diagonal terminal, two-state common quantile laws attain the pairwise continuation vector at every horizon (Proposition 16). Three states are therefore the smallest actual-state count admitting a strict gap under these assumptions.

An eight-stage route decision. Guarded probabilistic commands specify supplied marginals (PRISM Project, 2010; 2021). Consider eight fair two-port routing stages, with three replaced by calibrated probability rows (Appendix L.2). Thus $r = 3$, with no holding layers. A uniform selector permits cyclic routing with probabilities x_j and unit budget $x_0 + x_1 + x_2 \leq 2$. Pairwise optimization selects $(1, 0, 1)$, with sharp guarantee 0.00655; the 21-vector compiled function selects $(181/385, 204/385, 1)$, giving 0.00756. Appendix L.2 reports all 24 programs and both budgets. Shared responses are unspecified; stages are independent and untied.

3 Exactness without enumerating mechanism vertices

For any feasible mechanism parameter define

$$R_t(q) = K_t(q)\phi_{t+1} - \phi_t \geq 0, \qquad d_{t+1} = K_t(q_t)^T d_t, \quad d_0 = d. \tag{6}$$

A coordinate is *occupied* when its induced mass is positive. We use *occupied Bellman residual* for the residual weighted by that mass; it is not a different Bellman operator. The telescoping device is an occupancy-weighted performance-difference argument (Kakade & Langford, 2002, Lemma 6.1), here applied to a finite-horizon relaxed continuation.

Proposition 1 (Occupied-residual identity)**.** *Every legal word satisfies*

$$J(w) - L_{\mathrm{rel}} = \sum_t d_t^T R_t(q_t), \qquad L - L_{\mathrm{rel}} = \min_{w \text{ legal}} \sum_t d_t(w)^T R_t(q_t). \tag{7}$$

Each summand is $d_{t+1}^T \phi_{t+1} - d_t^T \phi_t$, so the identity telescopes. An incumbent's residual includes avoidable choices; only the minimum is the unavoidable relaxation gap. For the fair three-state table $0 \to \{1,2\}, 1 \to \{0,2\}, 2 \to \{0,1\}$ with strict-order terminal and cyclic masses a, b, c, direct common-response-law calculation gives

$$L_{\mathrm{pair}} = c/2, \qquad L = c/2 + \tfrac{1}{2}\min\{a,b,c\}. \tag{8}$$

Unused diagonal mass is allowed. This example distinguishes support-dependent exactness from weight-dependent positive cost (Appendix B.2.4).

A face tests both present and future compatibility. For occupied $S \subseteq E_t$ and allowed next support $T \subseteq E_{t+1}$, put

$$\Psi_{t,S,T}(q) = \sum_{e \in S} R_t(q)_e + \sum_{\substack{e \in S \\ e' \notin T}} K_t(q)_{e,e'}. \tag{9}$$

Every term is nonnegative. The LP minimum is zero exactly when *one feasible mechanism parameter* eliminates the residual on each occupied coordinate and sends no mass outside T. Its nonempty zero set is a face of Q_t; we call this a *zero-residual face*, or *zero face*. A rational primal feasible witness certifies acceptance; a positive dual lower bound certifies rejection. The sum is used only for an exact zero test, not as a calibrated magnitude of future risk.

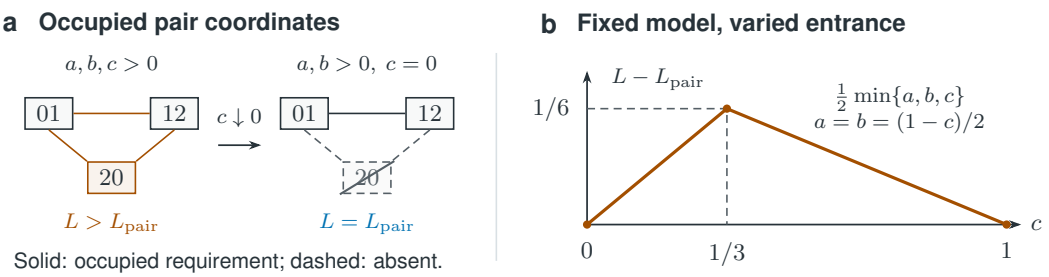


Figure 2: Occupied support controls exactness; weights control the magnitude of a positive gap. **a**, Nodes denote pair coordinates, not states in a causal graph. Removing 20 at $c = 0$ removes the conflict. **b**, With $d_{01} = d_{12} = (1 - c)/2$ and $d_{20} = c$, the curve is exactly (8). Both endpoints have zero gap; the transition table and terminal remain fixed.

Acceptable supports form a downward-closed family. Store its maximal members $\mathcal{A}_t$, with $\mathcal{A}_n = \{E_n\}$. Backward recursion is

$$W_t(S) \iff \exists T \in \mathcal{A}_{t+1} : \min_{q \in Q_t} \Psi_{t,S,T}(q) = 0. \tag{10}$$

Theorem 2 (Implicit-polytope exactness)**.** *For the general interface of Section 2, every suffix and $x \geq 0$ satisfy*

$$L_t(x) = x^T \phi_t \quad \iff \quad W_t(\operatorname{supp} x). \tag{11}$$

An accepting path returns a rational complete mechanism word. A direct implementation uses at most $\sum_t [m_t + 2^{m_t} |\mathcal{A}_{t+1}|]$ LPs over the supplied polytopes, without enumerating their vertices. For fixed continuation dimension its bit cost is polynomial in the supplied encoding and horizon, charging all mechanism variables, constraints and affine coefficients. Variable-dimensional support families can be exponential.

Proof. An accepted support has a feasible parameter in the zero-residual face and an accepted next target. Induction concatenates one feasible suffix with zero residual on that target; it works for all nonnegative weights supported there. Conversely, an attaining word with zero gap has zero residual on each occupied coordinate by Proposition 1. Nonnegativity makes its next support the union of positive outgoing entries. Its tight suffix support lies in a maximal accepted target, so the current law has zero residual and no leakage. This proves (10). Rational LP witnesses and accumulated denominator lengths give the stated encoding bound; Appendix B.1 includes empty supports and degeneracies. □

On the common embedded domain of Appendix I, all-state simultaneous solvability implies exactness for a supplied entrance, which implies current-layer compatibility; both converse implications fail. The no-leakage recursion distinguishes the latter two conditions: in the two-stage joint-response example of Appendix K, a current zero-residual face exists but cannot reach an acceptable next support.

For original vertex menus, the same test accepts S iff some vertex has zero residual on S and an accepted positive post-support (Appendix B.2.2). A fixed greedy choice among tied local laws is insufficient: it can destroy future viability. Conflict-directed support search and uncertain-entrance quantifiers are detailed in Appendix B.1.

Three optimization domains. Global bivariate consistency is stronger than separate Bellman minimization. For each unordered pair, choose a two-bit joint law with the given row marginals, use its transpose for the reverse pair, and optimize the entire untied time word offline. These pair laws need not extend jointly to all three responses. Projection of complete laws and Bellman induction give $L_{\mathrm{Bellman}} \le L_{\mathrm{bivariate}} \le L_{\mathrm{joint}}$. A legal four-layer threshold example satisfies

$$L_{\mathrm{Bellman}} = 4/405 < 17/1620 = L_{\mathrm{bivariate}} < L_{\mathrm{joint}} = 1/90.$$

Appendix E.1 specifies all domains and exact witnesses. The bivariate idea of Haugh & Singal (2023) is evaluated here on the common unconditioned, untied-time domain.

4 A SHARP LOWER-BOUND FUNCTION WITH BOUNDED TEMPORAL REPRESENTATION

Write $m_t(i)$ for row i's omitted successor and $k = \#\{t : \exists i,\ p_{t,i} \neq 1/2\}$.

Definition 3 (Compressible layers)**.** *A layer is D-fair if its three omissions differ and all rows are fair; X-matched if exactly two omissions agree and those rows have identical nominal distributions; or S-arbitrary if all omissions agree, with arbitrary interior row probabilities. Let r count all remaining layers.*

Recognition uses exact supports and marginals, not equality of realized responses. Every fair layer is included, so $r \le k$. For example, omissions 010 and probabilities (a_t, b_t, a_t) remain X-matched as both parameters vary: every layer can be nonfair while $r = 0$. The count records failures of these structural predicates.

Fair commands and duplicate-row interfaces impose D/X equalities by specification; estimated rows can instead be assessed through a same-support reference. For $0 \le g_{ij} \le G$, $g_{ii} = 0$, and a normalized entrance, complete-atom transport certifies $\max\{0, \bar{L} - G\Gamma\} \le L_p^g(d) \le J_p^g(M_{\mathrm{lift}}; d) \le \min\{G, \bar{L} + G\Gamma\}$, where $\bar{L} = L_{\bar{p}}^g(d)$ and $\Gamma = 1 - \prod_t (1 - \min\{1, \sum_i |p_{t,i} - \bar{p}_{t,i}|\})$. For entrance mass m, range and error bounds use mG. The transported SCM is feasible in the original model. Transport requires bounded costs; efficient reference optimization additionally uses the structure and terminal class of Theorem 4 (Appendix M.1).

Compressibility does not imply pairwise exactness. The all-D-fair family of Proposition 17 has $r = 0$ and $L = 2L_{\mathrm{pair}}$ at every odd horizon, with a decaying absolute gap; at even horizons it is exact. Thus the representation theorem also covers arbitrarily long, strictly loose queries.

A sufficient family H contains attained continuation values and satisfies

$$\min_{v \in H} a^T v = \min_{w \text{ legal suffix}} a^T v_w \qquad (a \ge 0). \tag{12}$$

Each $v \in H$ comes with an attaining word. Its region $C_v = \{a \ge 0 : a^T v \le a^T u \text{ for all } u \in H\}$ identifies the entrances where that word is optimal. These regions cover the entrance cone, allowing ties on their boundaries. Thus the output specifies sharp values and attaining mechanisms throughout the domain, including entrances not supplied during construction. The quantifier is $\forall a\, \exists w_a$. A fixed prefix sends a to a nonnegative vector, so sufficiency composes through that prefix and minimization over prefixes.

Theorem 4 (Refined temporal compression)**.** *For the three-state model of Section 2, let the supplied nonnegative rational zero-diagonal terminal g have at most one negative directed triangle slack. A constructible sufficient family has at most $(392n + 81)490^r$ values; if g is a directed quasimetric, $7 \cdot 490^r$ suffice. The sharp expected terminal-cost minimum, an attaining complete SCM and its certificate can be constructed and checked in $O(490^{2r}(n + B^* + 1)^{12})$ bit operations.*

Repair the domain before pruning. Let $\mathcal{C}_0$ be the nonnegative zero-diagonal directed quasimetric cone, defined by $\tau_{ij}(v) = v_{i\ell} + v_{\ell j} - v_{ij} \ge 0$. Let $\mathcal{C}_1$ permit at most one negative slack. Below, *metric* and *nonmetric* abbreviate membership and nonmembership in the directed quasimetric cone; symmetry and separation are not assumed. This class contains all nineteen threshold masks: zero, stars and strict orders are metric; a single edge has one defect. It also contains costs such as $(5/2, 1, 1, 1, 1, 1)$, whose only negative slack is $\tau_{01} = -1/2$. Common maps preserve $\mathcal{C}_0$, and maps with image size at most two send any nonnegative zero-diagonal cost into $\mathcal{C}_0$.

A binary three-state support permits zero or two permutation atoms. If both have positive mass, remove their common minimum and redistribute it to coalescing atoms. The marginals remain fixed; triangle slacks or opposite-pair sums show that the backup does not increase. In the distinct-omission case, one coalescer avoids the unique possible negative slack. This gives a no-worse *face point*. A face vertex is then selected for the induced scalar entrance; it need not dominate that point coordinatewise. At most seventy original face vertices suffice, and each preserves $\mathcal{C}_1$ (Appendix C.2).

Bound nonmetric survival. At X-matched, conditional repair of the duplicate responses gives four feasible laws with kernels $G + \lambda K_s^{\mathrm{fair}}$. The common term Gv is metric, so same-parent differences retain their fair signs. Two surviving nonmetric children have bad edges on different unordered pairs. The next matched layer can preserve a defect on only one required pair; thus at most one parent splits again. There are at most two active nonmetric histories *per input continuation*, even as labels and probabilities change. D-fair and S-arbitrary layers end nonmetric survival in this construction (Appendix C.4).

Complete a metric history without enumerating a block. A D-fair run has seven sufficient actual words, proved from the canonical identities $A_i^3 = A_i$, $A_j A_i \geq A_i^2$ and $N^2 = N + 2I$ with the normalization $A_i = 2K_i$. With a non-D layer, the rightmost such layer turns the D-suffix representatives into directed path/cut costs. Earlier common-quantile updates (Fill & Machida, 2002, Section 1.2) simultaneously minimize those path costs and preserve their class. Hence a metric block needs at most seven completions. The conclusion over arbitrary fair laws is scalar lower-envelope equality; wordwise coordinate domination is used only on the projected four-letter words, not on every convex mixture.

Why the length factors add. Let U, M count nonmetric and metric records before optional cross-record deletion. A block of length b and one exceptional layer satisfy, respectively,

$$U' \leq 2U, \qquad M' \leq 7M + (56b + 7)U, \tag{13}$$

$$U' \leq 70U, \qquad M' \leq 70M + 70U. \tag{14}$$

The factor $56 = 2 \cdot 4 \cdot 7$ counts two parents, four local choices and seven metric completions. There is no $M \to U$ term. A history pays its block-length multiplicity only upon its first backward entry into $\mathcal{C}_0$, never again.

Figure 3 depicts the absorbing structure. Summing over the unique first entry into $\mathcal{C}_0$, rather than multiplying bounds for successive block lengths, gives

$$N_{\mathrm{rep}} \leq (392n + 81)490^r. \tag{15}$$

A metric terminal needs at most $7 \cdot 490^r$ values. Appendix C.4.4 gives the complete summation and intermediate-record bounds, including empty blocks and $r = 0$. These constants bound representation and bit cost.

Correctness, bit cost and recovery. Local replacements preserve (12); feasible stored words provide the reverse inequality and attainment. For $G = \max g$, suffix coordinates lie in $[0, G]$; both numerator and denominator lengths are $O(B^* + n)$ along a word. Appendix C.4.4 accounts for intermediate records, repeated prefix construction, exact comparisons, provenance and complete-table output. Replay verifies every nominal row and the factual/query/query worlds, while lower-envelope coverage separately verifies optimality. These are distinct obligations.

A guarantee for generic exact pruning. Let X_t be the original-vertex suffix values and H_t the refined family. Multi-affinity gives $H_t \subseteq \mathrm{conv}(X_t)$, even for nonvertex matched laws. Sufficiency then implies

$$\mathrm{conv}(X_t) + \mathbb{R}^6_+ = \mathrm{conv}(H_t) + \mathbb{R}^6_+.$$

Its vertices belong to $H_t \cap X_t$. Sequential irredundant full-cone pruning therefore inherits the size bound, and $490^{O(r)}\,\mathrm{poly}(n + B^*)$ bit time when a polynomial-time exact LP subroutine is used (Appendix C.5). The argument operates in value space.

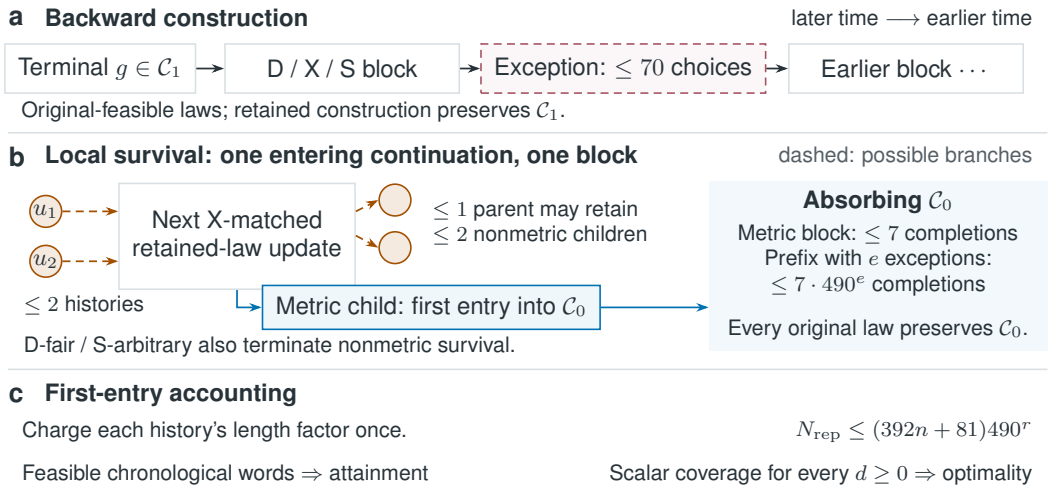


Figure 3: Compression proof. **a**, Retained original-feasible choices preserve $\mathcal{C}_1$. **b**, Each entering continuation has at most two nonmetric histories per block; only one parent can retain nonmetric descendants at the next X layer. Dashed arrows denote possibilities. Metric blocks need seven completions; a prefix with e exceptions needs at most $7 \cdot 490^e$. **c**, Charge first metric entry once; feasible words and scalar coverage establish attainment and optimality.

Nonmetric values need not disappear quickly. For every prescribed finite $L \geq 1$, a legal $n = L+2$ family with $k = r = 2$ has exactly four necessary values at each of its L prefix boundaries, two nonmetric. At the uniform entrance,

$$L(d) = \frac{1}{20\,4^L}, \qquad L_{\text{pair}}(d) = \frac{1}{30\,4^L}, \qquad L(d) - L_{\text{pair}}(d) = \frac{1}{60\,4^L}.$$

Appendix E.2 proves coverage and necessity with $O(L)$ suffix precision. This sustained survival explains the need to track nonmetric histories.

Exception count forces growing representation. For each $\ell \geq 1$, repeating the fixed three-state marginals $(1/10, 2/5, 3/5)$ for $n = r = 2\ell + 1$ layers with the disagreement terminal requires at least ℓ distinct attained values. Positive rational entrances uniquely expose these values over the full original law domain (Proposition 18). Thus necessary family size grows at least linearly with r, even for a fixed metric terminal.

5 A BOUNDARY WHEN THE ACTUAL STATE COUNT GROWS

Known MAX-CUT hardness in multimarginal optimal transport (MOT) and symmetric Bernoulli cut geometry give a complementary state-count result (Altschuler & Boix-Adserà, 2021; Huber & Marić, 2017); its scope is the present threshold/shared-mechanism model.

Proposition 5 (One-step state-count boundary)**.** *For every even $N = 2m \geq 4$, fix a one-step core whose row s has fair successors $s \bmod m$ and $s \bmod m + m$, rewards $h(s) = s/(2m-1)$, and threshold $1/(2m)$. For a nonnegatively weighted graph G on the lower m states with total weight $W > 0$, put its normalized edge weights on entrances $i > j$. Then*

$$L_G = 1 - \text{MaxCut}(G)/(2W), \qquad L_{\text{pair,G}} = 1/2.$$

Exact sharp-value computation is NP-hard in the compact input encoding; the unweighted threshold decision subfamily is NP-complete. Pairwise exactness for these graph entrances is equivalent to bipartiteness of the positive-weight support graph. For the same fixed core and terminal at each N, the minimum size of a finite attained family sufficient for all nonnegative directed core entrances *is exactly 2^{N-1}.*

A maximum cut mixed with its complement attains the graph value; directed entrances expose each complement-pair law, proving the family bound (Appendix F). Here N counts actual states and rows are compactly encoded. The family bound measures explicit attained representations.

6 EXACT OUTPUTS, DECISION EFFECTS AND COMPUTATION COST

Feasible words establish attainment; coverage establishes optimality. Structural certificates check original-menu reductions, noncircular convex deletions and pending branches (Appendix G). Shared rational primitives and exact checking of numerical proposals define the trust boundary.

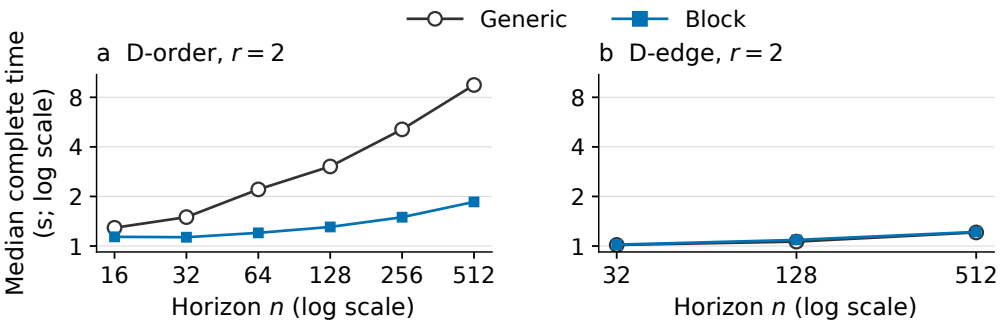


Figure 4: Current-source complete-function costs on both controlled D horizon slices: fixed $r = 2$, dispersed placement and generation parameter 3, with all horizons retained. Five processes per input/method. Historical curves and varying-r slices remain in Appendix N.

1. Complete all-entrance cost and avoided work. Five fresh GMP processes per method on each of 72 inputs give block/generic cold-cost ratio 0.6746, descriptive paired-input bootstrap interval $[0.6028, 0.7495]$, and 65/72 lower block medians. Both methods return the same complete envelope and attainment information. Launch through output and exit includes construction, certification, serialization and checking; inputs, sources and selection rules were fixed before measurement. At $n = 512$, D-order generic/block medians are 9.509/1.853s, compared with 1.209/1.217s for the D-edge control (Figure 4).

Generic exact full-cone pruning inherits the representation bound, but its intermediate work depends on pruning order. Instrumented counts and exclusive phases identify avoided backup and LP work. On 24 order-check inputs, ascending/descending/sum generic use 9304/7203/9304 LPs versus block's 360. The corresponding timing and transfer cohorts remain separate in Appendix N. The historical 72-input ratio 0.5827 remains separate, with its interval, construction costs and every input preserved in Figure 5 and Appendix N.

2. Changes in sharp answers and route decisions. For a supplied rational route polytope X and legal affine entrance $d(x)$, a complete family gives the finite LP

$$\max_{x \in X, z} z \quad \text{subject to} \quad z \le d(x)^T v \quad (v \in H).$$

Coverage and exact LP optimality certify the global design; an active vector supplies its attaining mechanism. This is a repeated-design output, distinct from a supplied scalar query. The original and transfer timing cohorts have maximum normalized corrections 2.28×10^{-5} and 4.57×10^{-15}. The 24 short D/X programs in Table 2 connect computation with larger corrections: six have multiple necessary vectors, and eight of 48 route comparisons improve the decision; 40 tie. The other 18 programs are pairwise exact. In the eight-stage, 21-vector example, unit-cost design selects $(181/385, 204/385, 1)$ in place of the pairwise route $(1, 0, 1)$, gaining $6227/6149220$ in the sharp objective (Appendix L). On this same cohort, block/descending is 0.9928, interval $[0.9768, 1.0057]$: 0.9491 on the six multi-vector cases and 1.0078 on singletons.

The equalities come from specified fair commands and shared interfaces. For the indicator costs and normalized entrances here, a same-support compressible reference improves the pairwise bound when $\bar{L} - L_{\text{pair}}(p) > \Gamma$. Perturbing every layer can raise r to n; four-stage exact checks find 18/96 positive route gains, and reference certificates prove 16 against the complete perturbed pairwise-optimal face. Reference-route loss is at most 1.395×10^{-5} at the largest budget (Appendix M.1). Holding programs preserve gap 0.01960 through 512 stages by scaling a fixed core; fixed per-stage noise instead decays the signal (Appendix L).

3. Diagnosis, direct search, or compilation. An accepted query support yields $L(d) = d^T\phi$ and an attaining word. For unresolved queries, standard branch-and-bound (Lawler & Wood, 1966) searches original response-law prefixes. Prefix mass x gives lower bound $x^T\phi_t$; feasible words give incumbents. Exact support closures and valid bound pruning reduce the original word tree. A proof checker verifies mass propagation, complete menu coverage, exact-state reuse and attainment (Appendix N.3). The method reads no compiled family or saved answers.

Table 1: Exact queries on eight $n = 8, 12$ programs; K counts supplied queries. Seconds span four input medians (three processes each); ratios geometrically average paired inputs. D is multi-vector, X singleton. P tests support before block compilation.

Group	K	Direct (s)	Block (s)	P (s)	Direct/Block	P/Block
D	1	0.073–0.081	1.269–1.456	1.265–1.449	0.0578	0.9974
D	24	0.219–0.406	1.331–1.543	1.330–1.536	0.1954	0.9982
X	1	0.072–0.078	1.179–1.198	0.072–0.077	0.0626	0.0625
X	24	0.156–0.205	1.251–1.293	0.143–0.187	0.1416	0.1286

All eight existing $n = 8, 12$ order programs enter the matched query comparison. $K = 1$ uses $(1, 2, 4, 8, 16, 32)/63$, unresolved on all four D inputs; $K = 24$ uses six basis and 18 mixed entrances. Cold costs include preprocessing, proof serialization/readback/checking and every requested SCM replay. Numerical dependency loading belongs to this endpoint; ratios do not isolate search work. Singleton X inputs provide controls (Table 1). Direct search wins all four D inputs at both workloads, using 55–283 proof nodes for one query and 1992–6360 for 24. No crossover is observed through 24 queries. On X, support diagnosis closes every query; at 24 queries the shared singleton certificate is cheaper than separate direct proof traces. Thus these supplied batches favor diagnosis or direct search, while compilation produces the reusable complete function.

Across eight models and twelve supplied contracts each, sharp route design improves on the best shared value over the entire pairwise-optimal face in 23/96 cases; 73 tie. Positive gains are 1.64×10^{-6}–0.001013 in absolute probability and 0.676–33.416% relative to that face's best value. For twelve-task batches, direct/compile cold ratios are 0.1120 on D and 0.1569 on X (Appendix L.3); matching exact bounds certify each route.

Appendix N.2 compares enumeration and support/interval methods. All matched exact-query timings use $\tau = 0$.

Public protocol dynamics. The Israeli–Jalfon three-process ring supplies a native D-fair layer in its documented stable circulation phase (Israeli & Jalfon, 1990; PRISM Project, n.d.; Hartmanns et al., 2019). A bijection preserves its three token positions and fair transitions, giving $r = 0$ for every horizon. We pose a new untied shared-response completion query on these marginals. All 90 queries over ten horizons are solved exactly; 20 have positive pairwise gaps, and every compiled envelope has three values (Appendix L.6).

7 Related work

Appendix I compares the exactness criterion and compression theorem with shared-response optimization, simultaneous solvability and value-vector pruning.

AI USE STATEMENT

The informal lemmas were written by the authors. We used generative AI tools to produce complete formalizations of these lemmas, to polish the writing, and to retrieve related work. All AI-assisted material was reviewed by the authors, who take responsibility for the final content.

REPRODUCIBILITY STATEMENT

The appendices supply complete proofs and the dependency table. The source contains editable vector figures and exact construction checks. The anonymous supplement provides generic, block and direct algorithms, independent certificate checkers, complete measurement ledgers and regeneration commands. It includes unchanged public protocol sources, a checked adapter, exact query certificates, and a generator and checker for the parameter lower bound. Lean sources and historical build and axiom-audit records cover the theorem endpoints identified in Appendix J; executable certificate checks cover the Python outputs. Historical timings, fresh correctness checks and analytic proofs have separate evidence entries in the claim map.

ETHICS STATEMENT

The computational studies use synthetic inputs and a public probabilistic protocol model. The retained historical randomized-email materials contain sufficient counts and reconstruction records, without customer-level rows. The research involved no new participant contact or intervention and makes no individual treatment recommendations.

REFERENCES

Jason M. Altschuler and Enric Boix-Adserà. Hardness results for multimarginal optimal transport problems. *Discrete Optimization*, 42:100669, 2021. doi: 10.1016/j.disopt.2021.100669. URL https://arxiv.org/abs/2012.05398v1. Section locators refer to arXiv:2012.05398v1.

Jason M. Altschuler and Enric Boix-Adserà. Polynomial-time algorithms for multimarginal optimal transport problems with structure. *Mathematical Programming*, 199:1107–1178, 2023. doi: 10.1007/s10107-022-01868-7. URL https://doi.org/10.1007/s10107-022-01868-7.

David L. Applegate, William Cook, Sanjeeb Dash, and Daniel G. Espinoza. Exact solutions to linear programming problems. *Operations Research Letters*, 35(6):693–699, 2007. doi: 10.1016/j.orl.2006.12.010. URL https://doi.org/10.1016/j.orl.2006.12.010.

Alexander Balke and Judea Pearl. Counterfactual probabilities: Computational methods, bounds and applications. In *Proceedings of the Tenth Conference on Uncertainty in Artificial Intelligence*, pp. 46–54, 1994. doi: 10.1016/B978-1-55860-332-5.50011-0. URL https://arxiv.org/abs/1302.6784.

Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. doi: 10.1017/CBO9780511804441. URL https://web.stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf.

Anthony R. Cassandra, Michael L. Littman, and Nevin L. Zhang. Incremental pruning: A simple, fast, exact method for partially observable Markov decision processes. In *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence*, pp. 54–61, 1997. URL https://arxiv.org/abs/1302.1525.

Krishnendu Chatterjee, Laurent Doyen, Thomas A. Henzinger, and Jean-François Raskin. Algorithms for Omega-Regular games with imperfect information. *Logical Methods in Computer Science*, 3(3), 2007. doi: 10.2168/LMCS-3(3:4)2007. URL https://lmcs.episciences.org/1094. Article 4.

Krishnendu Chatterjee, Martin Chmelík, Raghav Gupta, and Ayush Kanodia. Optimal cost almost-sure reachability in POMDPs. *Artificial Intelligence*, 234:26–48, 2016. doi: 10.1016/j.artint.2016.01.007. URL https://doi.org/10.1016/j.artint.2016.01.007.

John W. Chinneck and Erik W. Dravnieks. Locating minimal infeasible constraint sets in linear programs. *ORSA Journal on Computing*, 3(2):157–168, 1991. doi: 10.1287/ijoc.3.2.157. URL https://doi.org/10.1287/ijoc.3.2.157.

Yesong Choe, Yeahoon Kwon, Min Woo Park, and Sanghack Lee. Canonical domain reduction for partial counterfactual identification. In *Proceedings of the 42nd Conference on Uncertainty in Artificial Intelligence*, volume 337 of *Proceedings of Machine Learning Research*, pp. 1302–1326. PMLR, 2026. URL https://proceedings.mlr.press/v337/choe26a.html.

Johan de Aguas, Sebastian Krumscheid, Johan Pensar, and Guido Biele. The probability of tiered benefit: Partial identification with robust and stable inference. In *Proceedings of the Fourth Conference on Causal Learning and Reasoning*, volume 275 of *Proceedings of Machine Learning Research*, pp. 90–113. PMLR, 2025. URL https://proceedings.mlr.press/v275/aguas25a.html.

Guilherme Duarte, Noam Finkelstein, Dean Knox, Jonathan Mummolo, and Ilya Shpitser. An automated approach to causal inference in discrete settings. *Journal of the American Statistical Association*, 119(547):1778–1793, 2024. doi: 10.1080/01621459.2023.2216909. URL https://doi.org/10.1080/01621459.2023.2216909.

Zhengzhu Feng and Shlomo Zilberstein. Region-based incremental pruning for POMDPs. In *Proceedings of the Twentieth Conference on Uncertainty in Artificial Intelligence*, pp. 146–153, 2004. URL https://arxiv.org/abs/1207.4116.

James Allen Fill and Motoya Machida. Realizable monotonicity and inverse probability transform. In Carles M. Cuadras, Josep Fortiana, and José A. Rodriguez-Lallena (eds.), *Distributions With Given Marginals and Statistical Modelling*, pp. 63–71. Springer, Dordrecht, 2002. doi: 10.1007/978-94-017-0061-0_8. URL https://doi.org/10.1007/978-94-017-0061-0_8. Section and equation locators refer to arXiv:math/0010026v1.

Robert Ganian, Marlene Gründel, and Simon Wietheger. Gateways to tractability for satisfiability in Pearl's causal hierarchy. arXiv:2511.08091v2, 2026. URL https://arxiv.org/abs/2511.08091v2. Accepted at ICML 2026; Theorem 11 refers to the version 2 preprint.

Zijun Gao, Shu Ge, and Jian Qian. Bridging multiple worlds: Multi-marginal optimal transport for causal partial-identification problem. In *Proceedings of the 28th International Conference on Artificial Intelligence and Statistics*, volume 258 of *Proceedings of Machine Learning Research*, pp. 721–729. PMLR, 2025. URL https://proceedings.mlr.press/v258/gao25b.html. The formulation comparison uses the full version, arXiv:2406.07868v2.

Vineet Goyal and Julien Grand-Clément. Robust Markov decision processes: Beyond rectangularity. *Mathematics of Operations Research*, 48(1):203–226, 2023. doi: 10.1287/moor.2022.1259. URL https://arxiv.org/abs/1811.00215.

Julien Grand-Clément, Nian Si, and Shengbo Wang. Dynamic programming-compatible uncertainty sets in robust Markov decision processes. arXiv:2411.08435v2, 2026. URL https://arxiv.org/abs/2411.08435v2. First version 2024; theorem locators refer to the version 2 PDF.

Arnd Hartmanns, Michaela Klauck, David Parker, Tim Quatmann, and Enno Ruijters. The quantitative verification benchmark set. In *Tools and Algorithms for the Construction and Analysis of Systems*, volume 11427 of *Lecture Notes in Computer Science*, pp. 344–350. Springer, 2019. doi: 10.1007/978-3-030-17462-0_20. URL https://link.springer.com/chapter/10.1007/978-3-030-17462-0_20.

Martin B. Haugh and Raghav Singal. Counterfactual analysis in dynamic latent state models. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 12647–12677. PMLR, 2023. URL https://proceedings.mlr.press/v202/haugh23a.html.

Mark Huber and Nevena Marić. Bernoulli correlations and cut polytopes. arXiv:1706.06182v2, 2017. URL https://arxiv.org/abs/1706.06182v2.

Amos Israeli and Marc Jalfon. Token management schemes and random walks yield self-stabilizing mutual exclusion. In *Proceedings of the Ninth Annual ACM Symposium on Principles of Distributed Computing*, pp. 119–131. ACM, 1990. doi: 10.1145/93385.93409. URL https://doi.org/10.1145/93385.93409.

Garud N. Iyengar. Robust dynamic programming. *Mathematics of Operations Research*, 30(2): 257–280, 2005. doi: 10.1287/moor.1040.0129. URL https://doi.org/10.1287/moor.1040.0129.

Leslie Pack Kaelbling, Michael L. Littman, and Anthony R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1–2):99–134, 1998. doi: 10.1016/S0004-3702(98)00023-X.

Sham M. Kakade and John Langford. Approximately optimal approximate reinforcement learning. In *Proceedings of the Nineteenth International Conference on Machine Learning*, pp. 267–274, 2002. URL https://homes.cs.washington.edu/~sham/papers/rl/aoarl.pdf.

Ege Can Kaya, Mahsa Ghasemi, and Abolfazl Hashemi. Joint MDPs and reinforcement learning in coupled-dynamics environments. In *Proceedings of the 42nd Conference on Uncertainty in Artificial Intelligence*, volume 337 of *Proceedings of Machine Learning Research*, pp. 2877–2893. PMLR, 2026. URL https://proceedings.mlr.press/v337/kaya26b.html. The sampler definition and scope comparison refer to arXiv:2603.06946v1.

Jessica Lally, Milad Kazemi, and Nicola Paoletti. Robust counterfactual inference in Markov decision processes. In *Proceedings of the 25th International Conference on Autonomous Agents and Multiagent Systems*, pp. 1527–1535. IFAAMAS, 2026. doi: 10.65109/TXUQ4572. URL https://www.ifaamas.org/Proceedings/aamas2026/pdfs/TXUQ4572.pdf. Theorem D.7 is in the full version, arXiv:2502.13731v5.

E. L. Lawler and D. E. Wood. Branch-and-bound methods: A survey. *Operations Research*, 14(4): 699–719, 1966. doi: 10.1287/opre.14.4.699. URL https://pubsonline.informs.org/doi/10.1287/opre.14.4.699.

Cheuk Ting Li and Venkat Anantharam. Pairwise multi-marginal optimal transport and embedding for earth mover's distance. arXiv:1908.01388v2, 2019. URL https://arxiv.org/abs/1908.01388v2.

Mengmeng Li, Daniel Kuhn, and Tobias Sutter. Policy gradient algorithms for robust MDPs with nonrectangular uncertainty sets. *SIAM Journal on Optimization*, 36(1):120–151, 2026. doi: 10.1137/24M1631250. URL https://doi.org/10.1137/24M1631250. The full-version locator in the text refers to arXiv:2305.19004v4.

Yan Li and Alexander Shapiro. Rectangularity and duality of distributionally robust Markov decision processes. *Mathematical Programming*, 2025. doi: 10.1007/s10107-025-02297-y. URL https://doi.org/10.1007/s10107-025-02297-y. Online publication; locators refer to the publisher version.

Sirui Lin, Zijun Gao, José Blanchet, and Peter Glynn. Causal partial identification via conditional optimal transport. In *Proceedings of the 29th International Conference on Artificial Intelligence and Statistics*, volume 300 of *Proceedings of Machine Learning Research*, pp. 1189–1197. PMLR, 2026. URL https://proceedings.mlr.press/v300/lin26b.html. The estimation comparison uses the full version, arXiv:2506.00257v2.

Guy Lorberbom, Daniel D. Johnson, Chris J. Maddison, Daniel Tarlow, and Tamir Hazan. Learning generalized Gumbel-max causal mechanisms. In *Advances in Neural Information Processing Systems*, volume 34, pp. 26792–26803. Curran Associates, Inc., 2021.

Shie Mannor, Ofir Mebel, and Huan Xu. Lightning does not strike twice: Robust MDPs with coupled uncertainty. In *Proceedings of the 29th International Conference on Machine Learning*, pp. 385–392, 2012. URL https://icml.cc/2012/papers/215.pdf.

Vrettos Moulos. Bicausal optimal transport for Markov chains via dynamic programming. In *2021 IEEE International Symposium on Information Theory (ISIT)*, pp. 1688–1693. IEEE, 2021. doi: 10.1109/ISIT45174.2021.9517977. URL https://doi.org/10.1109/ISIT45174.2021.9517977. Theorem locators refer to the full version, arXiv:2010.06831v1.

Arnab Nilim and Laurent El Ghaoui. Robust control of Markov decision processes with uncertain transition matrices. *Operations Research*, 53(5):780–798, 2005. doi: 10.1287/opre.1050.0216. URL https://doi.org/10.1287/opre.1050.0216.

Michael Oberst and David Sontag. Counterfactual off-policy evaluation with Gumbel-Max structural causal models. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 4881–4890. PMLR, 2019. URL https://proceedings.mlr.press/v97/oberst19a.html.

Judea Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, 2 edition, 2009. doi: 10.1017/CBO9780511803161. URL https://doi.org/10.1017/CBO9780511803161.

PRISM Project. The PRISM language: Commands, 2010. URL https://www.prismmodelchecker.org/manual/ThePRISMLanguage/Commands. Official manual; page last modified 2010-12-03; accessed 2026-09-26.

PRISM Project. The PRISM language: Modules and variables, 2021. URL https://www.prismmodelchecker.org/manual/ThePRISMLanguage/ModulesAndVariables. Official manual; page last modified 2021-03-18; accessed 2026-09-26.

PRISM Project. Randomised self-stabilising algorithms, n.d. URL https://www.prismmodelchecker.org/casestudies/self-stabilisation.php. Official case study, Israeli and Jalfon section; accessed 2026-09-26.

Martin L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, 1994. doi: 10.1002/9780470316887. URL https://doi.org/10.1002/9780470316887.

Igal Sason. Entropy bounds for discrete random variables via maximal coupling. *IEEE Transactions on Information Theory*, 59(11):7118–7131, 2013. doi: 10.1109/TIT.2013.2274515. URL https://arxiv.org/abs/1209.5259v5. Theorems 1–2 refer to the final author version, arXiv:1209.5259v5.

Richard D. Smallwood and Edward J. Sondik. The optimal control of partially observable Markov processes over a finite horizon. *Operations Research*, 21(5):1071–1088, 1973. doi: 10.1287/opre.21.5.1071. URL https://doi.org/10.1287/opre.21.5.1071.

Jin Tian and Judea Pearl. Probabilities of causation: Bounds and identification. In *Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence*, pp. 589–598, 2000. URL https://arxiv.org/abs/1301.3898. The cited version is the UAI proceedings paper.

Wolfram Wiesemann, Daniel Kuhn, and Berç Rustem. Robust Markov decision processes. *Mathematics of Operations Research*, 38(1):153–183, 2013. doi: 10.1287/moor.1120.0566.

Marco Zaffalon, Alessandro Antonucci, Rafael Cabañas, David Huber, and Dario Azzimonti. Efficient computation of counterfactual bounds. *International Journal of Approximate Reasoning*, 171:109111, 2024. doi: 10.1016/j.ijar.2023.109111. URL https://arxiv.org/abs/2307.08304.

Junzhe Zhang, Jin Tian, and Elias Bareinboim. Partial counterfactual identification from observational and experimental data. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 26548–26558. PMLR, 2022. URL https://proceedings.mlr.press/v162/zhang22ab.html.

The appendices give the model and exactness proofs, temporal compression, explicit separation constructions, the variable-state boundary, the certificate contract and the primary protocol. Historical and optional branches are indexed separately in the source bundle.

A Model, response vertices and SCM construction

A.1 Complete model and the entrance wrapper

The input is a complete finite rational table $P_s(y \mid x, a)$, $0 \le s < T$, on finite states $\mathcal{S}$ and two actions, with an initial state, fixed factual actions and trajectory, two fixed open-loop query sequences, and a supplied rational nonnegative terminal pair cost g on the core, with $g_{ii} = 0$. Event queries specialize to reward $h : \mathcal{S} \to [0, 1]$, threshold $\eta > 0$, and $g_{ij} = \mathbf{1}\{h(i) - h(j) \ge \eta\}$. Every factual-path row assigns probability one to the stated successor, so the factual event is certain under every compatible mechanism and has no unknown conditioning denominator.

A response is a complete deterministic table $F_s : \mathcal{S} \times \mathcal{A} \to \mathcal{S}$. Its law reproduces every nominal row, and its realized table is shared by the factual and both query worlds. Time laws are independently chosen with independent noises and no cross-time parameter ties. Query actions are fixed before these noises are realized.

Both queries first read the same selector row, obtaining z with known probability λ_z. A second transition routes z deterministically to $r_1(z), r_2(z) \in \{0, 1, 2\}$ under their respective actions, giving

$$d_{ij} = \sum_{z:\, r_1(z)=i,\ r_2(z)=j} \lambda_z, \qquad i, j \in \{0, 1, 2\}. \tag{16}$$

One selector outcome per positive pair represents any rational nine-pair entrance. The wrapper, including all deterministic rows, is part of the complete input.

The full wrapper has $T \ge 3$ and $n = T - 2 \ge 1$ core transitions. Index them by $t = 0, \dots, n - 1$. At each such time both queries use the same action. Each of the three actual core rows has exactly two distinct core successors, listed in increasing core-label order as $s_{t,i}(0), s_{t,i}(1)$, with second-successor probability $p_{t,i} \in (0, 1) \cap \mathbb{Q}$. These supports may change with time. Except for the selector row and these three rows per core time, every actual row is deterministic. The three core states are actual states of the table, not three independently assigned coordinates for each world.

A complete core response atom is $z \in \{0, 1\}^3$, with map $f_{t,z}(i) = s_{t,i}(z_i)$. Thus the entire local domain is

$$\mathcal{Q}_t = \left\{ q \in \mathbb{R}^8_+ : \sum_z q_z = 1, \quad \sum_z z_i q_z = p_{t,i} \quad (i = 0, 1, 2) \right\}. \tag{17}$$

Every compatible complete mechanism induces such a law; conversely, each such law extends to a complete mechanism by the specified deterministic responses on the other rows. For a pair value v, put

$$[K_t(q)v]_{ij} = \sum_z q_z v_{f_{t,z}(i), f_{t,z}(j)}. \tag{18}$$

In particular, the same coordinate z_i is reused when $i = j$. For the supplied cost g, the sharp compatible expected-cost minimum is exactly

$$\min_{q_t \in \mathcal{Q}_t} d^{\mathsf{T}} K_0(q_0) \cdots K_{n-1}(q_{n-1}) g. \tag{19}$$

Here and below the minimum chooses one law per time for the entire entrance mass. Diagonal values are zero throughout, because equal states remain equal when both worlds read the same action and response table. We therefore use the six-coordinate order $(01, 02, 10, 12, 20, 21)$. Its operators are substochastic; the full nine-pair operators are stochastic. Diagonal entrance mass is retained and contributes zero, without renormalization.

Let B^* denote the complete binary input length, including the full table, wrapper, actions, factual path, and terminal numerators and denominators. For event-specialized statements we retain B for the encoding of rewards and threshold; generating its six-entry mask has polynomial cost. The parameter is

$$k = \#\{t : \text{some } p_{t,i} \ne \tfrac{1}{2}\}. \tag{20}$$

Equality with $1/2$ is exact rational equality. Fair layers still have unknown common response couplings. The selector is not counted in k.

A.2 Original response vertices and attainment

The equality matrix in (17) has columns $(1, z_0, z_1, z_2)^{\mathsf{T}}$ and rank four, witnessed by $000, 100, 010, 001$. A feasible point is a vertex exactly when its positive-support columns are independent: dependence gives a supported null direction with feasible small perturbations of both signs; independence precludes a nontrivial convex decomposition. Every independent support extends to a four-column basis. Solving the $\binom{8}{4} = 70$ candidate systems, setting other entries to zero, testing feasibility and deduplicating therefore enumerates all vertices, including degenerate ones.

Each $\mathcal{Q}_t$ is compact and nonempty (it contains the independent-bit law). The continuous, separately affine objective (19) attains a minimum. Replace each time law of a minimizer by a vertex minimizing that coordinate's linear objective with the other laws fixed. Its value cannot rise or fall below the global minimum. Thus an original-vertex word attains the optimum, using the product time domain without correlating noises.

A.3 Recovery and verification of a complete common SCM

Expand the selected word and suffix pointers to one law q_t per core time. Original vertices have at most four positive atoms and fair letters have two; a general retained law uses eight entries. For every atom z, fill core rows by $f_{t,z}(i) = s_{t,i}(z_i)$ and all other rows by their prescribed deterministic successors. At the selector time use the complete table of mass λ_z for each outcome; at routing use the single deterministic table. These specify all T laws.

Partition one uniform variable into intervals of the atom masses at each time, sharing that variable across worlds but using independent variables across times. The resulting SCM neither observes the realized entrance pair nor randomly selects a correlated sequence of laws. Explicit word length is polynomial. Expanding complete tables also has polynomial output cost in the full input; selector expansion can be quadratic in the auxiliary-state count and is included in Theorem 4.

For direct attainment verification, first sum the atom masses producing each successor at every original row and compare exactly with its nominal probabilities. Next propagate the full triple of factual and query states. If $\rho_s(x_0, x_1, x_2)$ is its current distribution and ν_s is the chosen complete-table law, update by

$$\rho_{s+1}(y_0, y_1, y_2) = \sum_{x_0, x_1, x_2} \rho_s(x_0, x_1, x_2) \sum_F \nu_s(F) \prod_{j=0}^{2} \mathbf{1}\{F(x_j, a_s^{(j)}) = y_j\}, \tag{21}$$

Here $a^{(0)}$ is factual. Initialize the triple at the supplied initial state and check that the factual coordinate follows its stated path with total mass one. Sum $\rho_T(y_0, y_1, y_2) g_{y_1 y_2}$ and compare with the returned objective. For event masks this equals the terminal mass satisfying the threshold. A nonnormalized nonnegative entrance is normalized for the selector and the objective multiplied by its total mass; the zero entrance is immediate. The replay uses one complete map in each summand and checks every nominal row and all three worlds. Its state space has size at most $|\mathcal{S}|^3$, with polynomial product-denominator bit lengths.

This forward computation proves attainment. Optimality comes separately from the lower-envelope induction and complete menu coverage; an attaining witness alone would supply only an upper bound on the minimum. Distinct queries or entrance laws may select different attaining SCMs. The construction does not identify the unknown true response coupling.

B Tightness and support-dependent certificates

B.1 Implicit mechanism polytopes and zero-face exactness

We prove Theorem 2 on the linear-input interface of Section 2. The finite-menu and three-state recurrences below are specializations, not extensions of the temporal compression theorem.

B.1.1 Complete statement and the local face test

Let $Q_t \subseteq \mathbb{R}^{p_t}$ be nonempty compact rational polytopes, supplied by linear constraints, and let $K_t(q)$ be rational affine, nonnegative, substochastic $m_t \times m_{t+1}$ kernels. The terminal is $g \in \mathbb{Q}_+^{m_n}$ and choices range over the product $\prod_t Q_t$. Define ϕ_t and R_t by (2) and (6). Each coordinate optimum is an LP; hence $R_t(q) \geq 0$ for every feasible q. For a legal suffix and $x \geq 0$, telescoping gives

$$x^T K_t(q_t) \cdots K_{n-1}(q_{n-1}) g - x^T \phi_t = \sum_{s=t}^{n-1} x_s^T R_s(q_s), \qquad x_{s+1} = K_s(q_s)^T x_s.$$

No diagonal reduction or common query action is needed. Distinct state–action response variables must nevertheless remain distinct, even when their marginals coincide.

Every term in $\Psi_{t,S,T}$ is nonnegative, so its zero constraints are exactly

$$R_t(q)_e = 0 \quad (e \in S), \qquad K_t(q)_{e,e'} = 0 \quad (e \in S, e' \notin T).$$

Their feasible set is a supporting face of Q_t, possibly the whole polytope. Compactness makes the minimum strictly positive when this set is empty. Rational primal feasibility or a positive LP dual lower bound certifies the alternatives; this uses standard LP weak and strong duality (Boyd & Vandenberghe, 2004, Sections 5.1–5.2). Only simultaneous zero is used; residual and leakage scales do not quantify the gap.

B.1.2 Backward induction, witnesses and bit complexity

Accepted supports are downward closed: a law and suffix valid on S remain valid on any subset. Thus only their maximal members $\mathcal{A}_t$ are needed, with $\mathcal{A}_n = \{E_n\}$.

For sufficiency, select a zero-face law from S into an accepted T. Every nonnegative entrance supported on S has zero current residual and next mass supported on T. The same recorded suffix works for all such next masses by induction. Product feasibility permits concatenation, and telescoping proves equality. The word can depend on S, but not on a realized pair.

For necessity, compactness supplies an attaining word. Zero total gap forces each occupied residual to vanish. Nonnegativity gives the exact next support

$$S' = \{e' : K_t(q)_{e,e'} > 0 \text{ for some } e \in S\}.$$

Its attaining suffix is tight, so induction puts S' in an accepted maximal T. The current law has zero residual and no leakage outside T. Empty S accepts since $\Psi \equiv 0$; empty T is allowed for substochastic kernels and has zero terminal contribution.

The direct bound is $\sum_t (m_t + 2^{m_t} |\mathcal{A}_{t+1}|)$ LPs plus subset bookkeeping. All dimensions, constraints and any categorical atom expansion are charged to input size. A rational-polytope vertex has polynomial local encoding by determinant bounds. Each coordinate optimum can select such a vertex; using a common denominator for ϕ_{t+1}, its backup adds local kernel denominator factors. Taking common multiples across at most m_t choices adds their bit lengths, rather than squaring the suffix denominator. Substochasticity bounds $\|\phi_t\|_\infty$ by $\|g\|_\infty$. Nonempty zero faces contain vertices, and optimal rational LP duals also have polynomial encoding. This proves polynomial bit cost at fixed coordinate dimension, not for an unrestricted support lattice. Avoiding vertex enumeration does not remove the cost of an explicitly supplied K^p-atom model.

B.1.3 Conflict-directed compilation and certificates

Representing downward-closed knowledge sets by maximal antichains is established in imperfect-information games (Chatterjee et al., 2007, Section 3.2); filtering minimal infeasible constraint sets is established in LP diagnosis (Chinneck & Dravnieks, 1991). The construction below applies these tools to target-specific zero-residual and no-leakage requirements. Fix the target T and begin with E_t. Acceptance covers every subset. On rejection, greedily remove requirements while infeasibility persists to obtain a core $C \subseteq S$. A feasible subset must omit some $e \in C$: including all of C makes its nonnegative residual/leakage sum at least the positive core optimum. Branches $S \setminus \{e\}$ therefore cover every feasible subset. A deletion once found feasible remains feasible after further removals, so greedy shrinking is inclusion-minimal. Minimality improves branching; certified positivity alone gives soundness.

Algorithm 1 Zero-face compilation for one time and one accepted target

1 Initialize a queue with E_t, accepted maxima $\mathcal{F} = \varnothing$, and rejected cores $\mathcal{C} = \varnothing$.
2 Remove S; skip it if already visited or contained in a member of $\mathcal{F}$.
3 If a saved $C \in \mathcal{C}$ is contained in S, use that target-specific positive certificate.
4 Otherwise solve the face LP on S. A zero witness inserts S into $\mathcal{F}$ and removes its subsets.
5 On rejection, shrink to a core C with certified positive optimum, caching all tests.
6 In this rejection branch, insert the sets $S \setminus \{e\}$, $e \in C$, into the queue.
7 After exhaustion return the feasible maxima, law witnesses and rejected-core certificates.

Run the search separately for each $T \in \mathcal{A}_{t+1}$ and retain maximal accepted sets in their union. Cache keys are (t, T): rejection for one target is not global rejection. The checker verifies rational certificates and, for each target, that every support is contained in an accepted maximum or contains a rejected core. Global rejection needs this coverage for every next target. These antichain, LP-duality and core-branching tools are applied to the specific zero-residual/no-leakage condition; work-count savings are not runtime evidence.

B.1.4 Independently specified entrance uncertainty

For a nonempty compact rational polytope $\mathcal{D} \subseteq \mathbb{R}_+^{m_0}$ independent of future domains, let

$$p_{\mathcal{D}} = \min_{x \in \mathcal{D}} x^T \phi_0, \quad \mathcal{D}^\star = \{x \in \mathcal{D} : x^T \phi_0 = p_{\mathcal{D}}\}, \quad L_{\mathcal{D}} = \min_{x \in \mathcal{D}, w} x^T V_w.$$

Corollary 6 (Robust-minimum versus uniform exactness)**.** *$L_{\mathcal{D}} = p_{\mathcal{D}}$ if and only if some $S \in \mathcal{A}_0$ admits $x \in \mathcal{D}^\star$ with $x_e = 0$ outside S. Separately, every $x \in \mathcal{D}$ is pointwise tight if and only if $W_0(S_{\mathcal{D}})$ holds, where $S_{\mathcal{D}} = \bigcup_{x \in \mathcal{D}} \operatorname{supp} x$.*

Proof. For each feasible pair, $x^T V_w - p_{\mathcal{D}} = (x^T \phi_0 - p_{\mathcal{D}}) + (x^T V_w - x^T \phi_0)$ is a sum of nonnegative terms. A compact-domain minimizer gives equality precisely on the relaxed optimal face at a tight support, proving the first claim by Theorem 2. For the second, average one feasible entrance for each coordinate in $S_{\mathcal{D}}$. Convexity yields an entrance with exactly that support. Its tightness and downward closure imply tightness for every other entrance; the converse follows by applying uniform tightness to this average. □

Compute $p_{\mathcal{D}}$ by LP and minimize $\sum_{e \notin S} x_e$ over $\mathcal{D}^\star$ for each accepted maximal S. A zero optimum accepts; positivity for every target rejects. Coordinate maximization finds the support union. These tests neither quantify a positive gap nor infer a joint entrance from marginals. A certified all-entrance family also gives $L_{\mathcal{D}} = \min_{h \in H} \min_{x \in \mathcal{D}} x^T h$.

B.2 Occupied conflict and exact support certificates

These refinements use the original three-state response domains, nonnegative zero-diagonal terminals and unnormalized six-coordinate entrances. A general terminal's encoding is included in input size; the one-defect assumption is not needed here.

B.2.1 Residual identity and the gap functional

Let $V_{t,w} = K_t(q_t) \cdots K_{n-1}(q_{n-1}) g$ and $D_{t,w} = V_{t,w} - \phi_t$. Any coupling of two distinct rows extends to a complete law by appending an independent third bit, so the coordinate relaxation is exactly the pairwise backup. Applying the residual identity to basis entrances gives $D_{t,w} \geq 0$ and

$$G_t(x) = \min_{w \text{ legal}} x^T D_{t,w} \geq 0. \tag{22}$$

It is monotone, positively homogeneous and superadditive, since $\min_w (x^T D_{t,w} + y^T D_{t,w}) \geq G_t(x) + G_t(y)$. Compactness and separate affinity give an original-vertex minimizer, hence

$$G_t(x) = \min_{q \in \mathcal{V}_t} \{x^T R_t(q) + G_{t+1}(K_t(q)^T x)\}. \tag{23}$$

This recursion acts on a deterministic mass distribution induced by a prefix, not an observed pair; its continuous domain is not a small quantitative state space.

B.2.2 Proof and complexity of the support criterion

Every nonempty zero face contains an original vertex. The general induction therefore gives the original-vertex recurrence: nonnegativity identifies $\operatorname{supp}(K_t(q)^T x)$ with $\operatorname{post}_t(\operatorname{supp} x, q)$, and downward closure allows the actual post-set in place of a maximal target. Any feasible vertex suffix handles empty support.

For bit accounting, let D_t be the common denominator of the three marginals times a fixed common multiple of nonzero four-column determinants of the 4×8 equality matrix. Every original vertex and kernel denominator divides D_t, and the sum of the bit lengths of the D_t is $O(B^* + n)$. If D_g is a terminal denominator, ϕ_t and original-vertex residuals use $D_g \prod_{s=t}^{n-1} D_s$. There are $64(n+1)$ support states and at most $64 \cdot 70n$ action tests; precomputed row masks give constant-dimensional post-sets. An accepting record stores a vertex index at each time. A rejecting Boolean table shows that every zero-current-residual action has a rejected successor. Regenerating the original menus verifies these implications in polynomial bit cost, without an exponential-r factor. Section A.3 supplies full-model attainment replay; support feasibility supplies optimality.

B.2.3 Proof of the quantitative support bound

For $a = K_t(q)^T \mathbf{1}_S$, put $S' = \operatorname{supp} a$ and $\beta_t(S, q) = \min_{e \in S'} a_e$, with $\beta = 0$ for empty S'. Define

$$c_n(S) = 0, \qquad c_t(S) = \min_{q \in \mathcal{V}_t} \left\{ \sum_{e \in S} R_t(q)_e + \beta_t(S, q) c_{t+1}(S') \right\}. \tag{24}$$

For distinct positive entries $0 = x_0 < x_1 < \cdots < x_m$ of x, set $S_j = \{e : x_e \geq x_j\}$ and

$$A_t(x) = \sum_j (x_j - x_{j-1}) c_t(S_j). \tag{25}$$

The $c_t(S)$ are support-dependent lower bounds, called support costs.

Theorem 7 (Quantitative support certificate)**.** *For every support, $0 \leq c_t(S) \leq G_t(\mathbf{1}_S)$ and $c_t(S) = 0$ exactly when $W_t(S)$ is true. For every $x \geq 0$, $L_t(x) \geq x^T \phi_t + A_t(x)$. The cost table and its checks have polynomial bit complexity in the present fixed-dimensional domain.*

For $x_q = K_t(q)^T x$, the lookahead bound is

$$B_t(x) = \max \left\{ x^T \phi_t + A_t(x),\ \min_{q \in \mathcal{V}_t} [x_q^T \phi_{t+1} + A_{t+1}(x_q)] \right\}. \tag{26}$$

It obeys $B_t(x) \leq L_t(x)$ and dominates ordinary one-step pairwise lookahead.

Proof. Assume $0 \leq c_{t+1}(S') \leq G_{t+1}(\mathbf{1}_{S'})$. Since $a \geq \beta \mathbf{1}_{S'}$, monotonicity and homogeneity give $G_{t+1}(a) \geq \beta G_{t+1}(\mathbf{1}_{S'}) \geq \beta c_{t+1}(S')$. Insert this into (23) at $\mathbf{1}_S$ and minimize to prove the support bound. All terms are nonnegative. When S' is nonempty, $\beta > 0$; when it is empty, the suffix contributes zero and is accepted. Thus an action has zero cost exactly when its current residual vanishes and its post-support is accepted. Induction gives $c_t(S) = 0 \iff W_t(S)$.

The identity $x = \sum_j (x_j - x_{j-1}) \mathbf{1}_{S_j}$ and superadditivity give $G_t(x) \geq \sum_j (x_j - x_{j-1}) G_t(\mathbf{1}_{S_j}) \geq A_t(x)$. No concavity or superadditivity of A_t is assumed. If $x \neq 0$ has rejected support, its first positive coefficient multiplies $c_t(\operatorname{supp} x) > 0$, detecting a strict gap.

Each lookahead candidate is at most $L_{t+1}(x_q)$. Applying it to the first law of a true optimal original-vertex word proves the upper inequality for the minimum in (26); nonnegative A_{t+1} proves one-step dominance. The maximum of the two valid bounds is valid. This uses an optimal vertex for the true objective, not a vertex-minimization claim for the nonconcave approximation. □

For bits, β has denominator dividing D_t. Induction gives $c_t(S)$ the same suffix denominator as ϕ_t; minimizing selects a rational candidate without new denominators. The bound $G_t(\mathbf{1}_S) \leq 6\|g\|_\infty$ controls numerators. Hence the 64-support table has polynomial encoding and checking cost. Lazy evaluation has the same guarantee when repeated states are memoized.

B.2.4 THE WEIGHTED-CYCLE CALCULATION

Figure 2 displays the resulting exact gap. For omission 012, the sorted row supports are $\{1,2\},\{0,2\},\{0,1\}$. At fairness, complement symmetrization preserves every pair marginal, so the four complementary complete response laws have the complete pair-kernel image of the original domain. In the established seed order $000, 100, 010, 001$, their backups on $g = (1,1,0,1,0,0)$ are

seed	$K_q g$ in $(01, 02, 10, 12, 20, 21)$	$d^T K_q g$
000	$(0, 0, 1/2, 0, 1, 1/2)$	c
100	$(1/2, 0, 1/2, 0, 1/2, 1/2)$	$(a+c)/2$
010	$(1/2, 0, 1/2, 1/2, 1, 1/2)$	$c + (a+b)/2$
001	$(0, 0, 1/2, 1/2, 1/2, 1/2)$	$(b+c)/2$

The componentwise pairwise minimum is $(0,0,1/2,0,1/2,1/2)$, and its entrance value is $c/2$. The double-permutation seed 010 is no better than the first row for $a, b \geq 0$. Taking the other three scalar minima proves (8). The formula holds also for $a+b+c \leq 1$ with the remaining mass on the absorbing diagonal. It is not a claim about arbitrary noncyclic entrances. The strict-order event is realized by $h = (1, 1/2, 0)$ and $\eta = 1/4$.

B.2.5 FINITE EXPLICIT KERNEL MENUS: A SCOPE CLARIFICATION

Let each $\mathcal{A}_t$ be an explicitly supplied finite menu of nonnegative rational $d \times d$ substochastic kernels $K_{t,a}$, with offline product choices and $g \geq 0$. Set $\phi_n = g$, $(\phi_t)_e = \min_{a \in \mathcal{A}_t} (K_{t,a}\phi_{t+1})_e$, and $R_{t,a} = K_{t,a}\phi_{t+1} - \phi_t$. The post-set recurrence gives

$$\min_{a_t, \ldots, a_{n-1}} x^T K_{t,a_t} \cdots K_{n-1,a_{n-1}} g = x^T \phi_t \quad \Longleftrightarrow \quad W_t(\operatorname{supp} x).$$

Embed the menu in its probability simplex with $K_t(q) = \sum_a q_a K_{t,a}$. Separate affinity preserves both discrete optima; a nonempty zero face contains a simplex vertex. The general induction therefore gives a word of original menu actions. Testing them costs $O(2^d d^2 \sum_t |\mathcal{A}_t|)$ operations. If D_t is a common denominator of the explicitly supplied kernels, suffix denominators divide $D_g \prod_t D_t$ and have polynomial total bit length. Substochasticity bounds values; entrance denominators add only their input length. Bit cost is the operation bound times a polynomial in the full encoding, not polynomial in variable d.

The quantitative proof transfers with six replaced by d: $a \geq \beta \mathbf{1}_{\operatorname{supp} a}$, at most d nested supports, and $G_t(\mathbf{1}_S) \leq |S| \|g\|_\infty$ give the same bounds and zero set. Taking $\mathcal{A}_t = \operatorname{vert}(Q_t)$ recovers the three-state result by the pair-extension argument above. Different actions require all relevant pair coordinates and correctly reconstructed menus; their construction cost and arbitrary-state compression are not bounded by this abstraction.

B.2.6 A TIGHT QUERY MISSED BY A FIXED GREEDY TIE RULE

This two-layer construction illustrates why the support criterion can certify a tight query that a fixed greedy tie rule misses; it does not assert failure of every tie rule. Use the complete wrapper from Appendix A.1, the strict-order terminal $g = e_{01} + e_{02} + e_{12}$, and entrance $d = (0, 1/3, 1/3, 0, 1/3, 0)$ in the six-coordinate order. The chronological layer specifications are

$$(m_0, p_0) = ((1,0,0), (2/3, 2/3, 1/3)), \qquad (m_1, p_1) = ((2,0,1), (1/3, 1/2, 1/2)).$$

The relaxed root vector is $\phi_0 = (2/9, 7/18, 0, 1/6, 0, 0)$, so $d^T \phi_0 = 7/54$. Writing δ_z for a unit mass on the original sorted-support atom z, consider

$$q_0^G = \tfrac{1}{3}(\delta_{010} + \delta_{101} + \delta_{110}), \qquad q_0^* = \tfrac{1}{3}(\delta_{000} + \delta_{110} + \delta_{111}), \qquad q_1 = \tfrac{1}{6}\delta_{000} + \tfrac{1}{2}\delta_{011} + \tfrac{1}{3}\delta_{100}.$$

These laws have all prescribed marginals. Both first laws have the same relaxed score $7/54$. Direct original-atom propagation gives

$$J(q_0^*, q_1) = 7/54 = L_{\text{pair}} = L, \qquad J(q_0^G, q_1) = 1/6.$$

Minimization over the full last-layer menu confirms that q_1 is a minimizing completion of each displayed first law. Their occupied residual sequences are $(0,0)$ and $(0, 1/27)$, respectively. The first law q_0^G reaches support $\{01, 10, 12, 20, 21\}$, which admits no zero-residual last law; q_0^* reaches $\{01, 10, 12, 21\}$, which does. Thus a rule selecting the former tied law misses exactness. The witness and pairwise lower bound prove $L = 7/54$. The exact check included in the source bundle reconstructs both original menus and verifies the displayed values, supports and last-layer minima.

C TEMPORAL COMPRESSION AND EXACT CONVEX PRUNING

Proof dependencies.

Hypothesis	Proof step	Role
Same core action; $g_{ii} = 0$	Section A.1	Equal states stay equal under a shared table; six off-diagonal coordinates suffice.
Three states; binary positive supports	Section A.2; Lemma 8	Eight response atoms, at most seventy vertices, and permutation/coalescing face repair.
D-fair	Section C.3	Complement symmetrization and seven-word fair-block reduction.
X-matched	Lemmas 10–11	Duplicate-row repair gives four scalar choices and at most two nonmetric histories.
S-arbitrary	(40); Lemma 12	One quantile law minimizes all coordinates and gives a directed path quasimetric.
$g \in \mathcal{C}_1$; retained-law closure	Lemmas 8, 13	Retained values stay in $\mathcal{C}_1$; $\mathcal{C}_0$ is absorbing under every original law.
Untied product-time domain	Sections A.2, A.3	Local replacements and legal prefix/suffix words compose with independent time noises.
r exceptional layers	Lemma 14	At most seventy face choices per exception; first-entry count $(392n + 81)490^r$.

These are the roles used in the proof. Definition 3 counts every layer outside D/X/S in r. Scalar sufficiency composes through $P^T d \geq 0$: $\forall d \geq 0\, \exists w_d$. Stored feasible words give attainment; all-entrance coverage gives optimality.

C.1 THE METRIC AND ONE-DEFECT CONTINUATION CLASSES

For $v \geq 0$ with zero diagonal, define

$$\tau_{ij}(v) = v_{i\ell} + v_{\ell j} - v_{ij}, \qquad \{i, j, \ell\} = \{0, 1, 2\}.$$

Let $\mathcal{C}_0$ be the cone on which all six slacks are nonnegative, and let $\mathcal{C}_1$ consist of values with at most one negative slack. Members of $\mathcal{C}_0$ are directed quasimetrics; symmetry is not required. The class $\mathcal{C}_1$ need not be convex.

If the only negative slack is at ab, let c be the third state and set

$$\delta = v_{ab} - v_{ac} - v_{cb} > 0, \qquad G = v - \delta e_{ab}.$$

The new entry $G_{ab} = v_{ac} + v_{cb}$ is nonnegative. Its own slack is zero. The only other decreased slacks, at ac and cb, become $v_{cb} + v_{bc}$ and $v_{ca} + v_{ac}$, respectively. Both are nonnegative, so $G \in \mathcal{C}_0$. Conversely, adding δe_{ab} to a quasimetric can decrease only its ab slack. Therefore

$$\mathcal{C}_1 = \mathcal{C}_0 \cup \bigcup_{a \neq b} \{G + \delta e_{ab} : G \in \mathcal{C}_0,\ \delta \geq 0\}. \tag{27}$$

The nineteen positive-threshold masks are zero, six single edges, six two-edge stars and six total orders. Order the states by reward: every event edge points down this order, and a two-edge chain forces its long edge because $\eta > 0$. This exhausts the possibilities, including ties and threshold equalities. Zero, stars and orders satisfy all triangle inequalities; a single edge has only its own negative slack. Thus every mask lies in $\mathcal{C}_1$.

For any deterministic common map f, an output triangle slack is

$$v_{f(i),f(\ell)} + v_{f(\ell),f(j)} - v_{f(i),f(j)}. \tag{28}$$

If $v \in \mathcal{C}_0$, this is nonnegative for every triple, including repeated images. Averaging proves that every common response law preserves $\mathcal{C}_0$. If f has image size at most two, the expression is either zero or a symmetric pair sum $v_{xy} + v_{yx} \geq 0$, even for arbitrary nonnegative zero-diagonal v. More specifically, the pullback of e_{ab} is zero or the directed cut from $f^{-1}(a)$ to $f^{-1}(b)$. A permutation pulls it back to a single edge. It follows from (27) that any law containing at most one positive permutation atom maps $\mathcal{C}_1$ into itself.

This is a property of the retained laws, not of all feasible images. For example, with omission map 012, masses $1/3$ on $(1,2,0)$ and $2/3$ on $(2,0,1)$ send e_{01} to $\frac{2}{3}e_{12}+\frac{1}{3}e_{20}$, which has two negative slacks. The feasible law with mass $1/3$ on each of $(2,0,1),(2,0,0),(1,2,1)$ has the same row marginals and sends e_{01} to $\frac{1}{3}e_{12}$.

Supplied terminal costs. All reductions below start from a nonnegative zero-diagonal continuation in $\mathcal{C}_1$ and use homogeneous comparisons and slack signs. Thus they apply to supplied costs, including $(5/2,1,1,1,1,1)$ with metric repair $(2,1,1,1,1,1)$. A metric terminal starts in the absorbing class and needs at most $7^{r+1}70^r = 7\cdot 490^r$ representatives. The upward-hull argument for pruning uses scalar sufficiency, not convexity of $\mathcal{C}_1$.

C.2 Full-domain face replacement

Write $m(i)$ for row i's omitted successor. Constant omissions allow no permutation. With two equal omissions, the duplicate rows must map to their two successors in opposite orders and the third row to the remaining state; there are exactly two permutations. With distinct omissions, source/successor relabelling gives omission 012 and the two inverse three-cycles.

Let $\mathcal{Z}_t$ be the union of the two coordinate-zero permutation faces of $\mathcal{Q}_t$, and $\mathcal{F}_t$ the union of their vertex sets. If no permutation is allowed, use $\mathcal{Z}_t=\mathcal{Q}_t$ and $\mathcal{F}_t=\mathrm{vert}(\mathcal{Q}_t)$. These are precisely the original vertices with at most one positive permutation coordinate, so $|\mathcal{F}_t|\le 70$ for the union.

Lemma 8 (Face point and scalar vertex replacement)**.** *For every $v\in\mathcal{C}_1$ and $q\in\mathcal{Q}_t$, some $q'\in\mathcal{Z}_t$ satisfies $K_t(q')v\le K_t(q)v$ coordinatewise. Every $q^*\in\mathcal{F}_t$ sends v into $\mathcal{C}_1$, and, for every $d_t\ge 0$,*

$$\min_{q\in\mathcal{Q}_t} d_t^{\mathsf{T}}K_t(q)v=\min_{q^*\in\mathcal{F}_t} d_t^{\mathsf{T}}K_t(q^*)v. \tag{29}$$

The menu $\mathcal{F}_t$ depends only on the nominal layer.

Proof. Only the case with two positive permutation masses requires a change. Remove their common smaller mass α from both atoms. For distinct omissions, use canonical omission 012 and the four complementary bit classes $000,100,010,001$, indexed $0,1,2,3$. Write $A_i=2K_i$ and $N=A_2$. Direct expansion in six-coordinate order gives

$$\begin{aligned}(N-A_0)v&=(\tau_{10},0,\tau_{01},\tau_{21},0,\tau_{12}),\\(N-A_1)v&=(0,\tau_{20},0,\tau_{21},\tau_{02},\tau_{12}),\\(N-A_3)v&=(\tau_{10},\tau_{20},\tau_{01},0,\tau_{02},0).\end{aligned}$$

Each difference omits both orientations of one unordered edge. Choose the coalescing class omitting the unique possible negative slack's edge, or any class when all slacks are nonnegative. Add mass α to each of its two complementary atoms. Removed and inserted pairs place mass α on each allowed successor of every actual row. All marginals and total mass are therefore unchanged. The old value minus the new value is $\alpha(N-A_i)v\ge 0$; this formula uses the unnormalized operators and has no missing factor of two.

For two equal omissions, write the permutation maps in the order of the two duplicate rows and the third row as (A,B,C) and (B,A,C). Replace the removed masses by mass α on (A,A,C) and (B,B,C). Every row marginal is unchanged. The sum of costs for each pair involving the third row is unchanged. Each ordered pair between the duplicate rows decreases by $\alpha(v_{AB}+v_{BA})\ge 0$. Constant omissions require no change. In each case at least one permutation mass becomes zero and no new permutation is introduced. This constructs the required feasible face point. The preservation of $\mathcal{C}_1$ follows from (27).

For a fixed $d_t\ge 0$, decompose the repaired point within its face as $q'=\sum_j \lambda_j q^{(j)}$. Linearity gives some original face vertex with

$$d_t^{\mathsf{T}}K_t(q^{(j)})v\le d_t^{\mathsf{T}}K_t(q')v\le d_t^{\mathsf{T}}K_t(q)v.$$

The reverse scalar inequality follows from $\mathcal{F}_t\subseteq\mathcal{Q}_t$. The selected vertex may depend on d_t, v and need not dominate $K_t(q')v$ coordinatewise; it is nevertheless one complete response law for all mass in d_t.

□

C.3 COMPLETE FAIR-BLOCK PROOF

The four exact pair projections. At fairness, any two response bits satisfy $\Pr(00) = \Pr(11)$ and $\Pr(01) = \Pr(10)$. Complement symmetrization $\bar{q}(z) = (q(z) + q(\mathbf{1} - z))/2$ therefore preserves the entire pair kernel. Its weights $q(b) + q(\mathbf{1} - b)$ express that kernel as a convex combination of the four laws uniform on $\{b, \mathbf{1} - b\}$, for

$$b^0 = 000, \quad b^1 = 100, \quad b^2 = 010, \quad b^3 = 001.$$

Thus the full fair pair-kernel image is their convex hull. Separate affinity makes its scalar block minima equal to minima over four-letter words. All coordinatewise word comparisons below concern those projected words, not arbitrary mixtures or equality of higher-order laws. Write $A_i = 2K_i$: displayed backups and slacks are unnormalized, and a length-L operator is $2^{-L} A_{i_0} \cdots A_{i_{L-1}}$.

Seven operators for distinct omissions. For omission 012 and $v = (a, b, c, d, e, f)$, the operators are

$$\begin{aligned} A_0 v &= (c, c+f, a, f, a+d, d), \\ A_1 v &= (d+e, e, b+f, f, b, d), \\ N v = A_2 v &= (d+e, c+f, b+f, a+e, a+d, b+c), \\ A_3 v &= (c, e, a, a+e, b, b+c). \end{aligned}$$

N is the sum of the two inverse three-cycle pullbacks. Therefore

$$N^2 = N + 2I, \qquad N^\ell = a_\ell I + b_\ell N, \quad a_\ell = \frac{2^\ell + 2(-1)^\ell}{3}, \quad b_\ell = \frac{2^\ell - (-1)^\ell}{3}. \tag{30}$$

For $i \in \{0, 1, 3\}$ and every $j \in \{0, 1, 2, 3\}$,

$$A_i^3 = A_i, \qquad A_j A_i \geq A_i^2 \quad \text{entrywise.} \tag{31}$$

Indeed, $A_0^2 v = (a, a+d, c, d, c+f, f)$ and $A_0^3 v = A_0 v$. For $j = 0, 1, 2, 3$, the respective differences $(A_j A_0 - A_0^2)v$ are

$$\begin{gathered} (0, 0, 0, 0, 0, 0), \quad (d+f, 0, d+f, 0, 0, 0), \\ (d+f, 0, d+f, a+c, 0, a+c), \quad (0, 0, 0, a+c, 0, a+c). \end{gathered}$$

They have nonnegative coefficients. Simultaneous source/successor permutations fix N and transitively permute the coalescers, proving all cases. Right multiplication by $A_i \geq 0$ also gives $A_j A_i^2 \geq A_i$.

For $L \geq 2$, retain the seven actual words

$$\{A_i^L : i = 0, 1, 3\} \cup \{A_i^{L-1} N : i = 0, 1, 3\} \cup \{N^L\}. \tag{32}$$

An all-N word is retained. Otherwise write the word $U A_i N^\ell$ using its rightmost coalescer. Successive application of (31) gives $U A_i \geq A_i^{L-\ell}$. For $\ell = 0, 1$ the result is already in (32). If $\ell \geq 2$ is even, $a_\ell \geq 1, b_\ell \geq 0$ give $A_i^{L-\ell} N^\ell \geq A_i^{L-\ell} = A_i^L$. If $\ell \geq 3$ is odd, $b_\ell \geq 1, a_\ell \geq 0$ give a value no smaller than $A_i^{L-\ell} N = A_i^{L-1} N$. The equalities use $A_i^3 = A_i$ and parity. Every replacement has length L; for $L = 1$ use the four letters, and for $L = 0$ the identity.

If the omissions m_t are arbitrary permutations, choose boundary relabelings recursively by

$$\sigma_L = \mathrm{id}, \qquad \sigma_t = \sigma_{t+1} \circ m_t. \tag{33}$$

The transformed omission $\sigma_{t+1} \circ m_t \circ \sigma_t^{-1}$ is the identity. For each canonical atom $\widehat{f}_t$, recover its actual response by

$$f_t(i) = \sigma_{t+1}^{-1}\big(\widehat{f}_t(\sigma_t(i))\big). \tag{34}$$

Its position in the actual sorted support gives the response bit. Relabelling is bijective on each support, so complementary atoms remain complementary. These fixed boundary permutations transport the seven words; they are not additional actions.

A common quantile law for directed path costs. The shared inverse-transform construction is standard (Fill & Machida, 2002, Section 1.2, equation (1.5)). We give the directed-cut calculation needed for its use in the metric-block proof. For an ordering (u, v, w), a directed path metric is the nonnegative sum of directed crossing costs of the nested cuts $\{u\}$ and $\{u, v\}$. If a cut C has outward and inward costs $r_C, s_C \geq 0$, any common response law with row marginals μ_i has expected cut cost at least

$$r_C\big(\mu_i(C) - \mu_j(C)\big)_+ + s_C\big(\mu_j(C) - \mu_i(C)\big)_+. \tag{35}$$

For $X = \mathbf{1}\{F(i) \in C\}$ and $Y = \mathbf{1}\{F(j) \in C\}$, the difference of the two crossing probabilities is $\Pr(X = 1, Y = 0) - \Pr(X = 0, Y = 1) = \mu_i(C) - \mu_j(C)$, proving the bound. Generate all rows by quantiles of one uniform variable in the order (u, v, w). Each nested-cut event is an initial interval of length $\mu_i(C)$, so this single law attains (35) simultaneously for both cuts and every ordered pair. It therefore minimizes the path-cost value-vector backup coordinatewise.

At a fair binary ternary row with omission $m(i)$,

$$\mu_i(C) - \mu_j(C) = \tfrac{1}{2}\big(\mathbf{1}\{m(j) \in C\} - \mathbf{1}\{m(i) \in C\}\big).$$

Summing the directed cut costs proves the normalized formula

$$[K_{mquantile} v]_{ij} = \tfrac{1}{2} v_{m(j),m(i)}. \tag{36}$$

The law uses the two complementary complete maps selected on the lower and upper halves of the uniform variable. Its image is a relabelled directed path metric when m is bijective, a directed cut metric when m has image size two, and zero when m is constant. Directed cuts, including their zero degeneracies, are directed path metrics.

Seven completions on the entire metric cone. For canonical two-equal omission 001, write the four unnormalized images, in some order of the four complementary classes, as

$$P = (0, \alpha, 0, \alpha, \beta, \beta), \quad Q = (0, \gamma, 0, \gamma, \delta, \delta),$$
$$R = (s, \gamma, s, \alpha, \delta, \beta), \quad S = (s, \alpha, s, \gamma, \beta, \delta),$$

where $\alpha = v_{10}$, $\beta = v_{01}$, $\gamma = v_{12} + v_{20}$, $\delta = v_{21} + v_{02}$, and $s = v_{12} + v_{21}$. For $v \in \mathcal{C}_0$ the two relevant triangle inequalities give $\alpha \leq \gamma$ and $\beta \leq \delta$. Thus P is componentwise no larger than any of the four images and is an attained directed cut metric. In actual labels its common response law synchronizes the duplicate rows and aligns their common successor with the third row. Independent source and successor permutations prove this statement for every two-equal omission. Constant omissions permit the two constant maps and hence the zero backup.

The metric-block construction is proved once, in Lemma 12: it uses the seven D words, the common X/S minimizing law and the quantile identities above. At fairness all these laws are projected four-letter laws, and each replacement is coordinatewise no larger for the supplied continuation. Thus its seven outputs also give wordwise coverage of the projected fair menu.

If the chronological omission composition $M = m_{L-1} \circ \cdots \circ m_0$ is constant, apply the quantile construction to the strictly positive off-diagonal distances on 0–1–2. Iterating (36) gives its pullback by M, transposed for odd L and scaled by 2^{-L}, hence zero. A nonnegative operator annihilating this positive six-vector is the zero operator. The same attained word therefore annihilates every zero-diagonal continuation.

The two-equal formulas with their normalization. For $v = (a, b, c, d, e, f)$ set $s_{01} = a + c$, $s_{02} = b + e$, $s_{12} = d + f$. In omission 010, the four actual classes give

$$A_0 v = (c, 0, a, a, 0, c),$$
$$A_1 v = (d + e, d + f, b + f, a, d + f, c),$$
$$A_2 v = (d + e, 0, b + f, b + f, 0, d + e),$$
$$A_3 v = (c, d + f, a, b + f, d + f, d + e).$$

Classes $0, 2$ consist only of maps with image size at most two and yield cut metrics. Subtraction and the full slack vectors give

$$(A_1 - A_0)v = (\tau_{10}, s_{12}, \tau_{01}, 0, s_{12}, 0),$$
$$(A_3 - A_0)v = (0, s_{12}, 0, \tau_{01}, s_{12}, \tau_{10}),$$
$$\tau(A_1 v) = (\tau_{20}, \tau_{21}, \tau_{02}, \tau_{01} + s_{12}, \tau_{12}, \tau_{10} + s_{12}),$$
$$\tau(A_3 v) = (\tau_{10} + s_{12}, \tau_{12}, \tau_{01} + s_{12}, \tau_{02}, \tau_{21}, \tau_{20}).$$

Values and slacks are normalized by dividing these expressions by two. Unless the unique bad-edge support is $01/10$, the coalescing child $A_0 v$ is no larger than either potentially nonmetric child. Otherwise each mixed child has at most one negative slack, and the two possible surviving edges lie on different unordered pairs. This local calculation, with actual boundary relabelling, is the basis of Lemma 11; its proof uses these identities, not the fair-block statement below.

Fair-block construction. Starting from one $v \in \mathcal{C}_1$, form the four children backward. Delete only a child with a same-parent coordinatewise no-larger replacement. Complete each metric child through the remaining prefix by the metric-block construction and keep only nonmetric children pending. Lemma 11 gives at most two pending histories under arbitrary actual boundary changes. At D-fair, the face/slack identities remove N in favor of an attained metric coalescer; at constant fair support the zero law is a projected letter. Thus neither type carries a nonmetric history past that layer. The fair specialization has only wordwise coordinatewise deletions; it does not rely on selecting a vertex that dominates a mixture.

Lemma 9 (Fair-block representatives)**.** *For any length-L fair block $\mathcal{B}$ and $v \in \mathcal{C}_1$, the preceding construction returns at most $56L + 7$ attained normalized value-vector backups, all in $\mathcal{C}_1$, with at most two outside $\mathcal{C}_0$. Every four-class word w has a returned value u with $u \le K_w v$ coordinatewise. Hence, for every $d \ge 0$, the returned family has the same minimum of $d^{\mathsf{T}} u$ as the full fair law domain. For $v \in \mathcal{C}_0$, seven representatives suffice. Empty blocks return the identity word.*

Proof. At most two pending parents, four children each and seven metric-prefix completions give at most $56L$ metric outputs and two final nonmetric values for $L > 0$. The metric-input and empty-block cases also satisfy $56L + 7$. Each local deletion is coordinatewise valid, and nonnegative earlier kernels preserve it. The metric compiler supplies a no-larger replacement for every earlier projected word, proving the stated wordwise coverage. Feasibility gives the reverse scalar bound; the four-law convex-hull identity extends scalar equality to the full fair domain, without asserting domination of every mixture by a single word. □

Actual-label implementation forms the four children and tests their slacks and same-parent domination; Lemma 11 proves the pending bound rather than imposing a numerical cap. Each metric-prefix call, including word assembly, costs $O(L^2)$ rational operations. There are $O(L)$ calls and fixed-dimensional pending updates, for $O(L^3)$ operations and $O(L^2)$ explicit law records.

A fair-mixture quantifier regression. The directed-quasimetric fair-block proof uses coordinatewise replacement for words in the four-letter projected menu. Arbitrary original fair laws can project to mixtures, and the full-domain conclusion is equality of scalar lower envelopes. For example, the original fair vertex assigning masses $1/4$ to $001, 010, 100, 111$ sends the strict-order terminal to $(1/4, 0, 1/2, 1/4, 3/4, 1/2)$. None of the four pure projected backups dominates it coordinatewise. A convex combination certifies coverage, consistently with the scalar theorem. The exact-algebra check included in the source bundle verifies this distinction.

C.4 Refined temporal compression on matched layers

We now extend the fair construction to the compressible layers of Definition 3, retaining the same original domains, six coordinates, continuation classes and probability scale.

C.4.1 Response-law repair at a matched X layer

Let i, k be the duplicate-support rows and j the singleton row. Their supports are $\{c, u\}$, $\{c, u\}$, and $\{c, v\}$, where c, u, v are distinct. Let U, V indicate use of u by the duplicate rows and let W indicate use of v by the singleton row. These are semantic indicators; conversion to sorted-support bits is part of the original-label decoding. Write

$$\mathbb{E}U = \mathbb{E}V = a, \qquad \mathbb{E}W = b, \qquad 0 < a, b < 1,$$

and set

$$\ell_X = \max\{0, a - b\}, \qquad h_X = \min\{a, 1 - b\},$$
$$w_X = h_X - \ell_X = \min(a, b, 1 - a, 1 - b) > 0, \qquad \lambda = 2w_X \le 1.$$

In semantic order (U, W, V), define a subprobability law ρ by

z	$(0,0,0)$	$(1,0,1)$	$(0,1,0)$	$(1,1,1)$
ρ_z	$1-b-h_X$	ℓ_X	$b-a+\ell_X$	$a-h_X.$

Every other mass is zero. The four displayed masses are nonnegative, sum to $1-\lambda$, and synchronize the duplicate rows in every atom. Consequently each associated complete map has image size at most two. If $q_0^{\mathrm{fair}}, \dots, q_3^{\mathrm{fair}}$ are the four complementary fair laws on these same supports, define

$$q_s = \rho + \lambda q_s^{\mathrm{fair}}, \qquad 0 \le s < 4. \tag{37}$$

Each q_s is a normalized original response law with the required marginals.

Lemma 10 (Four-law lower-envelope reduction)**.** *For every $q \in Q_t$, there is $q' \in \mathrm{conv}\{q_0, q_1, q_2, q_3\}$ such that $K_t(q') \le K_t(q)$ entrywise on the six non-diagonal coordinates. Hence, for every $v \ge 0$ with zero diagonal and every nonnegative entrance $d \ge 0$,*

$$\min_{q \in Q_t} d^T K_t(q) v = \min_{0 \le s < 4} d^T K_t(q_s) v. \tag{38}$$

Proof. Keep the complete pair laws (U, W) and (V, W) fixed. For $y \in \{0, 1\}$, put $M_y = \Pr(W = y)$, $A_y = \Pr(U = 1, W = y)$, and $B_y = \Pr(V = 1, W = y)$. Replace the conditional coupling of U, V by the slice

$$q'(1,y,1) = \min(A_y, B_y), \quad q'(1,y,0) = (A_y - B_y)_+,$$
$$q'(0,y,1) = (B_y - A_y)_+, \quad q'(0,y,0) = M_y - \max(A_y, B_y).$$

This requires no division, preserves all three marginals and the two pair laws involving the singleton row, and minimizes both off-diagonal disagreement probabilities of the duplicate rows conditional on W. The only changed non-diagonal pair-kernel entries therefore decrease. Mass may move to the absorbing diagonal, whose continuation is zero; no entrance is renormalized.

Put $x = \Pr(U = 1, W = 0)$ and $y = \Pr(V = 1, W = 0)$. Both lie in $[\ell_X, h_X]$. The repaired law is affine on each triangle of this square. When $x \ge y$, it is a convex combination of the corner laws $(\ell_X, \ell_X), (h_X, \ell_X), (h_X, h_X)$ with weights

$$(h_X - x)/w_X, \quad (x - y)/w_X, \quad (y - \ell_X)/w_X.$$

For $x \le y$, use corners $(\ell_X, \ell_X), (\ell_X, h_X), (h_X, h_X)$ and weights $(h_X - y)/w_X, (y - x)/w_X, (x - \ell_X)/w_X$. Substitution identifies these four corners with (37). For a fixed scalar objective, some corner is no worse than the mixture. All corners are original-feasible, proving the reverse inequality in (38). □

The repaired mixture supplies the entrywise statement. A single selected corner supplies the scalar statement at the given entrance distribution; it need not entrywise dominate every original response law. Nor does this identify the full higher-order law polytope with a four-vertex polytope.

C.4.2 Arbitrary-boundary nonmetric survival

Extend the six-coordinate kernel formula K_t linearly to subprobability laws, without renormalization. Let $\mathsf{G}_t = K_t(\rho)$ and $K_s^{\mathrm{fair}} = K_t(q_s^{\mathrm{fair}})$. The same subprobability law occurs in all four menus, so

$$K_t(q_s) = \mathsf{G}_t + \lambda K_s^{\mathrm{fair}}, \qquad K_t(q_s) - K_t(q_{s'}) = \lambda(K_s^{\mathrm{fair}} - K_{s'}^{\mathrm{fair}}). \tag{39}$$

Pullbacks by maps with image size at most two send every nonnegative zero-diagonal value into $\mathcal{C}_0$. Thus $\mathsf{G}_t v \in C_0$: it cannot add a negative triangle slack. The same-parent comparisons of the fair X formulas retain their directions. The coalescing letters remain metric, and each mixed letter still contains at most one positive permutation atom, so its image of $\mathcal{C}_1$ remains in $\mathcal{C}_1$.

Use the omission-010 identities in Section C.3, in fair action order $000, 100, 010, 001$. In addition to the two displayed mixed-child differences there,

$$(A_2 - A_0)v = (\tau_{10}, 0, \tau_{01}, \tau_{01}, 0, \tau_{10}).$$

Unless the unique bad-edge support is $\{0, 1\}$, child 0 is metric and no larger than all children. For parent edge 01, mixed children $1, 3$ can have bad edges $12, 10$; for parent 10, they can have $21, 01$. These are different unordered pairs. The common $\mathsf{G}_t v$ preserves the comparisons and can only remove negative slacks. Transport is performed separately at every actual source/successor boundary, with no globally fixed support sequence.

Lemma 11 (Two nonmetric survivors under actual boundary changes)**.** *Starting from one continuation in $\mathcal{C}_1$, a run of X-matched layers can be compiled while retaining at most two active nonmetric histories at every actual boundary, with no bound on the number of support or probability changes.*

Proof. Use the canonical fair-X child and slack identities with their actual source/successor label transport. A parent can have a retained nonmetric child only when the unordered support of its unique bad edge equals the pair of distinct omitted successor labels at this layer. Its two possible nonmetric children have bad edges on the two different source pairs joining the singleton row to one duplicate row. Equation (39) preserves each comparison, while its common quasimetric term can only remove negative slacks. If two nonmetric parents are pending, their unordered bad-edge supports differ. The next layer has only one required unordered pair, so at most one parent can have nonmetric descendants, and it has at most two. Induction proves the claim. Zero slacks terminate a nonmetric history; metric children are retained through their metric completions, not replaced by zero. □

A D-fair layer removes the double-permutation image by the original one-defect face/slack comparison, leaving only metric coalescing images. An S-arbitrary layer has a metric image by the following construction. Thus neither can carry a pending nonmetric history past that layer.

C.4.3 Arbitrary common-support probabilities and seven completions

At an S-arbitrary layer with ordered support $\{u, v\}$, write p_i for $\Pr(F(i) = v)$. Couple all row indicators by one uniform variable. Its non-diagonal output probabilities are simultaneously

$$\Pr(F(i) = v, F(j) = u) = (p_i - p_j)_+, \qquad \Pr(F(i) = u, F(j) = v) = (p_j - p_i)_+.$$

These are the pairwise lower bounds, attained by the same complete response law. For any nonnegative zero-diagonal continuation,

$$V'_{ij} = (p_i - p_j)_+ V_{vu} + (p_j - p_i)_+ V_{uv}. \tag{40}$$

In the order of the three p_i, this is a directed path quasimetric; it need not be zero. The quantile construction is a standard component, not an additional generic optimization principle.

At X-matched with an input in $\mathcal{C}_0$, the fair minimizing coalescer stays componentwise minimal by (39). Each of its atoms, and each atom of ρ, synchronizes the duplicate rows. The resulting value is a directed cut on this same row partition.

Lemma 12 (Metric-block closure)**.** *For any length of compressible block and any input in $\mathcal{C}_0$, at most seven attained complete-word values suffice for minimization at every nonnegative entrance.*

Proof. If all layers are D-fair, use the original seven D words with their actual boundary transport. Otherwise locate the rightmost non-D layer. Its D-fair suffix has at most seven representatives, each in $\mathcal{C}_0$. At the non-D layer, apply the common minimizing X or S law just described, obtaining a directed path quasimetric. For each earlier layer use the common quantile minimizer for the two nested cuts of that path. A D-fair layer gives the original scaled, relabelled or transposed path image. Identical duplicate-row marginals at X-matched force duplicate quantile responses, giving a directed cut. At S-arbitrary, (40) gives a directed path. Thus every earlier update remains within this domain and is a simultaneous coordinatewise minimum. The seven suffix representatives yield at most seven complete words. □

Lemma 13 (Compressible-block representatives)**.** *A length-L compressible block applied to one $v \in C_1$ has an attained all-entrance sufficient family of at most $56L + 7$ values in $\mathcal{C}_1$, at most two outside $\mathcal{C}_0$. An empty block is the identity.*

Proof. At each position there are at most two nonmetric parents, at most four retained local choices, and at most seven completions of each new metric child through the remaining prefix. Immediately invoke Lemma 12 for such a child. The resulting count is bounded by $56L + 7$, including the final pending values and the metric-input and empty-block cases. The original face/slack step, the repaired X-domain scalar identity, and the common S minimizer preserve the full nonnegative entrance lower envelope. Nonnegative prefix operators permit their use at every earlier boundary. Actual chronological word provenance is retained for every candidate, providing the reverse bound by feasibility. □

Why scalar sufficiency composes. Let H be a sufficient suffix family and let P be any fixed legal prefix operator ending at its boundary. For every entrance distribution $a \geq 0$, nonnegativity gives $P^T a \geq 0$, so

$$\min_{h \in H} a^T P h = \min_{w \text{ legal suffix}} a^T P V_w.$$

Taking the minimum over legal prefixes on both sides preserves this equality. Thus a repaired mixture may select a corner using the single induced entrance $P^T a$, and backward induction may use a different complete word for a different entrance. The quantifiers are $\forall a \geq 0\, \exists w_a$, not $\exists w\, \forall a \geq 0$. For any fixed entrance, all coordinates and random images still use the same selected word. No coordinatewise corner domination is needed.

C.4.4 First-entry counting, exactness, and bit cost

Let $\ell_0, \ldots, \ell_r$ be the lengths of the compressible blocks separated by the r exceptional times.

Lemma 14 (First-entry counting)**.** *Metricity is absorbing under every original response law. Classifying a generated history by its first backward entry into $\mathcal{C}_0$ yields*

$$\begin{aligned} N_{\mathrm{rep}} &\leq 7 \sum_{j=0}^{r} (56\ell_{r-j} + 8) 140^j 490^{r-j} + 2 \cdot 140^r \\ &\leq (392n + 81) 490^r. \end{aligned} \tag{41}$$

The family can be generated in this form without first enumerating all histories.

Proof. At each exceptional layer retain the at-most-seventy original face vertices, preserving $\mathcal{C}_1$ and scalar sufficiency. After j exceptions, at most 140^j nonmetric histories enter a block: each intervening block has two survivors and each exception seventy choices. The first metric entry in the length-ℓ_{r-j} block gives at most $56\ell_{r-j} + 8$ possibilities per such history; the remaining metric prefix has at most $7 \cdot 490^{r-j}$ completions. Metricity is absorbing, so each history pays the length factor only once. At most $2 \cdot 140^r$ histories stay nonmetric throughout. This proves the first sum in (41). Since $\sum_i \ell_i \leq n$,

$$\frac{N_{\mathrm{rep}}}{490^r} \leq 392 \sum_{j=0}^{r} \ell_{r-j} (2/7)^j + 56 \sum_{j=0}^{r} (2/7)^j + 2(2/7)^r \leq 392n + 81. \tag{42}$$

Generate a history only until its first metric entry, then invoke the metric-prefix construction. This implements the count without enumerating all words. Local scalar substitutions compose by the preceding argument; stored feasible words give attainment and the reverse bound. The result bounds a sufficient family, not the set of all feasible values. □

Intermediate records and bit cost. For nonmetric and metric record counts U, M before optional cross-record deletion, a length-ℓ block and one exception give

$$U' \leq 2U, \qquad M' \leq 7M + (56\ell + 7)U, \tag{43}$$

$$U' \leq 70U, \qquad M' \leq 70M + 70U. \tag{44}$$

First entries at exceptional layers are included in the $70U$ term of (44); after entering the absorbing class $\mathcal{C}_0$, such records use MetricPrefix to complete their remaining prefixes. There is no $M \to U$ term. Consequently the same first-entry count bounds every unpruned chunk boundary by $R = O(490^r(n+1))$; deletion only reduces it. For completeness, index blocks $b_0, \ldots, b_r$ in backward order. For a nonmetric terminal, $U_0 = 1, M_0 = 0$ and the block–exception recurrence is

$$U_{j+1} \leq 140 U_j, \qquad M_{j+1} \leq 490 M_j + (3920 b_j + 630) U_j.$$

Thus $U_j \leq 140^j$ and $M_r \leq \sum_{j=0}^{r-1} 490^{r-1-j} (3920 b_j + 630) 140^j$. The final block gives the sharper intermediate estimate used by this implementation,

$$N_{\mathrm{rep}} \leq 7M_r + (56b_r + 9)140^r \leq 490^r \sum_{j=0}^{r} (56b_j + 9)(2/7)^j \leq 490^r (56n + 63/5).$$

A metric terminal needs at most $7 \cdot 490^r$ values. Empty blocks and $r = 0$ are included in both counts.

Recognition uses exact supports and marginals. Exceptional menus solve fixed-size rational systems; matched laws use fixed-size arithmetic and quantile laws a constant number of breakpoints. Let $G = \max g$. Normalized suffix coordinates lie in $[0, G]$. A common denominator D is the product of terminal denominators and the fixed-size local-law denominators, so $\log D = O(B^* + n)$. Numerators are bounded by DG and have $O(B^* + n)$ bits as well. Doubled-operator intermediates add only $O(n)$ bits. This charges terminal magnitude in binary, with no inverse-probability or magnitude-sized loop. Metric-prefix construction costs $O(L^3)$ rational/word operations per block. Since $\sum_i \ell_i^3 \le n^3$ and exceptional stages add $O(rRn)$, the total pre-deletion records and primitive word operations are $Q = O(490^r(n+1)^4)$. The unpruned first-entry construction therefore has cost $O(490^r(n+1)^4(B^* + n + 1)^3) + \mathrm{poly}(B^*)$.

Put $N = n + B^* + 1$. Charging all exact duplicate, domination and coverage comparisons by $O(Q^2)$ and rational work by $O(N^3)$ gives $O(490^{2r}N^{11})$. A further conservative polynomial factor covers serialization, backpointers, parser checks, certificates and complete-table output, yielding $O(490^{2r}N^{12})$ with degree independent of r. No favorable pruning or product of block lengths is assumed.

Restore the selector, routing and deterministic unused rows as in Section A.3. A selected law need not be a vertex or a fair two-atom law, but has an eight-entry encoding. Each time's complete table is shared across worlds; noises remain independent across times. Mixtures in the proof are not correlated random draws of entire words. Nominal-row/joint-world replay verifies attainment; lower-envelope coverage proves optimality.

C.4.5 Constructive interface

For $v \in \mathcal{C}_0$, MetricPrefix$([0,t), v)$ processes the entire recognized prefix backward through its compressible blocks and exceptional layers. On each block it applies Lemma 12 to every current continuation; at each exceptional layer it backs up every current continuation through all retained face vertices of Lemma 8. Every original response law preserves $\mathcal{C}_0$, so all intermediate continuations remain metric. With e exceptional layers and at most $e+1$ possibly empty blocks, it returns at most $7^{e+1}70^e = 7 \cdot 490^e$ attained values. Each returned pair (u, p) retains the complete original chronological response-law word p, with $u = K_p v$; scalar sufficiency composes as above for every nonnegative entrance. The empty prefix returns $\{(v, ())\}$, the identity. This realizes the metric branch already counted in Lemma 14.

Algorithm 2 processes the recognized blocks and exceptional layers backward, retaining laws and suffix pointers with every value. Optional deletion requires an attained coordinatewise no-larger replacement, with all necessary coverage records saved. Minimizing against the original unnormalized off-diagonal entrance and decoding one word gives the full SCM; neither pairwise equality nor exhaustive word enumeration is needed. The certificate obligations and available verification material are indexed in Appendices G and J.

C.4.6 The two-survivor mechanism is reachable from a threshold

The following retained construction checks the interface against a legal terminal, rather than an arbitrary input in $\mathcal{C}_1$. Fix rational $0 < \theta < 1/6$, rewards $(1, 0, 1/2)$ and threshold $3/4$, so $g = e_{01}$. The last two layers have chronological specifications

$$(m, p) = ((1,2,1), (1/5, 2/5, 3/5)), \qquad ((1,0,0), (1/4, 1/4, 3/4 - \theta)).$$

At the last layer only row 0 can output terminal state 0. Its two nonzero pairwise minima are $1/2$ and θ, simultaneously attained by

$$q_{\mathrm{last}} = \theta\delta_{000} + (1/2 - \theta)\delta_{001} + \tfrac{1}{4}\delta_{011} + \tfrac{1}{4}\delta_{100}.$$

Here δ_z is the unit mass on a sorted-support atom. The resulting value is $u_\theta = (1/2, \theta, 0, 0, 0, 0)$. At the preceding layer, aligning the three nonzero-output indicators with one uniform variable minimizes their directed outgoing costs simultaneously. The law

$$q_{\mathrm{middle}} = \tfrac{2}{5}\delta_{000} + \tfrac{1}{5}\delta_{001} + \tfrac{1}{5}\delta_{011} + \tfrac{1}{5}\delta_{111}$$

Algorithm 2 Sharp lower bound with compressible blocks

Parse the original table; use six off-diagonal coordinates without renormalizing d.
Form Q_t, g; recognize compressible blocks and exceptional layers.
At each exceptional layer, enumerate the face vertices $\mathcal{F}_t$ (Lemma 8).
Initialize pending records $P \leftarrow \{(n, g, ())\}$ and complete records $F \leftarrow \varnothing$.
while $P \neq \varnothing$:
  Remove a record (t, v, w), where w is one complete suffix word.
  if $v \in \mathcal{C}_0$:
    for $(u, p) \in$ MetricPrefix$([0, t), v)$:
      Add (u, pw) to F; the empty prefix is the identity.
  else if $t = 0$: add (v, w) to F.
  else:
    Let $[s, t)$ be the rightmost remaining chunk.
    if this is an exceptional layer:
      $H \leftarrow \{(K_s(q)v, (q)) : q \in \mathcal{F}_s\}$.
    else: $H \leftarrow$ CompressC1$([s, t), v)$ (Lemma 13).
    for each attained continuation vector $(u, b) \in H$:
      Add (s, u, bw) to P, keeping value and word together.
end while
Select $(v^*, w^*) \in \arg\min_{(v,w)\in F} d^T v$.
Decode w^* into original-time laws and complete every nominal row.
return $d^T v^*$ and the resulting SCM with independent time noises.

attains $v_\theta = (1/10, 2\theta/5, 0, \theta/5, 0, 0)$. Its ij cost is the weight of row j times $(p_j - p_i)_+$. Nonnegative earlier kernels preserve both coordinatewise replacements; v_θ is therefore a simultaneous minimum of this complete two-layer suffix.

Prepend an X-matched layer with omissions 010 and probabilities $(1/3, 1/3, 1/3)$. Its four complete response laws in Lemma 10 are $\delta_{000}/3 + (2/3)q_s^{\mathrm{fair}}$, and their backups are

$$\begin{aligned} u_0 &= (0, 0, 1/15, 1/15, 0, 0), \\ u_1 &= (\theta/15, \theta/15, 1/30 + 2\theta/15, 1/15, \theta/15, 0), \\ u_2 &= (\theta/15, 0, 1/30 + 2\theta/15, 1/30 + 2\theta/15, 0, \theta/15), \\ u_3 &= (0, \theta/15, 1/15, 1/30 + 2\theta/15, \theta/15, \theta/15). \end{aligned}$$

The first and third are metric. Each of u_1, u_3 has exactly one negative slack:

$$\tau_{12}(u_1) = \tau_{10}(u_3) = \theta/5 - 1/30 < 0.$$

Their other slacks are nonnegative. The two nonmetric children are incomparable; u_2 is larger on coordinate 21 than u_1 and on 01 than u_3, and u_0 is larger on 10 or 12, respectively. Neither nonmetric child has a sibling that dominates it coordinatewise.

Prepend a D-fair 012 layer. Its retained coalescers send both histories into $\mathcal{C}_0$, after the face lemma removes the double-permutation choice. All earlier common kernels preserve metricity. Additional D-fair prefixes extend the family to arbitrary horizons with $k = 3, r = 2$. This is not a claim that every globally minimal convex representation must contain both histories.

C.5 Refined-parameter guarantee for exact convex pruning

Value-vector backup and witness-region pruning follow the classical POMDP framework (Smallwood & Sondik, 1973; Cassandra et al., 1997; Feng & Zilberstein, 2004). The upward-hull argument below uses ordinary separating-hyperplane geometry (Boyd & Vandenberghe, 2004, Section 2.5) to transfer the structural cardinality bound to that framework.

Corollary 15 (Full-cone generic pruning in the refined parameter)**.** *Consider per-layer backups over all original response vertices, followed by deduplication and sequential exact irredundant convex pruning on the entire nonnegative entrance cone. At a suffix with n_t core times and r_t layers outside Definition 3, the retained cardinality is at most*

$$(392n_t + 81)490^{r_t},$$

and every retained vector lies in $\mathcal{C}_1$. With a polynomial-bit exact linear-programming subroutine, this algorithm has cost $490^{O(r)}\,\mathrm{poly}(n+B^*)$ and recovers an attaining original-vertex common word.

Proof. Let X_t be the finite set of all values of original-vertex suffix words, and let H_t be the attained sufficient family from Theorem 4, applied to this suffix. Its members lie in $\mathcal{C}_1$ and satisfy the stated refined cardinality bound. The terminal boundary is the singleton $\{g\}$. We do not assume $H_t \subseteq X_t$: a refined representative law can be a nonvertex.

For any legal suffix law sequence, decompose each local law as

$$q_s = \sum_{\nu \in \mathrm{vert}(Q_s)} \lambda_{s,\nu}\nu, \qquad \lambda_{s,\nu} \geq 0, \qquad \sum_{\nu} \lambda_{s,\nu} = 1.$$

Linearity of each K_s and expansion of the time product give

$$\begin{aligned} &K_t(q_t)\cdots K_{n-1}(q_{n-1})g \\ &\quad = \sum_{\substack{\nu_s \in \mathrm{vert}(Q_s) \\ t \leq s < n}} \left(\prod_{s=t}^{n-1} \lambda_{s,\nu_s} \right) K_t(\nu_t)\cdots K_{n-1}(\nu_{n-1})g. \end{aligned} \tag{45}$$

The coefficients are nonnegative and sum to one. Consequently every legal suffix value, including every member of H_t, belongs to $\mathrm{conv}(X_t)$. Theorem 4's all-nonnegative-entrance sufficiency therefore implies

$$H_t \subseteq \mathrm{conv}(X_t), \qquad \min_{x \in X_t} a^T x = \min_{h \in H_t} a^T h \quad (a \geq 0).$$

Set $P_t = \mathrm{conv}(X_t) + \mathbb{R}^6_+$. The inclusion above gives one direction of

$$P_t = \mathrm{conv}(H_t) + \mathbb{R}^6_+. \tag{46}$$

For the other, suppose $x \in P_t$ lies outside the right-hand closed polyhedron. Strict separation from that upward polyhedron supplies a functional $a \geq 0$: a negative coordinate would make its infimum equal to $-\infty$. It would give $a^T x < \min_{h \in H_t} a^T h = \min_{y \in X_t} a^T y$, contradicting $x \in P_t$.

Every vertex of $\mathrm{conv}(H_t) + \mathbb{R}^6_+$ belongs to H_t. A nonzero added orthant vector gives a nontrivial midpoint decomposition, and a convex combination of distinct generators contradicts extremality. The same argument applied to X_t shows that every vertex also belongs to X_t. Hence

$$\mathrm{vert}(P_t) \subseteq H_t \cap X_t \subseteq \mathcal{C}_1, \qquad |\,\mathrm{vert}(P_t)| \leq |H_t| \leq (392 n_t + 81)490^{r_t}.$$

This conclusion does not require $\mathcal{C}_1$ to be convex.

Sequential exact irredundant pruning retains precisely these vertices. For the current deduplicated finite set X, a candidate x is removable exactly when $x \in \mathrm{conv}(X \setminus \{x\}) + \mathbb{R}^6_+$. Such a deletion preserves the upward hull. This polyhedron is pointed with recession cone $\mathbb{R}^6_+$, so every nonvertex is generated by its vertices and a nonnegative vector; a vertex cannot be deleted. Equivalently, for nonempty competitors y, solve

$$\max \delta \quad \text{subject to} \quad a \geq 0, \quad \sum_{i \neq j} a_{ij} = 1, \quad a^T(y - x) \geq \delta \quad \text{for every } y \neq x.$$

A positive optimum strictly exposes a necessary vertex; otherwise the candidate is redundant. An isolated candidate is retained. Duplicate values keep one original-word provenance. Deletion certificates use only currently retained candidates and so have no circular dependencies.

At a new layer, the raw backup set has at most seventy times the preceding retained cardinality. Nonnegative common kernels preserve the all-entrance lower-envelope substitution, so these local backups reproduce P_t inductively. The refined bound controls every retained frontier and hence every raw set. Original-vertex words have $O(B^* + n)$-bit rational continuation coordinates. Polynomial-bit exact LP, at most one sequential test per raw candidate, and polynomial-size word pointers and output give $490^{O(r)}\,\mathrm{poly}(n + B^*)$ bit time. A necessary candidate cannot become redundant when competitors are deleted, so tests need not be repeated. Every surviving vector still carries one original-vertex word, yielding an attaining SCM with independent time noises. □

Equation (45) is an identity in value space, not a prescription to implement a correlated random choice of an entire time word. The generic algorithm's witness is one retained original-vertex word. The result requires pruning on the full fixed nonnegative entrance cone and a polynomial-time exact LP subroutine. Checking a floating-point proposal exactly, or using a rational-simplex fallback, does not by itself give that particular backend a worst-case polynomial guarantee. The nominal-k cardinality bound follows as a weaker consequence of $r_t \le k_t$.

D State-count examples and initial-route design

D.1 The two-to-three-state boundary

D.1.1 Two-state simultaneous realizability

The two-state counterpart retains the same actual action in both worlds at every core time, independent untied time laws, a known normalized entrance on four pairs, and a nonnegative zero-diagonal terminal value. The certain selector/routing wrapper is unchanged in role. Write $p_{t,i} = \Pr(F_t(i) = 1)$ for $i \in \{0, 1\}$.

Proposition 16 (Two-state collapse)**.** *In this two-state class, $L = L_{\mathrm{pair}}$ for every finite horizon and every known entrance. A sequence of common quantile laws attains the complete pairwise continuation vector.*

Proof. At one time, put $x = \Pr(F(0) = 0, F(1) = 1)$ and $y = \Pr(F(0) = 1, F(1) = 0)$. Every complete response law obeys $x - y = p_1 - p_0$, hence

$$x \ge (p_1 - p_0)_+, \qquad y \ge (p_0 - p_1)_+.$$

For U uniform on $[0, 1]$, the single complete response $F(i) = \mathbf{1}\{U \le p_i\}$ attains both inequalities. For any $v \ge 0$ with zero diagonal,

$$(Kv)_{01} = xv_{01} + yv_{10}, \qquad (Kv)_{10} = yv_{01} + xv_{10}.$$

Thus this same law minimizes both coordinates. Starting from the terminal value, assume a common suffix attains the pairwise vector. Prepending this layer's quantile law attains its pairwise backup simultaneously; every other response-law backup is coordinatewise no smaller. Backward induction and the nonnegative entrance give equality of the minima. Use independent uniforms across times, with each time's response shared across worlds. Diagonal entrance mass is retained and contributes zero. □

This elementary proposition is a scope comparison, not a claim that all two-state counterfactual queries are easy. Different query actions, nonzero diagonal costs, or arbitrary factual conditioning can change its premises.

D.1.2 An all-compressible family with a parity-dependent strict gap

At every core time let the omission be 012, so the actual supports are $\{1, 2\}$, $\{0, 2\}$, and $\{0, 1\}$, all with second-successor probability $1/2$. Set $d_{01} = d_{12} = d_{20} = 1/3$ and all other entrance coordinates to zero. The complete wrapper in Appendix A.1 realizes this entrance. Rewards $h = (1, 1/2, 0)$ and threshold $\eta = 1/4$ give $g = e_{01} + e_{02} + e_{12}$. Every layer is D-fair, so $k = r = 0$.

Proposition 17 (Strict gaps inside the compressible class)**.** *For this input family and every $n \ge 1$, the sharp lower and pairwise values are*

$$L_n = \frac{2}{3 \cdot 2^n}, \qquad L_{\mathrm{pair},n} = \frac{1}{3 \cdot 2^n} \begin{cases} 1, & n \text{ odd}, \\ 2, & n \text{ even}. \end{cases} \tag{47}$$

Repeating any one canonical coalescing fair letter attains L_n. In particular, odd horizons have $L_n = 2L_{\mathrm{pair},n}$, whereas even horizons have equality.

Proof. Use the doubled operators A_0, A_1, A_3 and $N = A_2$ from Appendix C.3. In six-coordinate order, $d^\top = (1/3, 0, 0, 1/3, 1/3, 0)$ and $g = (1, 1, 0, 1, 0, 0)^\top$. Direct multiplication gives, for

$i \in \{0, 1, 3\}$,

$$d^\top A_i g = d^\top A_i^2 g = 2/3, \qquad d^\top A_i N g = d^\top A_i^2 N g = 4/3,$$
$$d^\top N = 2d^\top.$$

For $n \geq 2$, the seven sufficient words in (32) are A_i^n, $A_i^{n-1}N$, and N^n. The identities $A_i^3 = A_i$ and $N^2 = N + 2I$ hold exactly, and restoring 2^{-n} gives their scalar values

$$\frac{2}{3 \cdot 2^n}, \qquad \frac{4}{3 \cdot 2^n}, \qquad \frac{2}{3},$$

respectively. Their minimum is L_n. For $n = 1$, evaluate the four original fair pair-projection letters directly. Complement symmetrization covers the full fair law domain, so this is a minimum over all complete feasible laws, not just a face. Repeating a coalescer uses one feasible law per time with independent noises.

Let $g_{ij}^{\mathrm{rev}} = g_{ji}$ denote pair reversal, expressed in the same six-coordinate order. For pairwise backup T, the two endpoint couplings on each distinct pair give $Tg = \frac{1}{2} g^{\mathrm{rev}}$ and $Tg^{\mathrm{rev}} = \frac{1}{2} g$, with zero diagonal. Positive homogeneity and $d^\top g = 2/3$, $d^\top g^{\mathrm{rev}} = 1/3$ yield the parity formula by iteration. □

Together the two propositions locate the smallest actual-state count at which strict shared-law versus pairwise separation can occur within this same-action, zero-diagonal class. They do not establish a complexity lower bound. The odd-horizon gap is $1/(3 \cdot 2^n)$ and therefore decays; the family does not replace the separate nonfair examples distinguishing r from k.

Exact checks. The standalone construction checker included in the source bundle reconstructs the six full fair vertices from all 70 candidate bases, verifies the matrix identities, and evaluates all original-vertex words for $n = 1, \ldots, 5$. These finite checks supplement the arbitrary-horizon proof above; they are not benchmark runs.

D.2 A SELF-CONTAINED HETEROGENEOUS-COST ROUTING EXAMPLE

Take one D-fair layer with omission 012, rewards $h = (1, 1/2, 0)$ and threshold $\eta = 1/4$. A known selector first draws a uniform context $Z \in \{0, 1, 2\}$. In context j, the reference route starts at j and the treatment route moves to $j - 1 \pmod 3$ with probability x_j, staying at j otherwise. Realize this using known selector masses $x_j/3$ and $(1 - x_j)/3$ followed by deterministic move/stay routes, as in Appendix A.1. Thus

$$d_{01} = x_1/3, \qquad d_{12} = x_2/3, \qquad d_{20} = x_0/3, \qquad d_{jj} = (1 - x_j)/3.$$

The full entrance is normalized; diagonal mass is not renormalized away. Each selected core law serves the entire entrance and cannot condition on the realized selector. Equation (8) gives

$$L_{\mathrm{pair}}(x) = x_0/6, \qquad L_{\mathrm{shared}}(x) = (x_0 + \min\{x_0, x_1, x_2\})/6.$$

Use the rational route set

$$\mathcal{X}_\kappa = \{x \in [0, 1]^3 : \kappa x_0 + x_1 + x_2 \leq \kappa\}, \qquad \kappa > 2.$$

The unique pairwise optimum is $x^{\mathrm{PW}} = (1, 0, 0)$, since maximizing x_0 uses the whole budget. Its shared value is $1/6$. For $t = \min\{x_0, x_1, x_2\}$, feasibility implies

$$\kappa \geq \kappa x_0 + 2t = \frac{\kappa + 2}{2}(x_0 + t) + \frac{\kappa - 2}{2}(x_0 - t) \geq \frac{\kappa + 2}{2}(x_0 + t).$$

Consequently the unique shared optimum and its value are

$$x^\star = \frac{\kappa}{\kappa + 2}(1, 1, 1), \qquad L_{\mathrm{shared}}(x^\star) = \frac{\kappa}{3(\kappa + 2)}.$$

Equality in both inequalities forces $x_0 = t$ and $x_1 = x_2 = t$. The improvement over the most favorable pairwise-optimal design is therefore

$$\frac{\kappa}{3(\kappa + 2)} - \frac{1}{6} = \frac{\kappa - 2}{6(\kappa + 2)} > 0.$$

At $\kappa = 3$ this is $1/30$. This strict example does not assert a gain for every route budget: for the unit-cost set $\{x \in [0,1]^3 : x_0 + x_1 + x_2 \le 2\}$, $(1, 1/2, 1/2)$ is both shared-optimal and pairwise-optimal, with shared value $1/4$, so the analogous gain is zero. Indeed, $t = \min_j x_j \le (x_1 + x_2)/2$ gives $x_0 + t \le 1 + x_0/2 \le 3/2$, proving the upper bound $1/4$. These are supplied specifications, not measured treatment effects.

E Separation constructions and necessary representation size

E.1 A strict common-domain hierarchy

At time t let $p_{t,i} = \Pr(Z_i = 1)$ be the second-successor marginals. For each unordered pair $i < j$, specify $x_{ij} = \Pr(Z_i = Z_j = 1)$ in

$$[\max(0, p_{t,i} + p_{t,j} - 1), \min(p_{t,i}, p_{t,j})].$$

The four masses are $(1 - p_i - p_j + x_{ij}, p_j - x_{ij}, p_i - x_{ij}, x_{ij})$. Their product box $\mathcal{R}_t$ permits a separate choice for each unordered pair, with no requirement that all three pair laws extend to a common three-bit law. The reverse coordinate uses the transpose of the *same* two-bit law. The resulting nonnegative substochastic six-coordinate kernel is $K_t^{\mathrm{biv}}(x_t)$. Define

$$L_{\mathrm{bivariate}}(d) = \min_{(x_t) \in \prod_t \mathcal{R}_t} d^T K_0^{\mathrm{biv}}(x_0) \cdots K_{n-1}^{\mathrm{biv}}(x_{n-1}) g.$$

Choices are made offline, independent across time, and are not conditional on a realized pair. Projecting a complete feasible law gives a feasible box point at each time, hence $L_{\mathrm{bivariate}} \le L_{\mathrm{joint}}$. Coordinatewise Bellman minimization allows every coordinate to choose its own continuation minimizer; induction through nonnegative kernels gives $L_{\mathrm{Bellman}} \le L_{\mathrm{bivariate}}$. These are different quantifiers on the same supplied marginals, supports, entrance and terminal.

For a concrete strict separation, use chronological omissions and probabilities

$$(010, (1/3, 1/3, 1/3)), \quad (010, (1/3, 1/3, 1/3)),$$
$$(121, (1/5, 2/5, 3/5)), \quad (100, (1/4, 1/4, 2/3)).$$

Successors are the sorted complement of each omission. Take $g = e_{01}$, realized by rewards $(1, 0, 1/2)$ and threshold $3/4$, and $d_{01} = d_{12} = d_{20} = 1/3$. This has $n = k = 4, r = 2$. Separate affinity means a global minimum is attained by one vertex from each local domain. Exhaustive exact enumeration of the 4096 bivariate-box words and 2916 complete-law vertex words, together with coordinate minima, yields

$$L_{\mathrm{Bellman}} = \frac{4}{405} = \frac{16}{1620}, \qquad L_{\mathrm{bivariate}} = \frac{17}{1620}, \qquad L_{\mathrm{joint}} = \frac{1}{90} = \frac{18}{1620}.$$

The table lists minimizing words. Complete laws use atom order 000, 001, 010, 011, 100, 101, 110, 111; bivariate laws use (x_{01}, x_{02}, x_{12}) and need not extend to a complete law.

Time	Complete law	Global bivariate coordinates
0	$(1/3, 0, 1/3, 0, 0, 1/3, 0, 0)$	$(1/3, 1/3, 0)$
1	$(1/3, 0, 1/3, 0, 0, 1/3, 0, 0)$	$(0, 1/3, 1/3)$
2	$(2/5, 1/5, 0, 1/5, 0, 0, 0, 1/5)$	$(1/5, 1/5, 2/5)$
3	$(0, 1/2, 1/12, 1/6, 1/4, 0, 0, 0)$	$(0, 0, 0)$

The bundled construction checker reconstructs the menus, enumerates both full vertex-word domains and replays the displayed witnesses; enumeration covers the full domains by separate affinity, not only sampled entrances. This example separates the relaxations but does not assert strictness on every input. The bivariate formulation is inspired by Haugh & Singal (2023, Sections 4.1–4.3 and D.1–D.2). Their posterior-conditioned dynamic latent-state formulation, with its time-invariant copula constraints, is not identified with this deliberately common, unconditioned, untied-time specification.

E.2 Necessary nonmetric survival for any prescribed length

For $L \geq 1$ let $\theta = 2^{-(L+2)}$ and $M = 2^{L+1}$. Use terminal $g = e_{01}$, rewards $(1, 0, 1/2)$ and threshold $3/4$. The chronological suffix is

$$(121, (1/5, 2/5, 3/5)), \qquad (100, (1/4, 1/4, 3/4 - \theta)).$$

Prepend L fair layers whose omissions alternate $001, 022, 001, 022, \ldots$ in *backward* order; reverse that list in chronological inputs. Both suffix layers are exceptional and nonfair. Every prefix layer is X-matched. Thus $n = L + 2, k = r = 2$.

Full-domain suffix minimum. At the last layer the only possibly positive coordinates are 01 and 02. The relevant event has Fréchet lower probability $\max(0, 1-p_i-p_j)$, yielding $u_\theta = (1/2, \theta, 0, 0, 0, 0)$. The complete law

$$q_{\text{last}} = \theta\delta_{000} + (1/2 - \theta)\delta_{001} + (1/4)\delta_{011} + (1/4)\delta_{100}$$

attains both simultaneously with the prescribed marginals. At the preceding layer the second successors are $2, 1, 2$, and a pair's cost is $b_j \Pr(Z_i = 0, Z_j = 1) \geq b_j(p_j - p_i)_+$, where $b = (\theta, 1/2, \theta)$. The law

$$q_{\text{middle}} = (2/5)\delta_{000} + (1/5)\delta_{001} + (1/5)\delta_{011} + (1/5)\delta_{111}$$

attains all six bounds, giving

$$v_\theta = (1/10, 2\theta/5, 0, \theta/5, 0, 0) = (\theta/5)(M, 2, 0, 1, 0, 0).$$

Nonnegative earlier kernels preserve these bounds, so this is an original-domain suffix minimum, not just a feasible vector.

Four exact boundary representatives. After j prefix layers, let $A = 2^{j-1}$ and $c_j = \theta/(5 \cdot 2^j)$. In pair order $(01, 02, 10, 12, 20, 21)$ the following vectors, multiplied by c_j, are sufficient and necessary:

	odd j	even j
B	$(0, 0, 0, 0, M, M)$	$(M, M, 0, 0, 0, 0)$
S	$(0, 1, 0, 1, 2, 2)$	$(2, 2, 1, 0, 1, 0)$
N_1	$(A, 0, A, A, M, A + 1)$	$(A + 1, M, A, A, 0, A)$
N_2	$(A, A, A, 0, A + 1, M)$	$(M, A + 1, 0, A, A, A)$

The six vertices of the fair original law polytope, in atom order $000, 001, 010, 011, 100, 101, 110, 111$, are

$$q_0 = (0, 0, 0, 1/2, 1/2, 0, 0, 0), \qquad q_1 = (0, 0, 1/2, 0, 0, 1/2, 0, 0),$$
$$q_2 = (0, 1/4, 1/4, 0, 1/4, 0, 0, 1/4), \qquad q_3 = (0, 1/2, 0, 0, 0, 0, 1/2, 0),$$
$$q_4 = (1/4, 0, 0, 1/4, 0, 1/4, 1/4, 0), \qquad q_5 = (1/2, 0, 0, 0, 0, 0, 0, 1/2).$$

Completeness follows by exact solution of the seventy four-column bases of the rank-four equality matrix; every vertex support extends to such a basis. Let $T_q = 2K(q)$. At the base, omission 001 acting on $(M, 2, 0, 1, 0, 0)$ gives backups dominating, in menu order,

$$N_2, \quad N_1, \quad (B + S + N_1 + N_2)/4, \quad S, \quad (B + S + N_1 + N_2)/4, \quad B.$$

Here target $A = 1$, and $M \geq 3$. Targets are attained by indices $(5, 3, 1, 0)$ for (B, S, N_1, N_2).

For induction let A be the *old* parameter and target parameter be $2A$. On the right of the table all letters denote targets. Put $C = (B + 2S + 2N_1 + N_2)/6$ and $D = (2S + N_1 + N_2)/4$.

Transition	old	q_0	q_1	q_2	q_3	q_4	q_5
odd to even, 022	B	B	B	B	B	B	B
	S	S	S	S	S	S	S
	N_1	B	N_2	C	N_1	C	S
	N_2	S	S	C	S	C	B
even to odd, 001	B	B	B	B	B	B	B
	S	S	S	S	S	S	S
	N_1	S	S	D	S	D	S
	N_2	N_2	N_1	D	S	D	B

Each raw doubled backup dominates the listed convex combination when $A \geq 1, M \geq 4A + 1$. This is a finite polynomial identity check: substituting $A = 1 + x, M = 5 + 4x + y$ makes all coefficients of every coordinate difference nonnegative. The bundled checker reconstructs all 54 six-coordinate differences, including the base, directly from sorted supports and atom kernels and checks their polynomial coefficients exactly. No extrapolation from numerical horizons is needed. At an allowed step $j \to j + 1 \leq L$, $M \geq 8A \geq 4A + 1$.

In odd-to-even transitions, $(B, q_0), (S, q_0), (N_1, q_3), (N_1, q_1)$ attain the targets. In even-to-odd transitions use $(B, q_0), (S, q_0), (N_2, q_1), (N_2, q_0)$. These are exact polynomial equalities. Inductively every original suffix value dominates a convex combination of the retained four: nonnegative kernels preserve domination, arbitrary fair laws expand in the complete six-vertex menu, and the table supplies the next cover. Attainment concatenates actual laws. Therefore the four targets give the full all-entrance envelope; convex combinations certify coverage, not correlated time actions.

Necessity over the full cone. For odd parity, exposing entrances for B, S, N_1, N_2 are respectively

$$e_{02} + e_{12}, \quad e_{20} + e_{21}, \quad Ae_{02} + e_{21}, \quad Ae_{12} + e_{20}.$$

For even parity they are $e_{10} + e_{20}$, $e_{01} + e_{02}$, $e_{01} + Ae_{20}$ and $e_{02} + Ae_{10}$. Their three competitor-minus-target gaps before scaling are

$$(2, A, A), \quad (2M - 4, A + M - 3, A + M - 3), \quad (M - A - 1, 1, A^2 - A + M - 1),$$

with the last triple for both nonmetric values. All are positive for $M > 2A + 1$. If γ is the smallest gap and $D_0 = \max_{w \neq v} |\mathbf{1}^T(w - v)|$, perturb the entrance by $\gamma\mathbf{1}/[2(D_0 + 1)]$. It becomes strictly positive and still exposes the target. Since every actual value dominates a convex combination of targets, equality at this positive entrance forces precisely the target and zero coordinate slack. Thus any finite attained representation must contain it; irredundance is not inferred merely from absence of coordinate domination.

Each N has exactly one negative triangle slack, $2A + 1 - M$; its other five slacks and all slacks of B, S are nonnegative. The same checker expands all 48 slack polynomials. The Bellman vector, in the same normalization, is $(0, 0, 0, 0, 2, 2)$ at odd parity and $(2, 2, 0, 0, 0, 0)$ at even parity, by direct coordinate minimization of the six doubled backups. At the uniform entrance, S has sum 6, while B has sum $2M$ and either N has sum $M + 4A + 1$. At $j = L$, S is uniquely minimal, giving

$$L(d) = c_L = \frac{1}{20\,4^L}, \qquad L_{\text{pair}}(d) = \frac{2}{3}c_L, \qquad L(d) - L_{\text{pair}}(d) = \frac{1}{60\,4^L}.$$

The ratio is $3/2$, with decaying absolute gap. The denominator of θ has $L + 3$ binary digits; variable suffix precision and the explicit prefix description grow as $O(L)$. Exactly four values are necessary at every prefix boundary.

E.3 A LINEAR LOWER BOUND IN THE EXCEPTION COUNT

Proposition 18. *For every $\ell \geq 1$, a three-state instance with $n = r = 2\ell + 1$, a fixed symmetric-metric terminal and $O(\ell)$ input bits requires at least ℓ distinct vectors in every finite attained family representing its sharp minimum over all nonnegative entrances.*

Proof. Every layer has omission map $(0, 1, 2)$ and second-successor marginals $(1/10, 2/5, 3/5)$; let $g_{ij} = \mathbf{1}\{i \neq j\}$. These are positive binary supports with a common action and untied time laws. All layers are exceptional: distinct omissions and nonfair probabilities. The terminal and probabilities have constant encoding size.

Full-domain reduction. Write a symmetric metric as $t = (x, y, z) = (t_{01}, t_{02}, t_{12})$ and put $u = (x + y - z)/2$, $v = (x + z - y)/2$, $w = (y + z - x)/2$. These are its nonnegative coefficients on the cuts $(1, 1, 0)$, $(1, 0, 1)$ and $(0, 1, 1)$. Averages of metric pullbacks preserve this cone. The full local polytope has the following nine vertices in the usual eight-atom order:

$$\begin{aligned}
20q_0 &= (0, 10, 6, 2, 2, 0, 0, 0), & 20q_1 &= (0, 10, 8, 0, 0, 2, 0, 0),\\
20q_2 &= (0, 11, 7, 0, 1, 0, 0, 1), & 20q_3 &= (0, 12, 6, 0, 0, 0, 2, 0),\\
20q_4 &= (2, 10, 6, 0, 0, 0, 0, 2), & 20q_5 &= (6, 4, 0, 8, 2, 0, 0, 0),\\
20q_6 &= (6, 6, 0, 6, 0, 0, 2, 0), & 20q_7 &= (8, 2, 0, 8, 0, 2, 0, 0),\\
20q_8 &= (8, 4, 0, 6, 0, 0, 0, 2).
\end{aligned}$$

The seventy four-column bases of the rank-four marginal system give this complete list. Let $A = K(q_6)$, $B = K(q_8)$ and $C = K(q_5)$. On (x, y, z),

$$A = \frac{1}{10}\begin{pmatrix} 6 & 0 & 3 \\ 3 & 1 & 0 \\ 3 & 1 & 3 \end{pmatrix}, \quad B = \frac{1}{10}\begin{pmatrix} 6 & 0 & 3 \\ 4 & 0 & 1 \\ 2 & 0 & 4 \end{pmatrix}, \quad C = \frac{1}{10}\begin{pmatrix} 5 & 1 & 4 \\ 3 & 1 & 0 \\ 2 & 0 & 4 \end{pmatrix}.$$

For $q_0, \dots, q_8$, respectively, the operators $(A, B, A, A, B, C, A, B, B)$ are coordinatewise no larger on each of the three metric generators, hence on their entire cone. Thus every original vertex word dominates an $A/B/C$ word, by successive replacement from the terminal. Separate affinity extends this coverage to arbitrary original laws through convex combinations of vertex-word values.

The $A/B/C$ values also satisfy $u \geq w$ and $u > 0$. Initially $u = v = w = 1/2$; their coefficient maps, multiplied by ten, are

	$10u'$	$10v'$	$10w'$
A	$3u + 3v$	$3u + 6v + 3w$	$u + w$
B	$4u + 4v$	$2u + 5v + 3w$	$v + w$
C	$4u + 3v + w$	$2u + 6v + 4w$	0

and directly preserve these inequalities.

One-switch minimizers. For entrance weights $d = (\alpha, \beta, \gamma)$ in $\mathcal{R} = \{\gamma \geq \alpha \geq 3\beta \geq 0\}$, $dAt - dCt = (\gamma u - \alpha w)/5 \geq 0$. It is strict if $\gamma > \alpha$. The transformed entrance $d\overline{C} = \frac{1}{10}(5\alpha + 3\beta + 2\gamma, \alpha + \beta, 4\alpha + 4\gamma)$ stays in $\mathcal{R}$: $\alpha' - 3\beta' = (\alpha + \gamma)/5$ and $\gamma' - \alpha' = (2\gamma - \alpha - 3\beta)/10$. Strictness is preserved. Hence before the first B, each minimizing choice is C. After B, the entrance's y coefficient is zero. Moreover B simultaneously minimizes x, z among the three operators: their excesses are $(At - Bt)_{xz} = (0, u/5)$ and $(Ct - Bt)_{xz} = (w/5, 0)$. Its x, z evolution is

$$M = \frac{1}{10}\begin{pmatrix} 6 & 3 \\ 2 & 4 \end{pmatrix}.$$

Induction gives the simultaneous lower value $M^m(1, 1)^T$ for a suffix of length m. Thus all subsequent choices can be B. With positive x, z entrance weights, any minimizing suffix has exactly these two coordinates; the preceding B discards its y coordinate, producing the same complete vector. Thus, for $\alpha, \beta > 0$ and $\gamma > \alpha$ in $\mathcal{R}$, the full-domain minimum is represented by $V_j = C^j B^{n-j} g$, $0 \leq j \leq n$.

The exposed values. Set $(x_m, z_m)^T = M^m(1, 1)^T$. After B^m, $w = z_{m-1}/10$ for $m \geq 1$; C sets $w = 0$, or $y = x - z$. On this face, C and B have the same x, z evolution. Consequently, for $1 \leq j \leq n - 1$,

$$(V_j)_{xz} = M^n\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{z_{n-j-1}}{50} M^{j-1}\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad (V_j)_y = (V_j)_x - (V_j)_z.$$

Take $d_h = (3/4, 1/4, h + 1/4)$ with $h > 1/2$. Its objective at V_j is $x_n + hz_n + f_q(h)/50$, where

$$N = n - 2, \quad q = j - 1, \qquad f_q(h) = z_{N-q}(1, h) M^q \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad 0 \leq q \leq N.$$

The all-B objective is $x_n + hz_n + z_{n-1}/20$, strictly larger than the minimum of these values: $\min_q f_q \leq z_{n-2} < (5/2)z_{n-1}$. The all-$C$ objective is also larger than that of V_{n-1}, since its extra x, z vector exceeds that of V_{n-1} by $M^{n-2}(1/25, 1/50)^T$, and the objective on their common face is $x + hz$.

Put $s = \sqrt{7}$, $\lambda = (5 + s)/10$, $\mu = (5 - s)/10$ and $\rho = \mu/\lambda \in (0, 1)$. Direct diagonalization gives

$$z_m = Z\lambda^m + W\mu^m, \qquad (1, h) M^q \begin{pmatrix} 1 \\ 0 \end{pmatrix} = U\lambda^q + V\mu^q,$$

$$Z = \frac{s + 1}{2s}, \quad W = \frac{s - 1}{2s}, \quad U = \frac{s + 1 + 2h}{2s}, \quad V = \frac{s - 1 - 2h}{2s},$$

$$f_q/\lambda^N = ZU + WV\rho^N + ZV\rho^q + WU\rho^{N-q}.$$

For $1/2 < h < h_* = (s-1)/2$, $R(h) = ZV/(WU)$ decreases strictly from ρ to zero. The sign of $f_{q+1} - f_q$ is the sign of $-R(h)\rho^q + \rho^{N-q-1}$, which changes at most once. Therefore $q \geq 1$ is the unique global minimizer whenever $\rho^{N-2q+1} < R(h) < \rho^{N-2q-1}$. With $n = 2\ell + 1$, each $q = 1, \ldots, \ell - 1$ has such an interval intersecting $(0, \rho)$, since $N - 2q + 1 \geq 2$. For $q = 0$, take $h = 1$: $V < 0 < U$ makes every successive difference positive.

These exposing entrances can be rational with $O(n)$ bits. If $h_{q,r}$ is the rational intersection of the affine functions f_q and f_r, choose, for $q \geq 1$, $h_q = \frac{1}{2}(\max\{1/2, h_{q,q+1}\} + h_{q-1,q})$, with $h_0 = 1$. The strict intervals just proved order these endpoints. Powers of the fixed rational matrix M, their intersections and these midpoints all have $O(n)$ bit length. Split each unordered-pair weight equally between its two orientations and normalize by $5/4 + h_q$ to obtain a strictly positive rational probability entrance.

Each entrance uniquely minimizes an $A/B/C$ value vector, including all ties after its first B. Coordinatewise reduction and entrance positivity force every optimal original-vertex word to have zero slack. Every positive-weight term in a nonvertex law's convex expansion is optimal, with this same vector. Thus uniqueness holds on the original continuous domain: its ℓ distinct exposed vectors belong to every sufficient attained family. □

This linear lower bound on $n = r$ neither establishes exponential necessity in r nor excludes a polynomial-in-n bound for the full class. The artifact supplies the generator and exact checks.

F VARIABLE-STATE COMPLEXITY AND ATTAINED REPRESENTATION

We prove Proposition 5 inside the full complete-law domain. There are $N = 2m$ actual core states. Row s has sorted successors $s \bmod m$ and $s \bmod m + m$, both fair. All joint distributions on the N response bits with fair marginals are admissible. Both worlds use the same action and realized complete map. Since adjacent distinct rewards differ by $1/(2m-1) > 1/(2m)$, the terminal event is precisely $u > v$. No equality tolerance is involved.

F.1 MAX-CUT REALIZATION AND THE EXACTNESS DISTINCTION

For a loop-free nonnegatively weighted graph on lower-half vertices with $W = \sum w_{ij} > 0$, orient its positive-weight edges $i > j$ and set $d_{ij} = w_{ij}/W$. Four bit assignments give

$$g(i + mz_i, j + mz_j) = 1 - z_j + z_i z_j = 1 - \tfrac{1}{2}\mathbf{1}\{z_i \neq z_j\} + \tfrac{1}{2}(z_i - z_j).$$

Fairness cancels the last term. Thus the full objective is $1 - \mathbb{E}\,\mathrm{cut}_G(Z)/(2W)$, at least $1 - \mathrm{MaxCut}(G)/(2W)$. Choose a maximum cut and its complement with mass $1/2$ each, extend its unused upper-half bits arbitrarily in the first atom and complement them in the second. This is a feasible full law attaining equality. Each pair alone can use fair anti-correlated bits, attaining $1/2$; this pair law extends to the remaining rows. Hence the pairwise value is $1/2$.

For unweighted graphs, deciding $L_G \leq 1 - K/(2|E|)$ is equivalent to deciding whether a cut of size at least K exists. This is a polynomial reduction and a cut is a polynomial-size certificate checked by counting edges. This restricted threshold language is NP-complete; exact sharp evaluation is NP-hard. It asserts no NP membership for every general shared-mechanism formulation. The compact input lists supports, marginals, rewards, threshold and entrance, not all 2^N response columns. Polynomial time in a fully expanded LP would not contradict the result.

The support graph is bipartite iff its maximum cut has weight W, exactly the condition $L_G = L_{\mathrm{pair,G}}$. Thus exactness is easy to decide in this family although its sharp value is hard to compute as N grows. The known reduction of MAX-CUT to pairwise-interaction multimarginal costs is credited to Altschuler & Boix-Adserà (2021, Section 6.2, Proposition 6.3 of the cited arXiv version); the increment here is its explicit shared-mechanism threshold realization and query-exactness interpretation.

F.2 FULL NONNEGATIVE ENTRANCE DOMAIN: EXACT ATTAINED-FAMILY COUNT

An attained value is the complete pair-continuation vector of one feasible complete response law. The family must attain the lower envelope for every nonnegative directed core entrance. For source

rows s, t, write $a = s \bmod m$, $b = t \bmod m$ and $\delta_{st} = \Pr(Z_s \neq Z_t)$. Direct threshold evaluation and fairness give

$$v_{st} = \begin{cases} 1 - \delta_{st}/2, & a > b, \\ \delta_{st}/2, & a < b, \\ \delta_{st}/2, & a = b. \end{cases}$$

For $s = t$ this is zero. Complementing and averaging a law therefore preserves its entire value vector, even for distinct rows with identical supports. A symmetric law is a convex mixture of the 2^{N-1} laws uniform on $\{z, 1 - z\}$. Their attained values suffice for every nonnegative entrance. This is the symmetric Bernoulli/cut-polytope structure of Huber & Marić (2017, Theorem 1, Corollary 1, Proposition 2 and Theorem 4), expressed in threshold-cost coordinates.

To prove necessity using only nonnegative entrances, let E_0 consist of all unordered pairs of rows with different residues modulo m. This is the connected graph K_N minus a perfect matching. Fix any $z \in \{0, 1\}^N$, including the constant pattern. On each edge choose one directed entrance of weight $1/|E_0|$: put the higher-residue row first when $z_s \neq z_t$, and the lower-residue row first otherwise. For every full feasible law, its objective equals

$$b_z + \frac{1}{2|E_0|} \sum_{\{s,t\} \in E_0} \Pr(\mathbf{1}\{Z_s \neq Z_t\} \neq \mathbf{1}\{z_s \neq z_t\}), \quad b_z = \frac{|\{e \in E_0 : z \text{ separates } e\}|}{2|E_0|}.$$

The complement-pair law attains b_z. Equality forces every nonnegative mismatch probability to be zero. Connectedness forces $Z = z$ or $1 - z$; fairness forces equal probabilities. The full law, hence its full vector, is unique. The connected graph's disagreements distinguish all 2^{N-1} complement pairs. Every finite attained representing family must include every one, proving the exact minimum.

This result uses the full nonnegative directed entrance domain. Restricting entrances to lower-half coordinates $i > j$, as in the hardness reduction, is a different representation problem and does not give this necessity argument.

SCM and representation scope. A finite rational selector can route the query worlds to each entrance pair while a factual wrapper of probability one follows its fixed route, as in Appendix A.1. Selector/helper states have polynomial description and are not counted in N. There is no pair-observing action or restriction of admissible full laws to two-atom laws; such laws only witness extrema. The result concerns this family and finite explicit attained vectors, not every implicit representation or extended formulation. It is not fixed-four-state hardness, nor a lower bound in the three-state horizon n or exception count r.

G STANDALONE STRUCTURAL CERTIFICATE SOUNDNESS

The executable schema binds original tables, supplied six-coordinate terminal cost, reduced derivation, final values and chronological words. The structural input guard checks nonnegativity and at most one negative triangle slack; the diagonal is fixed to zero by this representation. It establishes the scalar lower envelope for *all* nonnegative entrances. It does not assert identity of full law distributions or coordinatewise domination by one retained word. The generic solver and structural solver share exact local algebra; this trust boundary is explicit.

G.1 ORIGINAL-DOMAIN LOCAL COVERAGE

For a continuation v and retained feasible local laws q_j, each original local vertex q has rational nonnegative certificate weights λ_j summing to one and satisfying

$$K(q)v \geq \sum_j \lambda_j K(q_j)v$$

coordinatewise. The checker recomputes every backup and inequality. The at-most-seventy four-column bases enumerate the complete local polytope, including degenerate vertices. Every original law therefore has a no-larger mixture of retained backups. For each fixed $d \geq 0$, some retained backup is no worse than that mixture; feasibility gives the reverse scalar inequality. The selected law can depend on d.

This implements exceptional face reductions and nonmetric D/X/S reductions. The face menu is a union of original vertices with at most one positive permutation atom, with seventy as the bound on the *whole union*. The checker requires coverage for every original vertex, including those outside the union. It checks $\mathcal{C}_1$ membership on the actual surviving values, not on arbitrary convex mixtures.

For X matching it reconstructs the duplicate sorted supports and exact marginals, semantic labels, a, b, $\ell = \max(0, a - b)$, $h = \min(a, 1 - b)$ and $\lambda = 2(h - \ell)$. It checks the coalescing subprobability law ρ, of mass $1 - \lambda$, and feasibility of the four laws $\rho + \lambda q_{\mathrm{fair}}$. Their sufficiency for the supplied continuation is established by the same full-original-vertex convex witnesses. It does not demand that a single corner dominate every original law.

G.2 Universal pure-D rule

For every actual fair D layer, the checker verifies that each original-vertex kernel equals the mixture of four complementary-law kernels with weights $q(b) + q(1 - b)$ for $b = 000, 100, 010, 001$. Original-law linearity and temporal separate affinity then reduce scalar minimization to projected four-letter words. This is not correlated selection of a whole temporal word.

The checker reconstructs the canonical unnormalized matrices $A_i = 2K_i$ and verifies $N^2 = N + 2I$, $A_i^3 = A_i$ and $A_j A_i \geq A_i^2$ for all $i \in \{0, 1, 3\}$ and $j \in \{0, 1, 2, 3\}$. The rightmost-coalescer induction in Appendix C.3, equations (30)–(32), proves the seven-word rule at every length, including the four-letter and identity boundary cases. The checker also verifies the actual boundary permutations (33), atom transport (34), and all distinct rule outputs. Thus long-word coverage depends on an explicit induction, not a finite long-word test.

G.3 Metric completion, first entry and global composition

The rightmost non-D boundary is derived from the original layers. Its pure-D suffix uses the previous rule. At each earlier metric completion, the recorded law must be feasible and its actual backup no greater coordinatewise than that of *every* original local vertex for that supplied continuation. This verifies a simultaneous minimum only where it actually exists. All six triangle slacks are checked.

For nonmetric parents, the trace accounts for every child. A coordinatewise deletion names a final kept child from the same parent, excluding circular chains. Every metric child has its explicit metric-prefix completion; every nonmetric child remains pending. The two-pending-record and one-splitting-parent properties are checked consequences, never caps used to discard records. A trace can stop only when no pending obligation remains.

A nonnegative prefix sends any nonnegative entrance to another such entrance. Consequently each certified all-entrance suffix substitution is valid after every feasible prefix. Reconstructing all chunks and exceptional boundaries and checking their reduced raw families gives a global induction. Sequential convex deletion certificates preserve the envelope, and final exposing entrances establish irredundance. Every retained law is feasible; concatenation is chronological; each output word is replayed against original marginals and terminal. Coverage gives optimality, while replay gives attainment. Neither obligation substitutes for the other.

G.4 Implementation and trust boundary

Floating-point proposals are validated by exact rational inequalities (Applegate et al., 2007). Construction alone proposes LP witnesses; normal checking invokes no optimizer, old compiler, external process or saved answer. It reconstructs local menus and traces. Regression tests disable generic/old entry points during generation and prune/redundancy/phase-I during checking; mutated laws, coverage, gauges, branches and outputs must be rejected.

The trust base includes rational arithmetic, exhaustive basis enumeration, linear algebra, atom-to-kernel mapping, the seven-word induction, parsing and chronology, and Python with assertions enabled. Optimized Python is rejected. The transport cache is cleared at checker entry; within a check it reuses only already reconstructed transport. Distributed certificates have separate checking and word-replay entry points. These finite tests do not establish formal refinement of Python; mathematical Lean scope is mapped in Appendix J.

H HISTORICAL COMPARISON AND PARSER SENSITIVITY

The historical 72-input manifest covers D/order and edge, X/S, nonzero all-nonfair and reachable nonmetric cases, horizons 16–512 and $r \in \{0, 1, 2, 4, 8\}$; the legacy 80-input cohort is separate. Five serial, seeded-interleaved processes per method/input used guest CPU 2, single-threaded GMP and 120-second/4-GiB limits. Parent wall includes parsing, construction/proofs, serialization, coverage and nine-pair replay. Swap or steal above 0.005 permits up to two retries; the first exact clean pass is selected. The complete plan retains 900/930 selected/attempted runs, including 720 primary.

The complete-cost geometric ratio is 0.5827289321, with descriptive input-bootstrap interval $[0.5065352196, 0.6634696294]$ and 66/72 lower block medians. The post-hoc zero-function split has 33 inputs, ratio 0.9179582239 and 27 lower; the other 39 give 0.3967096440, all lower. All six ratios above one belong to the zero group: four exceed one by less than 3×10^{-4}; D-edge $n = 512, r = 0, 2$ give $1.177217, 1.088697$. The construction-with-proofs ratio is 0.2553455725; complete cost is primary. Figure 5 displays all 72 inputs.

The typed rational-size parser excludes metadata strings and finds maxima of 1257/2149 numerator/denominator bits. A separate 13-input sensitivity study retains 260/264 selected/attempted runs with identical exact families. Corrected/old block-generic ratios are 0.43655/0.44399; corrected/old costs are 1.05066 for generic and 1.03305 for block. The separate current-source confirmation follows in Appendix N.

I RELATION TO PRIOR METHODS

Response laws and sharp bounds. Structural counterfactual semantics and finite response-function optimization originate in Pearl (2009, Chapter 7) and Balke & Pearl (1994, Sections 3.1–3.3). The threshold contrast relates to causation and ordered-benefit bounds (Tian & Pearl, 2000; de Aguas et al., 2025). Finite canonical SCMs preserve counterfactual distributions; their independent exogenous factors yield polynomial objectives (Zhang et al., 2022, Theorems 2.4 and 2.8). Polynomial programming supplies sharp bounds and globally valid outer approximations (Duarte et al., 2024), while credal formulations give exact and approximate algorithms for further SCM classes (Zaffalon et al., 2024). Objective-dependent merging of response states (Choe et al., 2026) reduces the canonical LP domain; our support and value reductions retain the original mechanism domain.

Shared noise and dynamic models. Gumbel-Max SCMs specify cross-world couplings through shared noise (Oberst & Sontag, 2019); learned Gumbel mechanisms adapt couplings to query distributions (Lorberbom et al., 2021). The latter's Proposition 2 establishes that a universal categorical coupler cannot always be maximal. Haugh & Singal (2023, Sections 4.1–4.3) optimize complete-law and linked bivariate uncertainty in posterior-conditioned dynamic models. Our hierarchy in Appendix E.1 uses an unconditioned core with independently selectable time laws. Lally et al. (2026, Theorem D.7) simultaneously realize an entailed counterfactual MDP under stability and monotonicity restrictions. Kaya et al. (2026, Definition 4.1) supply a shared one-step sampler and evaluate joint moments with conditionally independent future trajectories; our compatible response laws remain unknown and shared across worlds throughout the trajectory.

Bellman exactness. Rectangular robust uncertainty supports dynamic programming (Iyengar, 2005; Nilim & El Ghaoui, 2005; Wiesemann et al., 2013). Tractable factor coupling (Goyal & Grand-Clément, 2023) and budgets on departures from nominal transitions (Mannor et al., 2012) provide other structured regimes. For compact stationary uncertainty sets, Grand-Clément et al. (2026, Theorems 1 and 5) characterize Bellman evaluation uniformly over policies, rewards and initial distributions; Theorem 5 restricts rewards to be next-state independent. The common minimizing kernel may depend on the policy and reward. In finite horizon, Li & Shapiro (2025, Theorem 3) give a sufficient static/game equivalence condition: at each time a whole kernel solves all statewise Bellman problems for each policy. Their Theorem 4 preserves the game value on the enlarged statewise rectangular hull. Our query-specific zero-residual/no-leakage criterion uses the occupancy/performance-difference principle (Kakade & Langford, 2002, Lemma 6.1). Value-dependent nonrectangularity measures control approximate evaluation (Li et al., 2026, Section 3.2); the present support test characterizes exact equality.

Same-domain comparison. Complete each substochastic kernel with an absorbing zero-cost state and use its feasible image as a one-action robust-MDP ambiguity set. With zero intermediate costs, terminal $-g$ and a normalized entrance, the static and game values are $-L$ and $-L_{\mathrm{rel}}$. For a fixed model and terminal, let A require a zero-residual law on all coordinates at every layer, $E(x)$ mean $L_0(x) = x^T\phi_0$, and $C(x)$ require a first-layer zero-residual law only on $\operatorname{supp} x$. Li–Shapiro's Theorem 3(i) becomes A on this embedded subclass; their Theorem 4 gives the relaxed value on its rectangular hull. Concatenation and the occupied-residual identity yield $A \Rightarrow E(x) \Rightarrow C(x)$. Neither converse holds: the same two-stage model in Appendix K has $E(e_{01})$, whereas $x = (e_{01} + e_{02})/2$ satisfies $C(x)$ but not $E(x)$. Thus A is false for this model. Theorem 2 strengthens current compatibility to an exact support-viability condition, returning an original-domain chronological word. This is a query-specific refinement on the embedded subclass; the stationary robust-MDP results address different universal task quantifiers.

Value representation and pruning. Finite-horizon dynamic programming provides the baseline (Puterman, 1994); POMDP values admit piecewise-linear representations (Smallwood & Sondik, 1973; Kaelbling et al., 1998). After negating terminal cost, our pair process is a blind POMDP with nine hidden states, one observation and time-dependent complete-law actions. Exact vector backups and witness-region pruning (Cassandra et al., 1997; Feng & Zilberstein, 2004) supply the generic representation. Original-domain repair and first-entry counting bound this representation and generic pruning (Appendix C.5). Antichains encode downward-closed support sets (Chatterjee et al., 2007); qualitative reachability and quantitative POMDP cost remain distinct questions (Chatterjee et al., 2016). The zero-face search uses the established infeasible-core idea (Chinneck & Dravnieks, 1991) to filter and branch within a fixed time/target problem.

Coupling, transport and geometry. Simultaneous near-optimal pair couplings (Li & Anantharam, 2019) expose the compatibility issue. Bicausal Markov transport chooses a coupling for the current pair (Moulos, 2021, Theorem 1); our offline law serves the entire entrance mass. Multimarginal causal bounds (Gao et al., 2025), covariate-assisted conditional transport (Lin et al., 2026) and algorithms for structured multimarginal costs (Altschuler & Boix-Adserà, 2023) use different information or cost structures. The common-quantile construction (Fill & Machida, 2002, Eq. (1.5)) underlies our directed-path calculation. The variable-state boundary uses MAX-CUT-based multimarginal hardness (Altschuler & Boix-Adserà, 2021) and symmetric Bernoulli cut geometry (Huber & Marić, 2017), adding the legal threshold realization and exposing entrances for the exact attained-family count.

Certification and parameters. LP duality and separation (Boyd & Vandenberghe, 2004), and rational LP with floating-point assistance (Applegate et al., 2007), support the implementation. The executable rational-simplex fallback and the theorem's polynomial-bit exact-LP premise are distinct. Fixed-parameter linear counterfactual satisfiability in domain size and endogenous-variable count (Ganian et al., 2026, Theorem 11) charges a literal time expansion at every stage; our parameter charges exceptional layers while retaining independent time-law factors.

J ARTIFACT AND REPRODUCTION

The source supplies editable LaTeX, seven vector figures and display data. The claim map separates analytic proofs, executable output checks, finite exact tests and mapped Lean endpoints, with source identities and reproduction commands. All measured attempts and source/input/certificate identities remain versioned.

The 1,473 Lean sources retain pinned dependencies and source-bound historical Linux-build and 2,705-declaration audit records using `propext`, `Classical.choice` and `Quot.sound`. Their endpoints connect stored mathematical executions to attainment, coverage and charged exact-LP operations. Theorem 4 has an analytic proof for supplied rational $\mathcal{C}_1$ costs; the mapped Lean construction endpoints cover bounded one-defect core costs, and the encoded complete-model endpoint covers threshold terminals. Corollary 19 has an analytic bounded-cost transport proof. Its finite executable tests check terminal scaling, zero and unnormalized entrances, complete-law feasibility and chronological replay, alongside the preserved indicator cases. Python certificate checks and their finite tests are listed separately from these Lean endpoints. The supplementary index also identifies the retained study protocols and historical reconstructions.

K AN IMPLICIT MECHANISM DOMAIN IN OPERATION

The general interface accepts constraints beyond nominal marginals. Take four actual states, all four successors, $h = (0, 1/3, 2/3, 1)$ and $\eta = 1/6$. The marginal rows are $P = \frac{1}{10}[(1,2,3,4);(4,1,3,2);(2,4,1,3);(3,2,4,1)]$. The 256 response atoms $z \in \{0,1,2,3\}^4$ index variables $q_z \geq 0$; normalization and three marginals per row give 13 equalities. A second supplied domain adds $\sum_{z:z_0=z_1} q_z = 23/100$. The product law proves feasibility of this moment constraint but is only one admissible law. Neither domain is passed as a polytope-vertex list. The marginal/moment domains require 130/364 exact LPs and accept 1715/570 nonempty supports, represented by 8/42 maxima. Each includes 12 coordinate LPs.

The unchanged zero-face compiler obtains coordinate minima and branches on infeasible support cores. The marginal-domain classification matches its saved exhaustive 4096-support reference. In the moment-constrained domain, uniform mass on $\{01, 12, 13, 23\}$ is tight with value $307/1600$. Uniform mass on $\{01, 10\}$ is loose by $83/480$: the joint constraint couples the two opposite improvement events. Exact primal/dual witnesses certify both outcomes. A separate optimizer-free Fraction checker reconstructs each input and LP cost, verifies all 494 certificates and both complete support partitions, and rejects six intended corruptions. This illustrates constraint-form diagnosis rather than a vertex-enumeration speed benchmark. Code and witnesses are in the implicit-domain evidence.

Two stages: local tightness need not continue. Keep four states and the same terminal, but give rows $0, 1, 2, 3$ the respective ordered supports $012, 123, 023, 013$ at both times. Their probabilities are

$$P_0 = \tfrac{1}{4}[(1,1,2);(2,1,1);(1,2,1);(2,1,1)], \qquad P_1 = \tfrac{1}{4}[(1,2,1);(1,1,2);(2,1,1);(1,2,1)].$$

There are 81 atom variables and nine marginal equalities per layer. Add $\Pr(F_0(0) = F_0(1)) = 1/4$ and $\Pr(F_1(0) = F_1(1)) = 3/16$; the independent-row laws satisfy these moments. For $S = \{01, 02\}$, one feasible first law has zero residual on both coordinates, with $\phi_0(01) = 1/16$ and $\phi_0(02) = 1/8$. Yet every one of the ten maximal acceptable next supports T has $\min_{q \in Q_0} \Psi_{0,S,T}(q) > 0$, the smallest being $5/128$. Thus current tightness cannot be continued without leakage. This penalty certifies infeasibility of the zero face, not the numerical optimal gap. In contrast, entrance $d = e_{01}$ has a checked two-stage attaining word of value $1/16$. The moment-constrained/original domains require 1022/589 compilation LPs and have 16/14 maximal initial supports. The original domain also has a locally tight but dynamically rejected entrance. Readback checks all 1630 rational certificates, including explanatory probes, 69632 support–target classifications and both full words; 10 intended corruptions are rejected. No polytope vertices are enumerated. Inputs, all probes and witnesses are in the dynamic implicit-domain evidence.

L CERTIFIED FUNCTIONS AND SUPPLIED ROUTING DESIGNS

L.1 ROUTE OPTIMIZATION FROM AN ATTAINED ENVELOPE

The supplied selector gives $d_{j-1,j} = x_j/3$, with indices modulo three, $x \in [0,1]^3$ and the remaining mass diagonal. The two route polytopes use costs $(3,1,1)$ with budget 3 and $(1,1,1)$ with budget 2. Given a complete attained family, the design LP maximizes z subject to the route constraints and $z \leq \sum_j x_j v_{j-1,j}/3$ for every $v \in H$. The stored exact LPs include $z \geq 0$, preserving the optimum, and enumerate feasible intersections of four constraints and select the largest z. The polytope is bounded, so this certifies global optimality. A minimizing active vector and its chronological word certify attainment at the selected route.

L.2 NONHOLDING PROGRAMS AND CHANGED OPTIMAL ROUTES

The retained grid removes S layers: D/X matched stages, $n = 4, 8, 12$, $r = 2, 3$, order/edge terminals, dispersed calibrations and fixed seeded support permutations at denominator 11. All 24 cells are retained in the complete-function comparison. Six D-order families contain 7–21 values; four have maximum normalized gap at least 0.001. All 18 other cells are pairwise exact.

For the main-text $n = 8, r = 3$ case, row omissions are $012, 102, 102, 201, 102, 012, 201, 120$. Probabilities are $(1/2, 1/2, 1/2)$ except at zero-based times 2,4,6, where they are $(2/11, 5/11, 9/11)$.

Reward $(1, 1/2, 0)$ and threshold $1/4$ define the order terminal. The selector uses $d_{j-1,j} = x_j/3$ and diagonal remainder. Exact route LPs compare against the whole pairwise-optimal face. Across all 24 programs and both costs, eight combinations improve and 40 do not; the eight-stage unit-cost optimum is $372061/49193760$ at $(181/385, 204/385, 1)$. The best shared value on the pairwise-optimal face is $279/42592$ at $(1, 0, 1)$, giving gain $6227/6149220$. At $n = 4, r = 3$, the unit-cost gain is approximately 0.01013.

Block/descending is 0.9928, descriptive interval $[0.9768, 1.0057]$, 5/24 lower; the six multi-vector/18 singleton ratios are $0.9491/1.0078$. The singleton group has 17 non-faster cases; all individual costs are retained. Separate resident-use processes evaluate and recover words for 24 fixed entrances and replay their full SCMs. Per-query medians span 1.23–37.83/0.26–1.37 μs for evaluation/recovery and 0.11–0.34/0.35–0.93 ms for materialization/full replay. Cold includes import, decoding, certificate checking and word replay. The adaptively motivated grid entered the declared matched study in full.

L.3 Repeated exact route design

The same eight $n = 8, 12$ order models each receive twelve supplied contracts: unit costs at budgets $2, 1, 3/2$, then each permutation of $(3, 1, 1)$ at budgets $2, 5/2, 3$, in that order. $K = 1$ is the first contract; $K = 12$ is the full sequence. The complete-function method compiles once. The direct adapter shares original menus, support preparation and feasible-word cuts across tasks. A master LP uses these cuts as global upper planes; an exact prefix query at its chosen route gives a lower design value. Matching rational bounds prove global optimality; a violated attained word supplies the next cut. Master primal/dual and lower-query certificates are checked after readback; shared cuts and compiler coverage are checked once, with one full SCM replay per task. Both methods use the same four-variable exact vertex LP. Small-word exhaustion and corruption tests precede timing; direct solving reads no compiled family or saved answer.

The fixed 96-job campaign uses three fresh processes per cell, the preceding resource limits and first-clean rules. All 96 finish exactly without retries or timeouts. Table 3 includes every input. On D batches, direct search retains 5–8 cuts and uses 16–19 oracle calls; the complete families have 11–21 vectors. Resident master/search work takes 0.185–0.597s versus 1.204–9.837s, while preparation takes 0.009–0.010s versus 1.133–1.276s. These are phase ranges, separate from the cold endpoint. The LP's constraint count affects this implementation comparison. The measured contracts favor task-specific separation, while the representation theorem guarantees an explicit envelope for all entrances.

Across all 96 model/contract pairs, 23 sharp-optimal routes improve over the best shared value on the entire pairwise-optimal face; 73 tie. Positive gains range from 1.64×10^{-6} to 0.001013, or 0.676–33.416% relative to that face's best shared value (all denominators are positive). For $n = 8$, 13/48 improve; for $n = 12$, 10/48 improve with smaller absolute gains 1.64×10^{-6}–3.82×10^{-5}. Exact primal/dual certificates cover all three design LPs per case.

L.4 Holding structure and persistent corrections

A specified transition program with one exceptional stage. A fair D layer (omissions 012) precedes calibration $(121; 1/5, 2/5, 3/5)$. Each of $L = n - 2$ S holding stages has successors $\{0, 1\}$ and probabilities $(\epsilon, 1 - \epsilon, 1/2)$: binary labels are retained with error ϵ, and label 2 initializes fairly. Thus $r = 1$ for every duration. Reward $(1, 0, 1/2)$ and threshold $3/4$ give $g = e_{01}$; the selector is as above. The commands specify marginals, leaving their shared-response coupling unknown.

Why the signal persists. Put $u = (1, 1/2, 0, 0, 0, 1/2)$ and $a = (1 - 2\epsilon)^L$. A common uniform threshold coupling (Fill & Machida, 2002) has holding-row atom weights $(\epsilon, 0, 1/2 - \epsilon, 1/2 - \epsilon, 0, 0, \epsilon)$ in binary order. Its backup of either e_{01} or u is $(1 - 2\epsilon)u$, the coordinatewise Fréchet minimum. Nonnegative composition therefore gives the attained all-entrance holding continuation au. For the two initial stages, exact original-domain enumeration and rational coverage certificates give

$$H_0 = \frac{a}{20}\{(0, 0, 3, 3, 2, 2),\ (0, 2, 3, 2, 4, 1),\ (1, 0, 4, 2, 2, 1)\}, \qquad \phi_0 = \frac{a}{20}(0, 0, 3, 2, 2, 1).$$

The maximum normalized gap is $a/50$: the three residual sums are $a(d_{12} + d_{21})/20$, $a(d_{02} + d_{20})/10$ and $a(d_{01} + d_{10})/20$. Their maximizing group masses are $2/5, 1/5, 2/5$; a convex combination with

Table 2: The six non-singleton programs, joined by matched input identities. Δ is maximum normalized gap; Δ_H/Δ_U lists heterogeneous and unit-cost route gains, separated by a slash. G/D/B are original generic/descending/block median seconds (five processes). P/B is the separate matched precheck/block ratio (three processes, lightweight imports, 24 SCM queries). Time cohorts are not pooled. The other 18 programs are pairwise exact; 16 benefit from prechecking.

n	r	$\lvert H\rvert$	Δ	Δ_H/Δ_U	G (s)	D (s)	B (s)	P/B
4	2	9	0.03386	0/0.00537	1.06	1.05	1.05	0.999
4	3	7	0.02519	0.00067/0.01013	1.05	1.04	1.04	0.992
8	2	12	0.00181	0/0.00055	1.16	1.15	1.11	1.003
8	3	21	0.00535	0.00012/0.00101	1.38	1.34	1.28	0.993
12	2	11	0.00011	0/0.00002	1.26	1.25	1.11	1.003
12	3	16	0.00024	0/0.00003	1.50	1.45	1.29	0.998

Table 3: Repeated-design cold seconds: ranges of three-process input medians; ratios are geometric means of Direct separation/(Compile + LP) over four inputs per group. K counts supplied contracts. All outputs are exact.

Group	K	Direct separation	Compile + LP	Ratio
D	1	0.106–0.182	1.372–2.249	0.0787
D	12	0.336–0.852	2.510–11.288	0.1120
X	1	0.083–0.089	1.179–1.196	0.0726
X	12	0.185–0.214	1.214–1.237	0.1569

the same weights gives the matching upper bound. A fixed total noise budget $\epsilon = 1/(100L)$ gives $a \geq 49/50$ by Bernoulli's inequality, hence gap at least 0.0196. A fixed per-stage error instead gives $a = (0.98)^L$ and decaying signal. The duration dependence is an elementary holding calculation; the compiler still checks the complete original domain.

For costs $(3, 1, 1)$ and budget 3, the shared objective is

$$\frac{a}{60}\min\{2x_0 + 3x_2,\ 4x_0 + 2x_2,\ 2x_0 + x_1 + 2x_2\}.$$

Its optimum $3a/50$ occurs at $(2/5, 4/5, 1)$: weighting the second and third forms by $1/5, 4/5$ bounds their minimum by $\frac{4}{5}(3x_0 + x_1 + x_2) + \frac{6}{5}x_2 \leq 18/5$. The pairwise objective $a(x_0 + x_2)/30$ uniquely selects $(2/3, 0, 1)$, whose sharp value is $a/18$. The design gain is therefore $a/225$. Unit costs with budget 2 produce no design gain.

Joint evidence and controls. The $n = 128, 256, 512$ study uses both noise rules and both calibrations. The control $(012; 2/11, 5/11, 9/11)$ gives six exact singletons; six positive calibrations have three values and positive heterogeneous gains, while all 12 unit-cost gains vanish. Block/descending geometric ratio is 0.7797, descriptive paired interval $[0.7000, 0.8540]$, 12/12 lower; the positive/null groups give $0.8175/0.7436$. Block/ascending is 0.5884. The separate $n = 4, 16, 64$ study has ratio 0.9914, interval $[0.9822, 1.0002]$, 8/12 lower. Both studies retain all twelve adaptively selected inputs.

Several calibration decisions. The spaced and contiguous calibration families retain three-vector envelopes. The twelve contiguous programs contain six nonproportional cores, rescaled by holding duration; their adaptive construction and full controls are recorded. Both prior route costs give 14 positive and 10 zero gains. Matched complete-function costs give block/descending 0.9884, interval $[0.9807, 0.9964]$, 10/12 lower; The artifact reports complete cold and resident costs.

L.5 Additional contract controls

The eight fair-exclusion horizon/cost combinations in Appendix D.2 have three positive odd-horizon gains and five zero gains; all exact LPs remain in the artifact. The mixed five-stage X/S/D/X/E control has a singleton value $(19/2000, 0, 113/6000, 113/6000, 0, 19/2000)$ equal to its pairwise vector. At

$d_{10} = d_{21} = d_{02} = 1/3$, both values are $17/1800$. The model-source map supplies its complete specification and certificate.

L.6 Shared-response completion of a public token protocol

Source dynamics and exact state correspondence. The Israeli–Jalfon protocol moves a token to either neighbor with equal probability (Israeli & Jalfon, 1990; PRISM Project, n.d.). The PRISM case study explicitly identifies its stable phase: one privileged process circulates the token indefinitely (PRISM Project, n.d.). We use the unchanged three-process model distributed by the Quantitative Verification Benchmark Set (Hartmanns et al., 2019), with both its guarded-command and independently converted JANI representations. Writing the Boolean token variables as (q_1, q_2, q_3), the first command is

$$q_1 = 1 : \quad \tfrac{1}{2}(q_1' = 0, q_3' = 1) + \tfrac{1}{2}(q_1' = 0, q_2' = 1);$$

the other commands are cyclic renamings. Among seven reachable nonempty token configurations, the closed stable set $C = \{(1,0,0), (0,1,0), (0,0,1)\}$ contains exactly one enabled command per state. The bijection $\iota : C \to \{0,1,2\}$ records the unique token position. Each state moves to the other two with probability $1/2$, so its omissions are $(0,1,2)$ and every layer is D-fair. This is a native $r = 0$ family with $O(n)$ explicit input bits and constant probability precision. The constant auxiliary variable and single control location are retained in this full-state bijection.

The completion query. The public source specifies single-world transition marginals. Our new paired-world query starts two copies in supplied stable states and advances one protocol step per layer with a shared complete response mechanism. We optimize over all complete laws with the source marginals, independently and untied across layers. Let $h = (1, 1/2, 0)$ and $\eta = 1/4$; the terminal event compares the final token positions through $\mathbf{1}\{h(i) - h(j) \geq \eta\}$.

The core and full source-table completion queries are equivalent: Restricting a full-table response law to the three closed stable rows yields a feasible core law. Conversely, any core law extends to a full-table law by independently adjoining responses for all other rows with their specified marginals. Timewise extension preserves the product domain; closure and the bijection preserve entrances, terminal events, pairwise backups and chronological values. Thus restriction and extension preserve the sharp minimum and attainment. Shared responses preserve the diagonal.

Exact evaluation. The predetermined cohort uses horizons $1, \ldots, 8, 16, 32$ and nine entrances per horizon: the six oriented basis pairs, the uniform clockwise and reverse cycles, and the uniform off-diagonal entrance. Both complete-function compilers return three attained values at each horizon; direct prefix search agrees on all 90 exact minima. There are 20 positive pairwise gaps and 70 equalities. At $n = 1$, the clockwise, reverse and uniform entrances have pairwise values $(1/6, 1/3, 1/4)$ and common sharp value $1/3$; these illustrate the complete cohort. The gaps decay with horizon, reaching $1/12884901888$ and $1/25769803776$ on the reverse and uniform entrances at $n = 32$.

Exact PRISM/JANI comparison covers all 32 declared valuations, reconstructing seven reachable states and twelve enabled state–action rows. Original-word enumeration through horizon four checks the scalar answers; serialized coverage certificates, direct-search proofs, chronological replay, support recursion and nine-coordinate diagonal checks validate the outputs. Runtime comparisons use the separate cohorts in Section 6.

M Stability through a compressible reference

M.1 Certificates from nearby compressible structure

The refined solver and complete-atom marginal transport yield a certified interval for the original minimum. The Bernoulli channel and mismatch argument use standard coupling ingredients (Sason, 2013, Theorems 1–2) on the untied response-law domain.

M.1.1 Same-support transport and the enclosure

Fix the supports, actual row identities, scheduled actions, probability-one factual wrapper, and independent untied complete-response laws of Section 2. Fix a normalized joint entrance d (including

any zero-cost diagonal mass) and a supplied rational terminal with $0 \le g_{ij} \le G$, $g_{ii} = 0$, and rational $G \ge 0$. Let $p, \bar{p}$ be strictly interior rational core probability tables on these same supports. Write $\bar{r}$ for the reference exception count. The bit length B_* includes both probability tables, their common structural input, the supplied terminal, G, and the entrance. Define

$$c_t = \min\left\{1, \sum_{i=0}^{2} |p_{t,i} - \bar{p}_{t,i}|\right\}, \qquad \Gamma = 1 - \prod_{t=0}^{n-1}(1 - c_t). \tag{48}$$

Corollary 19 (Nearby-compressible-reference certificate)**.** *Put $\bar{L} = L_{\bar{p}}^{g}(d)$. There is a transported reference optimizer M_{lift} feasible for the original complete-response domain such that*

$$\max\{0, \bar{L} - G\Gamma\} \le L_p^g(d) \le J_p^g(M_{\mathrm{lift}}; d) \le \min\{G, \bar{L} + G\Gamma\}. \tag{49}$$

Thus $|L_p^g(d) - L_{\bar{p}}^g(d)| \le G\Gamma$; the displayed enclosure has width at most $2G\Gamma$, as does the enclosure with lifted value as its upper endpoint, and the lifted witness is at most $2G\Gamma$ suboptimal. The inequality only requires bounded costs. If additionally $g \in \mathcal{C}_1$ and $\bar{r} \le s$ under Theorem 4's structural assumptions, reference optimization, transport and checking take $O(490^{2s}(n + B_* + 1)^{12})$ bit operations under that theorem's implementation bound. The original exceptional count may be large. For a nonnegative unnormalized entrance of mass $m = \|d\|_1$, the same statements hold with G replaced by mG in every range and error bound, with both objectives evaluated on d.

Proof. For current bit marginal $a \in (0,1)$ and target $b \in (0,1)$, flip 0 to 1 with probability $(b - a)/(1 - a)$ when $b \ge a$, or 1 to 0 with probability $(a - b)/a$ otherwise. The unconditional change probability is $|b - a|$. Apply these channels coordinatewise to a complete atom, with independent channel randomness. Other marginals are preserved; aggregating outputs gives a normalized law on the same eight atoms with the target marginals.

The original and transported complete atoms disagree with probability at most c_t. Couple these atom pairs independently across time and retain the same selector, routing and entrance draw. With probability at least $\prod_t (1 - c_t)$, all core atoms agree, so both query trajectories and their terminal costs agree. Otherwise the two costs, both in $[0, G]$, differ by at most G. Hence every transported complete-law word changes the objective by at most $G\Gamma$. A complete atom is shared across worlds throughout; no channel depends on the visited pair.

The compact product of response-law polytopes has optimizers. Transporting a reference optimizer to p gives a feasible original witness with value at most $\bar{L} + G\Gamma$. Transporting an original optimizer to $\bar{p}$ gives $\bar{L} \le L_p^g(d) + G\Gamma$. Nonnegativity and the range $[0, G]$ prove the enclosure, absolute error, width and suboptimality bounds. For $m > 0$, apply this argument to d/m and rescale by m; at $m = 0$ all objectives vanish. At $G = 0$ the terminal is zero and all objectives vanish, without scaling by G. Shared maps preserve equality of the two states, so diagonal mass always has zero cost. Deleting that mass yields the six-coordinate substochastic representation with the same objective; it must not be renormalized after a transition.

Transport uses at most 8×8 rational atom entries per time. The channels, cost evaluations and products have polynomial encoding lengths: denominator lengths add through time, and rational multiplication by G and m costs polynomially in their binary encodings. Restoring deterministic rows and chronological factual/query/query replay add polynomial overhead to reference construction and checking. There is no iteration over numerical magnitudes, inverse probabilities, or G; events specialize to $G = 1$. □

A separately valid original pairwise bound strengthens the left endpoint to $\max\{0, \bar{L} - G\Gamma, L_{\mathrm{pair}}^g(p; d)\}$ (use $mG\Gamma$ for mass m). The interval encloses the *minimum*, not the sharp compatible range $[L(p), U(p)]$; the lifted SCM need not attain that minimum.

M.1.2 Recognizing a low-exception reference

At a fixed layer the smallest sum of changes to the sorted second-successor probabilities that makes the layer compressible is

$$e_t = \begin{cases} \sum_i |p_{t,i} - 1/2|, & \text{D supports,} \\ |p_{t,i} - p_{t,j}|, & \text{X supports, duplicate rows } i, j, \\ 0, & \text{S supports.} \end{cases}$$

For D the fair table is the only compressible table on those supports. At X, setting both duplicate probabilities to any point between them achieves the triangle-inequality lower bound $|p_{t,i} - p_{t,j}|$; we choose their rational average and leave the singleton row unchanged. At S no change is needed. These projections stay strictly interior and have polynomial encoding length in the original input.

Keep at most s positive-cost layers unchanged, choosing those with the largest capped costs $\min(1, e_t)$ and resolving ties by time index. Project the others as above. All projected layers are compressible, so $\bar{r} \le s$. A pairwise exchange shows that retaining a larger cost rather than a smaller one weakly increases $\prod_t (1 - c_t)$; the rule therefore minimizes (48) among these local projection choices with this budget, including the case of unit costs. It does not minimize the actual interval width, reference risk, or running time. Writing the positive costs in descending order gives $\Gamma_s = 1 - \prod_{j>s}(1 - c_{(j)})$. The reference-optimum error is at most $G\Gamma_s$ and both enclosure width and transported-witness suboptimality are at most $2G\Gamma_s$. Hence $\min\{s : 2G\Gamma_s \le \tau\}$ is the smallest budget meeting this sufficient bound within the prescribed projection scheme. For mass m, the criterion is $2mG\Gamma_s \le \tau$; $\tau \ge 0$ is measured in cost units. At $mG = 0$ every objective is zero and budget zero suffices.

Nonfair matched X means and singleton probabilities are preserved. The guarantee uses known marginals and fixed supports; Γ can exceed the signal of interest. For normalized entrances the range is $[0, G]$; for mass m it is $[0, mG]$. The saved studies below use indicator costs, so their unit-scale values, margins and transported witnesses are unchanged. For a general bounded cost the sufficient pairwise-improvement test is $\bar{L} - L^g_{\mathrm{pair}}(p; d) > mG\Gamma$.

Information gained over pairwise. The 42-case D/X/S grid uses $n \in \{1, 4, 16\}$, changes $\epsilon \in \{1/1000, 1/10\}$ per layer or divided by n, six X-preserving controls and reference budget zero. Exact original minima are available for the 28 cases with $n \le 4$; all lifted laws are feasible and checked. The artifact supplies the complete channel, projection and pairwise-recursion checks. Only D at $n = 1, 4$, $\epsilon = 1/1000$ gains under both budget modes ($n = 1$ duplicates); the other 38 do not. At $n = 4$, gains are 0.002357/0.008397; combined lower/feasible upper/exact values are 0.033691/0.042042/0.041792 and 0.039668/0.041760/0.041698. All 42 cases remain in the supplement.

Holding perturbations. Holding remains S under strict fixed-support changes, so $r \le 2$. The 12 longer programs receive $\pm\beta(1, 0, -1)/(2n)$ at each layer, $\beta \in \{0.001, 0.01, 0.015\}$. At $d_{01} = d_{12} = 2/5, d_{20} = 1/5$, 20/72 cases improve pairwise: all 18 total-noise positive calibrations and two fixed-noise cases; all 36 null controls tie. Total-noise gains range 0.01857–0.01863,0.00936–0.00995,0.00427–0.00515; width/reference-to-original- pairwise margin is 4.9–5.3%,48.1–53.6%,71.4–80.8%. All original/holding ratios, undefined nonpositive margins and complete widths remain in the artifact.

Breaking every temporal structural equality. All 24 nonholding programs receive $\pm\beta(1, 0, -1)/(2n)$ at every layer, for $\beta \in \{10^{-5}, 10^{-4}, 10^{-3}\}$: all 144 remain interior and have $r = n$ (reference $r = 2, 3$). For both route contracts, fix the reference shared-optimal x_S and best shared x_P on the reference pairwise-optimal face. Use the combined lower and smallest transported-word upper at each route. Positive $\mathrm{lower}(x_S) - \mathrm{upper}(x_P)$ certifies this prescribed preference, without asserting perturbed pairwise optimality of x_P. At increasing budgets, improved pairwise lowers number 22/20/14 of 96, certified preferences 16/12/6, and width/reference-correction ranges are 0.05–20.93%,0.47–100.03%,4.73–100.34%. Widths divide by positive reference corrections at x_S; zero denominators remain undefined. All 16 signed positive reference-route comparisons survive 10^{-5}; six survive 10^{-3}. The other 254/288 are inconclusive, including all singleton controls. The artifact retains every interval, transported coupling and exact check.

Exact perturbed route optima at four stages. All 48 four-stage perturbations are solved by the unchanged generic compiler, with complete-domain certificates and chronological replay. Both route contracts are optimized over the true perturbed family and over the entire perturbed pairwise-optimal face. Of 96 comparisons, 18 have positive design gain and 78 tie. The reference lower at x_S exceeds the maximum transported-word upper over that new face in 16/18 positive cases: 6,6,4 at increasing budgets; two largest-budget cases remain uncertified. The reference x_S is exactly optimal in 60/96 cases; its maximum loss is $1.395 \times 10^{-7}, 1.395 \times 10^{-6}, 1.395 \times 10^{-5}$ at the three budgets. These upper-rounded losses compare actual optima, not just two fixed routes. Primal/dual LP certificates and

transported-word checks are supplied. The budgets are declared parameter tolerances, not sampling confidence levels.

N Complete-function and supplied-query computation

N.1 Complete-function cost and avoided work

The current-source 72-input confirmation uses five serial GMP processes per method/input, guest CPU 2, seed 20260925, first-clean selection and one permitted contamination/interruption retry. Unknown noise is ineligible; guest affinity does not imply physical exclusivity. Both 72-input bootstrap intervals use 10000 paired input draws with protocol-recorded seeds and quantiles. Seven confirmation inputs are slower: six D and one S, with ratios 1.0035–1.0452. The separate 24-input transfer grid at horizons 48/192 gives ratio 0.7537 and D/X/S ratios 0.6095/0.8956/0.7843; all four D-edge ratios 1.0026–1.0246 are slower.

Figures 5–7 show historical ratios (Appendix H), historical/current costs and the separate instrumented pass reproducing the original solves exactly. Shared backup counts include proof work. All 96

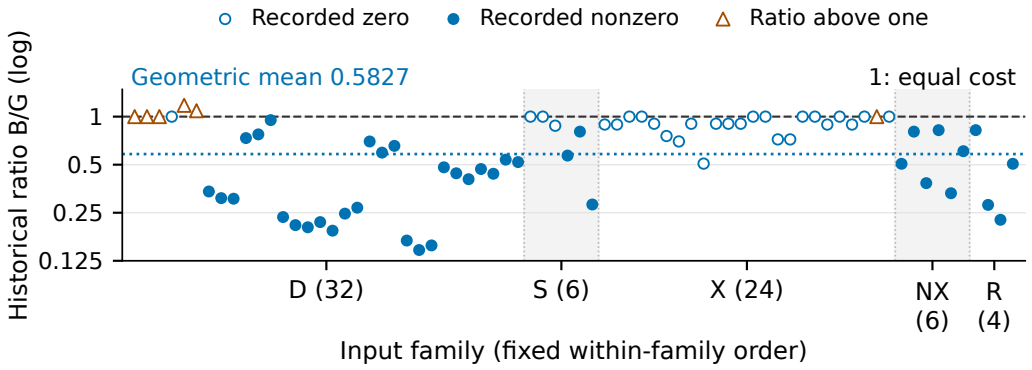


Figure 5: Historical complete computation cost on all 72 primary inputs. Each point is the ratio of the block and generic median complete wall times across five repetitions per method. Open markers denote recorded zero functions; filled circles denote recorded nonzero functions. Open triangles identify the six ratios above one, all in the recorded zero group; ratios below one favor block. Every input appears once. Parentheses give family sizes; NX denotes the six-input nonzero all-nonfair X family, not the 39-input nonzero subgroup; R denotes reachable nonmetric. The geometric mean is 0.5827.

pruning-order checks agree. Against descending-order generic pruning, block uses fewer/equal/more LPs on 4/16/4 inputs. Separate 180-run timing gives ratio 0.7041, interval $[0.5296, 0.8954]$ and 9/12 lower block medians, with three slower D-edge ratios 1.0068–1.0230. Full-entrance gaps are positive on 36/72 original and 4/24 transfer models, at most 2.28×10^{-5} and 4.57×10^{-15}.

Survival bounds on corrections. For feasible q, let s^q be its terminal noncoalescence vector. Zero diagonal cost gives $v^q \le \|g\|_\infty s^q$; with $\phi \ge 0$, normalized entrances satisfy $0 \le G(d) \le d^T(v^q - \phi) \le \|g\|_\infty d^T s^q \le \|g\|_\infty \max_e s^q_e$. Entrance mass m multiplies the final bound by m; diagonal mass contributes zero without renormalization, and zero cost or mass needs no division. Recorded cohorts use indicator costs, leaving their numbers unchanged. Mean-original-vertex words bound corrections below 0.01 on all 96 original/transfer models and below 0.001 on 89 (transfer maximum 6.00×10^{-7}); bounds are below 0.01 on 10/24 temporal and 4/12 long holding models. Exact checks cover all 132 laws, values, survival vectors and Fréchet recursions.

N.2 Support diagnosis and supplied accuracy contracts

Eight $n = 4$ programs use $K = 1, 6, 24, 96$: coordinate entrances followed by normalized $(1, j, \ldots, j^5)$, $j \ge 2$. Direct streams 1296 original words, block compiles, and hybrid tests full

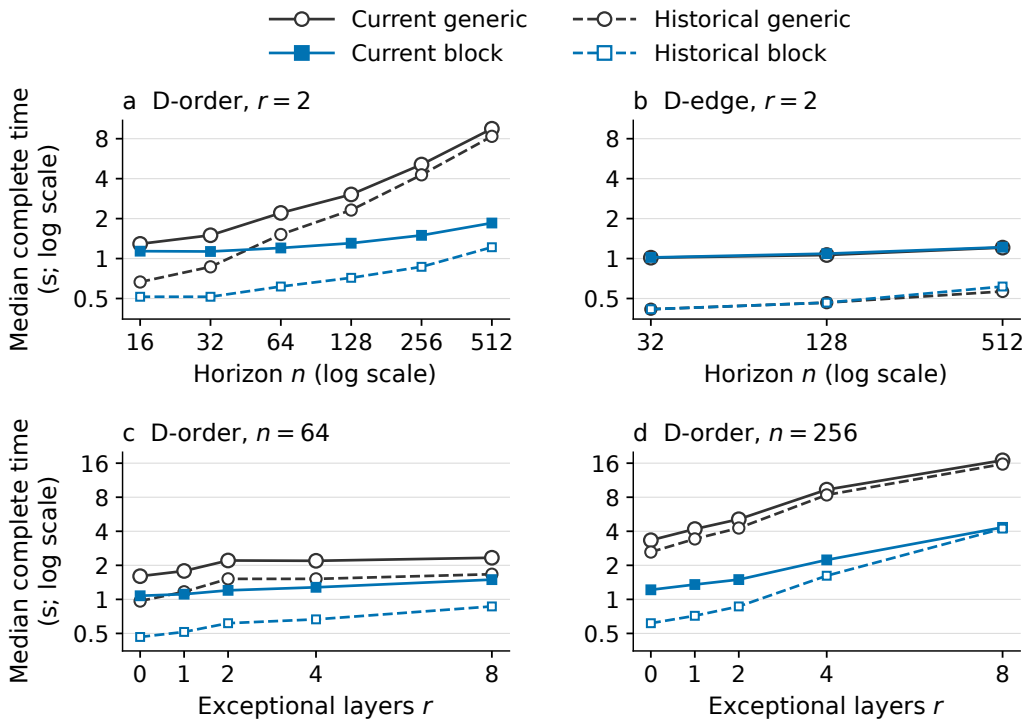


Figure 6: Controlled D slices with dispersed placement and generation parameter 3. Top: fixed terminal and $r = 2$, varying horizon. Bottom: fixed D-order terminal and horizon, varying r. Dashed historical and solid current curves are separate cohorts. The D-edge slice retains the unfavorable long-horizon behavior.

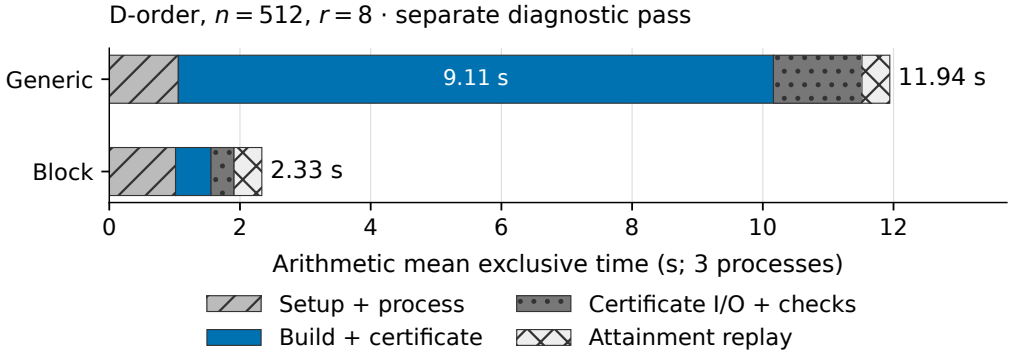


Figure 7: Exclusive phase means on one diagnostic input sum to the arithmetic mean parent total. Setup + process combines imports, parsing/menus and startup/output residual; certificate I/O + checks combines serialization/reload, coverage checks and typed diagnostics. Construction includes interleaved proof generation. Wrapper overhead belongs to this separate diagnostic pass; the primary endpoint uses uninstrumented delivered solvers.

support before compilation. All 288 runs share imports, timing rules, coverage/attainment checks and requested SCM replay; ratios geometrically average input ratios of three-process medians. At $K = 1$, accepted/rejected hybrid medians are 0.067–0.069/1.244–1.262s versus direct 0.745–0.808/0.756–0.787s. Hybrid/block ratios on rejection are 0.9986,1.0018,1.0007,0.9995; failed tests cost 5.84–5.91ms. Hybrid is slower in 6/32 cells.

Precheck across horizons. Across all 24 programs and 24 queries, hybrid/block ratios are 0.4274 overall, 0.3221 on 18 accepted and 0.9980 on six rejected programs, lower on 16/18 and 4/6. Nine acceptances avoid optimizer import. Complete costs and ratios, including all four slower cases, are in the artifact.

Query supports and accuracy contracts. All eight order-terminal $n = 8, 12$, $r = 2, 3$ programs enter a separate 192-run comparison: four multi-vector D cases and four singleton X controls, $K = 1, 24$, and declared tolerances $\tau = 0, 0.001, 0.01$. The policy first recovers accepted query-support words, then tests the mean-vertex-word interval; any unresolved query triggers one full compilation. Three processes per cell use the matched endpoint above, charging failed work and certificate/SCM checks; full compilation is shared across tolerances. At $\tau = 0$, basis query e_{01} is accepted on 2/4 D inputs; policy/full ratios are 0.2270 for one query and 1.0043 for 24. All four D batches still compile at $\tau = 0.01$: no mean-word interval resolves an additional query. Their batch ratios across tolerances are 0.991–1.013. All X queries are exact from support words; ratios are 0.058–0.060 for $K = 1$ and 0.126–0.131 for $K = 24$. The policy is slower in 8/48 input/batch/tolerance cells, with 18/48 fallbacks. All 48 cost rows remain in the artifact. Tolerances are declared sensitivity levels, not application-derived requirements.

N.3 Exact fixed-entrance search and matched comparison

For $d \geq 0$, propagate $x_0 = d$, $x_{t+1} = K_t(q_t)^T x_t$ and use lower bound $x_t^T \phi_t$. Search original vertices in successor-bound order, breaking ties by menu index. A greedy feasible word initializes the incumbent; accepted supports close with attaining suffixes. Prune bounds at least the incumbent and reuse closed identical (t, x_t) states within a query. Menus, relaxed values and support recovery are shared across queries. This is rational branch-and-bound (Lawler & Wood, 1966).

Optimality certificate. Original-vertex attainment reduces the domain to a finite word tree. Every node records a complete menu branch, valid bound prune, attained support closure, terminal, or closed-state reference. The checker reconstructs branches and masses from response atoms and independently evaluates suffix Fréchet bounds. Support closures have zero weighted residual throughout. A closed state certifies every branch against the incumbent at closure. Since incumbents decrease, exact-state reuse preserves the final lower bound, including pruned leaves. Only closed states are reusable. Induction over the checked tree shows every completion has value at least the final incumbent; a chronological feasible word attains that value. Generator and checker share rational domain primitives and independently implement mass propagation and bound generation.

At resource limits, pending-node bounds and the incumbent enclose the answer and determine completion status. Queries share a soft 90s deadline, with a 200000-node expansion threshold per query; pending fringe nodes are recorded. Process limits are 180s wall, 150s CPU and 8GiB. Every certificate is serialized and read back; each requested SCM is replayed. Validation checks 16 inputs against certified functions and exhaustively enumerates eight four-stage cases with compiler entry points disabled. Tests cover corruption, chronology, ties, zero mass and resource stops.

Matched measurement. Eight order programs ($n = 8, 12$, $r = 2, 3$), three methods and $K = 1, 24$ use three fresh processes per cell. One-query runs use the first declared mixed entrance; batches contain six basis and 18 mixed entrances. Block and precheck–block use the production compiler; accepted full support returns a singleton. Cold launch-to-exit includes preparation, generation, serialization/readback, checks and SCM replay. Compiler phase bins include lazy numerical imports, precluding isolated search timing. GMP, CPU placement and thread controls match the prior study. The 1800s budget uses first-clean selection and one contamination/interruption retry; all attempts are retained.

All 144 runs complete exactly; the ledger includes one contaminated attempt and its prescribed retry. Table 1 aggregates every input. Precheck–block is slower on one D input at each workload; at $K = 24$ it beats direct on all X inputs (one-query ordering mixed). D batch proofs contain 1565–4805 bound prunes, 99–490 exact-state reuses and 0–9 support closures; X closes at the root. Direct D certificates are 5.5–9.2kB and 101.1–180.1kB, versus block's 27.5–79.7kB and 38.8–95.5kB for the two workloads.