Classical Extensions and Strongly Positive Realizations of Decoherent History Probabilities

Zhenyu Sun

School of Mathematics and Statistics, Northeastern University at Qinhuangdao, Qinhuangdao 066004, China

hexagon:2611.00094v1

Abstract

Probabilities on several decoherent families of histories may agree on shared events yet admit no classical extension to the Boolean algebra that the families generate. We characterize the strongly positive decoherence functionals that realize given local probabilities on a finite algebra. A covariance decomposition allows any signed extension to be prescribed at any attainable rank and gives the resulting increase in minimum trace. For two contexts, agreement on their intersection suffices for a strongly positive realization. On the bipartite graph of context atoms, the minimum negative mass of a signed extension is $(W_1 - 1)/2$, where W_1 is the transport cost for graph distance. A graph Laplacian formula gives the minimum trace. Its minimizers form an affine copy of a rectangular contraction ball whose extreme points have the smallest realization rank. For a pure-state implementation with one intermediate projective decomposition and a final postselection, the optimal success probability of a realization is the reciprocal of the least trace of a diagonal matrix dominating it. In a family of classical examples, this probability tends to zero for the unique trace minimizer even though another realization succeeds with probability one.

Keywords. Decoherent histories, probability extension, strong positivity, signed measures, optimal transport, effective resistance, contextuality.

1 Introduction

A decoherence functional assigns probabilities to events in each medium-decoherent family of histories. Two such families agree on the probabilities of their common events. Does a single probability measure on the Boolean algebra they generate reproduce both sets of probabilities? Given only the local probabilities, which strongly positive decoherence functionals reproduce them and make each family medium decoherent? The distinction already arises for two binary contexts on three atoms. If each context assigns probability one to a different singleton, the two marginals agree on their common events. No probability measure can give unit mass to both singletons, but a rank-one decoherence functional realizes the two marginals (Section 8.2).

The classical probability rules within a consistent set of histories are well established [1, 2]. Isham studied presheaves on the partially ordered set of consistent Boolean subalgebras [3]. Halliwell asked whether incompatible consistent sets admit a *unifying probability*, and used Bell and Leggett–Garg inequalities to test its existence [4]. A common decoherent refinement supplies such a probability. The converse can fail even for histories of a single qubit, as the class operators in Section 8.5 show.

In the approach of Abramsky and Brandenburger, a noncontextual model is a probability distribution on global measurement assignments with the prescribed local marginals [5]. For

subalgebras of a fixed history algebra, one must also specify which global assignments the algebra permits. A choice of one atom from each context can agree on every overlap while the chosen atoms have empty intersection. The two deterministic singleton outcomes above give an example. We relate this distinction to the amalgamation of Boolean subalgebras [6, 7] and show exactly when the two assignment spaces yield the same classical probability polytope (Proposition 2.4).

On the quantum side, Dowker, Henson, and Wallden studied joint quantum measures whose restrictions to measurement contexts are classical, and related strong positivity to bounds on quantum correlations [8]. Strong positivity also yields Tsirelson bounds [9]. Hilbert-space representations of decoherence functionals and the role of strong positivity in the histories formulation are discussed in [10, 11]. Here we fix the local probabilities and determine which atomic decoherence matrices realize them, how small their rank and trace can be, and with what probability they can be implemented by postselection.

Let $\mathcal{M}$ be the family of contexts and X the set of atoms of the finite Boolean algebra they generate. For $C \in \mathcal{M}$, let P_C sum coordinates on X over each atom of C, and let e_C be its probability vector. A signed extension is a real vector q satisfying $P_C q = e_C$ for all C. It is a classical extension if $q \geq 0$. A strongly positive realization is a Hermitian matrix $H \succeq 0$ satisfying

$$P_C H P_C^{\mathsf{T}} = \operatorname{diag}(e_C) \qquad (C \in \mathcal{M}). \tag{1.1}$$

The diagonal entries give the context probabilities, and the off-diagonal equations express medium decoherence. Summing both indices gives $\mathbf{1}^* H \mathbf{1} = 1$. Each entry of $P_C H P_C^{\mathsf{T}}$ sums the entries of H over a pair of context atoms. Quantum coupling problems instead prescribe partial traces [12].

For any finite family of contexts, every realization has a unique decomposition

$$H = zz^* + K, \qquad z = H\mathbf{1}, \qquad K \succeq 0, \quad K\mathbf{1} = 0.$$

The equations for z are $P_C z = e_C$. Those for K prescribe the local covariances $\operatorname{diag}(e_C) - e_C e_C^{\mathsf{T}}$. Replacing z by a real signed extension q preserves the rank of H. Prescribing $H\mathbf{1} = q$ raises the minimum trace by $\|q - q_0\|_2^2$, where q_0 is the signed extension of least Euclidean norm (Theorems 3.1 and 3.3).

For two contexts, put an edge between two context atoms whenever they intersect. The edges of the resulting bipartite graph G are exactly the atoms of the generated algebra. Overlap agreement suffices for a strongly positive realization, and the smallest rank is the larger of the two marginal support sizes (Theorem 4.2). A classical extension must also satisfy the supported transport inequalities in Proposition 4.5. For compatible marginals p, q, let $\eta_G(p, q)$ be the minimum negative mass of a signed extension, a quantity used in studies of contextuality [13]. We show that

$$\eta_G(p, q) = \frac{W_1^G(p, q) - 1}{2},$$

where the cost of moving a unit of mass between vertices is their graph distance. Together with strong positivity, this identity gives bounds on negative real interference and trace that are attained on families of odd paths (Theorems 5.2 and 5.4 and Proposition 8.2).

Minimizing the trace uses the electrical minimum-energy principle [14] and a singular-value optimization of the type that occurs in the orthogonal Procrustes problem [15, 16]. For deterministic endpoints, the minimum trace is their effective resistance. For general compatible marginals, the minimizers are parametrized by a rectangular contraction ball, including when the graph is disconnected. Its extreme points attain minimum rank and minimum trace simultaneously. When every marginal entry is positive, the rank of the

bipartite adjacency matrix determines the entire interval of minimizing ranks (Theorems 6.1 and 6.3 and Corollaries 6.2 and 6.4). For coordinate covers of full product spaces, the minimum trace is the fraction of Fourier characters visible in a context (Theorem 6.5).

We also consider pure-state histories with one intermediate projective decomposition and a selected final event. This is a finite instance of the histories formalism with initial and final conditions [17]. A Gram-matrix construction [18] gives the exact success criterion. For fixed H, the largest success probability is $1/\Gamma(H)$, where $\Gamma(H)$ is the least trace of a diagonal majorant of H. After trace normalization, this optimization is related to the robustness of coherence [19, 20]. Optimizing over all realizations gives success probability one exactly for classically extendable data (Theorems 7.1 and 7.2). The asymmetric three-atom example gives the exact gap in the bound from negative mass. A separate family shows how far the trace minimizer can fall short of optimal success even for classical data (Propositions 8.1 and 8.3).

2 Finite history algebras and their extensions

2.1 Decoherence functionals and contexts

Let $X = \mathrm{At}(\mathfrak{G})$ be the atom set of a nontrivial finite Boolean algebra $\mathfrak{G}$. Each event is the union of its atoms, so we identify $\mathfrak{G}$ with $\mathcal{P}(X)$. Fix a nonempty finite family $\mathcal{M}$ of Boolean subalgebras, called contexts, that generates $\mathfrak{G}$. For $C \in \mathcal{M}$, write $X_C = \mathrm{At}(C)$ and let

$$r_C : X \longrightarrow X_C$$

send each ambient atom to the unique atom of C that contains it. The matrix that sums over the fibers of r_C is

$$(P_C)_{b,a} = \mathbf{1}_{\{a \le b\}}. \tag{2.1}$$

Here $b \in X_C$ and $a \in X$. Each column has exactly one nonzero entry.

Definition 2.1. A normalized decoherence functional on $\mathfrak{G}$ is a Hermitian map $D : \mathfrak{G} \times \mathfrak{G} \to \mathbb{C}$, additive on disjoint events in each argument, with $D(1,1) = 1$ and $D(A,A) \ge 0$ for every event A. It is *strongly positive* if $[D(A_i, A_j)]_{i,j}$ is positive semidefinite for every finite list of events. A context C is *medium decoherent* if $D(b,b') = 0$ for all distinct $b, b' \in X_C$.

Additivity determines D from its atomic matrix H by

$$H_{ab} = D(a,b), \qquad D(A,B) = \mathbf{1}_A^{\mathsf{T}} H \mathbf{1}_B. \tag{2.2}$$

Strong positivity is equivalent to $H \succeq 0$. The list of all atoms gives necessity. For sufficiency, take V with rows $\mathbf{1}_{A_i}^{\mathsf{T}}$ for any list of events. Their matrix is VHV^{T}. The indicator vectors are real, and normalization reads $\mathbf{1}^{\mathsf{T}} H \mathbf{1} = 1$.

For $A, B \in C$, expand $D(A,B)$ over the atoms of C. Medium decoherence leaves only the diagonal terms, giving

$$D(A,B) = \sum_{\substack{b \in X_C \\ b \le A \cap B}} D(b,b) = p_C(A \cap B), \qquad A, B \in C. \tag{2.3}$$

Thus $p_C(A) = D(A,A)$ defines a probability measure on C, as in the history sum rule [1, 3]. Restriction preserves medium decoherence, and the measures induced by a common D agree on intersections. The weaker condition $\mathrm{Re}\, D(b,b') = 0$ for $b \ne b'$ suffices for additivity of $D(A,A)$. We require medium decoherence to obtain (2.3) for every pair of events in C.

2.2 Classical extensions and strongly positive realizations

For each $C \in \mathcal{M}$, let e_C be a probability vector on X_C. Assume that the induced measures agree on pairwise intersections. Stacking the matrices P_C and vectors e_C in the same order gives M and e. Write

$$\mathcal{Q}(e) = \{q \in \mathbb{R}^X : Mq = e\}, \tag{2.4}$$

$$\mathcal{H}(e) = \{H = H^* \succeq 0 : P_C H P_C^{\mathsf{T}} = \operatorname{diag}(e_C) \text{ for every } C \in \mathcal{M}\}. \tag{2.5}$$

Both sets incorporate normalization. Summing $P_C q = e_C$ over a single context gives $\mathbf{1}^{\mathsf{T}} q = 1$. Summing both indices of its matrix constraint gives $\mathbf{1}^{\mathsf{T}} H \mathbf{1} = 1$.

Definition 2.2. A *classical extension* of e to $\mathfrak{G}$ is an element of $\mathcal{Q}(e)$ with nonnegative entries. A *strongly positive realization* of e is an element of $\mathcal{H}(e)$. The classical probability polytope relative to $\mathfrak{G}$ is

$$\mathsf{NC}_{\mathfrak{G}} = M\Delta(X), \qquad \Delta(X) = \{g \geq 0 : \mathbf{1}^{\mathsf{T}} g = 1\}.$$

A vector in $\Delta(X)$ defines a probability measure on $\mathfrak{G}$, and multiplication by P_C gives its restriction to C. A classical extension is therefore a unifying probability in Halliwell's sense [4]. Each classical extension g gives a realization $H = \operatorname{diag}(g)$. In the reverse direction, a realization always gives a signed extension.

Lemma 2.3 (Canonical signed extension)**.** *If $H \in \mathcal{H}(e)$, then $q_H = \operatorname{Re}(H\mathbf{1})$ belongs to $\mathcal{Q}(e)$.*

Proof. Since $P_C^{\mathsf{T}} \mathbf{1} = \mathbf{1}$,

$$P_C H \mathbf{1} = P_C H P_C^{\mathsf{T}} \mathbf{1} = e_C.$$

Taking real parts proves the assertion. □

For an event A, the canonical extension is $q_H(A) = \operatorname{Re} D(A, 1)$. For Heisenberg-picture class operators with $C_X = I$ and initial state ρ, this is the linear assignment $\operatorname{Re}\operatorname{tr}(C_A \rho)$ of Goldstein and Page [21]. Even if q_H takes negative values, another signed extension of the same data may be nonnegative [22].

2.3 Ambient atoms and compatible local valuations

Let $\mathcal{C}$ consist of all intersections of one or more members of $\mathcal{M}$, together with the two-element algebra, ordered by inclusion. For $C' \subseteq C$, restriction sends each atom of C to the atom of C' containing it. These maps define an inverse system whose limit

$$L = \varprojlim_{C \in \mathcal{C}} X_C$$

consists of the compatible local atom choices. Each choice is determined by its components on $\mathcal{M}$. An ambient atom a gives the choice $j(a) = (r_C(a))_{C \in \mathcal{M}}$.

Proposition 2.4 (Comparison of assignment spaces)**.** *The map j is injective. Its image consists of the compatible choices $(b_C)_{C \in \mathcal{M}}$ for which $\bigcap_C b_C \neq 0$. A probability on X corresponds to a probability on L supported on $j(X)$. In particular,*

$$\mathsf{NC}_{\mathfrak{G}} \subseteq \mathsf{NC}_L,$$

where NC_L is the convex hull of the deterministic probability families indexed by L. The two polytopes coincide if and only if j is surjective.

Proof. The nonempty intersections obtained by choosing one atom from each generating context partition X. Every generator is a union of blocks of this partition. Since the contexts generate $\mathfrak{G}$, the blocks are its atoms. This proves injectivity and the description of $j(X)$. Pushforward by j identifies probabilities on X with probabilities supported on $j(X)$, giving the inclusion and equality when j is surjective. If $\ell \in L \setminus j(X)$, its deterministic marginal assigns probability one to ℓ_C in every context. A measure realizing all these marginals would assign probability one to their finite intersection, which is empty. □

Surjectivity means that every compatible family of local ultrafilters extends to an ultrafilter on $\mathfrak{G}$, the amalgamation condition for commuting Boolean subalgebras [6, 7]. With more than two contexts the condition applies jointly to the whole family. Under finite Stone duality, j corresponds to the surjection from the Boolean colimit of the overlap diagram onto $\mathfrak{G}$. Its kernel ideal consists of the subsets of L disjoint from $j(X)$.

For a finite set Y of measurements with finite nonempty outcome sets O_y, let $X = \prod_{y \in Y} O_y$ and take the coordinate algebras of a cover of Y as contexts. Agreement on overlaps gives a unique outcome for each measurement, so j is bijective. Classical extension then reduces to the global-distribution criterion of Abramsky and Brandenburger [5].

Adjoining $\mathfrak{G}$ to $\mathcal{C}$ changes the global sections of the atomic presheaf from L to X, since the atom chosen at the top determines all restrictions. Likewise, (e_C) and its restrictions give a global section of the distribution presheaf on $\mathcal{C}$. Extending this section to the enlarged poset amounts to choosing a classical extension on X.

2.4 Support and finite duality

The assignments compatible with the local supports form

$$S_e = \{a \in X : e_C(r_C(a)) > 0 \text{ for every } C \in \mathcal{M}\}. \tag{2.6}$$

Every classical extension is supported on S_e. This set may be empty even though X is nonempty. Conversely, $S_e = X$ need not imply classical extendability, as the three-time model in Section 8.5 shows. The distinction between support and probability occurs in the contextuality hierarchy [5]. In histories, contrary inferences can arise when different consistent sets assign probability one to incompatible alternatives [23, 24].

Classical extendability amounts to feasibility of $Mg = e$, $g \geq 0$. By Farkas' lemma, this system is infeasible if and only if there is a real vector y, indexed by the context atoms, satisfying

$$M^{\mathsf{T}} y \geq 0, \qquad y^{\mathsf{T}} e < 0. \tag{2.7}$$

The inequality $y^{\mathsf{T}} M g \geq 0$ then holds for every $g \in \Delta(X)$, but the specified data violate it. If the data admit a common strongly positive realization, Lemma 2.3 already supplies a solution of the affine equations $Mg = e$.

3 Covariance parametrization of strongly positive realizations

For each C, the context probabilities e_C determine the covariance matrix $\operatorname{diag}(e_C) - e_C e_C^{\mathsf{T}}$ of its atom indicators. Define

$$\mathcal{K}(e) = \left\{ K = K^* \succeq 0 : K\mathbf{1} = 0, \quad P_C K P_C^{\mathsf{T}} = \operatorname{diag}(e_C) - e_C e_C^{\mathsf{T}} \text{ for all } C \in \mathcal{M} \right\}. \tag{3.1}$$

Write $\mathcal{Q}_{\mathbb{C}}(e) = \{z \in \mathbb{C}^X : Mz = e\}$ for the complex affine solution space.

Theorem 3.1 (Covariance parametrization and preservation of rank)**.** *The map*

$$(z, K) \longmapsto zz^* + K \tag{3.2}$$

is a bijection from $\mathcal{Q}_{\mathbb{C}}(e) \times \mathcal{K}(e)$ *onto* $\mathcal{H}(e)$*. Its inverse is*

$$z = H\mathbf{1}, \qquad K = H - (H\mathbf{1})(H\mathbf{1})^*. \tag{3.3}$$

For every such pair,

$$\operatorname{rank}(zz^* + K) = 1 + \operatorname{rank} K. \tag{3.4}$$

For each $H \in \mathcal{H}(e)$ *and each real signed extension* $q \in \mathcal{Q}(e)$*, there is an* $H_q \in \mathcal{H}(e)$ *with* $H_q\mathbf{1} = q$ *and* $\operatorname{rank} H_q = \operatorname{rank} H$*.*

Proof. Let $H \in \mathcal{H}(e)$ and put $z = H\mathbf{1}$. The calculation in Lemma 2.3 gives $P_C z = e_C$, and normalization gives $\mathbf{1}^* z = 1$. For $u \in \mathbb{C}^X$, Cauchy–Schwarz for the form induced by H gives

$$|u^* H\mathbf{1}|^2 \le (u^* H u)(\mathbf{1}^* H\mathbf{1}) = u^* H u.$$

It follows that $K = H - zz^* \succeq 0$. Since $z^*\mathbf{1} = 1$, we have $K\mathbf{1} = 0$. Applying P_C on the left and P_C^{T} on the right gives the prescribed context covariance.

Conversely, let $Mz = e$ and $K \in \mathcal{K}(e)$. Summing the equations for one context gives $\mathbf{1}^* z = 1$. Thus $H = zz^* + K$ is positive semidefinite, satisfies $H\mathbf{1} = z$, and obeys

$$P_C H P_C^{\mathsf{T}} = e_C e_C^{\mathsf{T}} + \operatorname{diag}(e_C) - e_C e_C^{\mathsf{T}} = \operatorname{diag}(e_C).$$

This proves the bijection. Since $K\mathbf{1} = 0$, its range lies in $\mathbf{1}^{\perp}$. The identity $\mathbf{1}^* z = 1$ puts z outside this range. The range of a sum of positive semidefinite matrices is the sum of their ranges, giving (3.4). Replacing z by q preserves K and hence the rank. □

If H is real, so is K, and the replacement $qq^{\mathsf{T}} + K$ yields a real realization of the same rank.

Corollary 3.2. *A strongly positive realization exists if and only if both* $\mathcal{Q}(e)$ *and* $\mathcal{K}(e)$ *are nonempty. In that case,*

$$\min_{H \in \mathcal{H}(e)} \operatorname{rank} H = 1 + \min_{K \in \mathcal{K}(e)} \operatorname{rank} K.$$

Prescribing any $q \in \mathcal{Q}(e)$ *as* $H\mathbf{1}$ *gives the same minimum rank.*

Proof. A realization gives $K \in \mathcal{K}(e)$ and $\operatorname{Re}(H\mathbf{1}) \in \mathcal{Q}(e)$. Conversely, any $q \in \mathcal{Q}(e)$ and $K \in \mathcal{K}(e)$ give a realization of rank $1 + \operatorname{rank} K$ by Theorem 3.1. □

3.1 The exact trace cost of a prescribed extension

For realizable data, define

$$\tau(e) = \min_{H \in \mathcal{H}(e)} \operatorname{tr} H, \qquad \kappa(e) = \min_{K \in \mathcal{K}(e)} \operatorname{tr} K. \tag{3.5}$$

The feasible sets are closed, and a trace bound bounds every eigenvalue of a positive semidefinite matrix. Thus both minima are attained. The unique signed extension of least Euclidean norm is $q_0 = M^+ e$, where M^+ is the Moore–Penrose inverse.

Theorem 3.3 (Trace and signed extension)**.** *If e admits a strongly positive realization, then*

$$\tau(e) = \|q_0\|_2^2 + \kappa(e), \tag{3.6}$$

$$\min_{\substack{H\in\mathcal{H}(e)\\ H\mathbf{1}=q}} \operatorname{tr} H = \tau(e) + \|q - q_0\|_2^2 \qquad (q \in \mathcal{Q}(e)). \tag{3.7}$$

A trace-minimizing realization can be chosen real. If

$$\mathcal{Q}_s(e) = \{q \in \mathcal{Q}(e) : \textstyle\sum_a (-q_a)_+ \le s\}$$

is nonempty, where $(t)_+ = \max\{t, 0\}$, then

$$\min_{\substack{H\in\mathcal{H}(e)\\ \sum_a(-\operatorname{Re}(H\mathbf{1})_a)_+\le s}} \operatorname{tr} H = \tau(e) + \operatorname{dist}_2(q_0, \mathcal{Q}_s(e))^2. \tag{3.8}$$

Proof. By Theorem 3.1, $\operatorname{tr} H = \|z\|_2^2 + \operatorname{tr} K$, and z and K can be chosen independently. Writing $z = q + it$ with q, t real gives $Mq = e$, $Mt = 0$, and $\|z\|_2^2 = \|q\|_2^2 + \|t\|_2^2$. Thus a trace minimum requires $t = 0$. Replacing K by its real part preserves positive semidefiniteness, the context covariances, the zero row sums, and the trace. This gives a real minimizer. Since $q - q_0 \in \ker M$ and $q_0 \perp \ker M$,

$$\|q\|_2^2 = \|q_0\|_2^2 + \|q - q_0\|_2^2.$$

Minimizing in K gives (3.7) and, with $q = q_0$, (3.6). The negative-mass constraint depends only on q. The remaining minimum is the squared distance to the nonempty closed set $\mathcal{Q}_s(e)$, giving (3.8). □

The trace $\operatorname{tr} H = \sum_a D(a, a)$ is the total diagonal mass on the chosen atomic refinement. Expanding $\mathbf{1}^\mathsf{T} H \mathbf{1} = 1$ gives

$$1 - \operatorname{tr} H = 2 \sum_{a<b} \operatorname{Re} H_{ab}. \tag{3.9}$$

Thus the trace also determines the net real interference on that refinement.

3.2 Compression to observable event functions

The context-atom indicators span $W = \operatorname{ran} M^\mathsf{T} \subseteq \mathbb{R}^X$. Let Π denote the orthogonal projection onto W for the counting-measure inner product, and extend it complex linearly to $\mathbb{C}^X$.

Proposition 3.4 (Compression)**.** *For $H \in \mathcal{H}(e)$, the matrix $\Pi H \Pi$ is another realization and*

$$\operatorname{tr}(\Pi H \Pi) \le \operatorname{tr} H.$$

Equality holds if and only if $\operatorname{ran} H \subseteq W_\mathbb{C}$*, the complexification of W. The compressed matrix satisfies $(\Pi H \Pi)\mathbf{1} = q_0 = M^+ e$.*

Proof. Each row of P_C belongs to W, so $P_C \Pi = P_C$. Compression preserves positive semidefiniteness and all context equations. Moreover,

$$\operatorname{tr} H - \operatorname{tr}(\Pi H \Pi) = \operatorname{tr}((I - \Pi)H) = \|(I - \Pi)H^{1/2}\|_F^2,$$

where $\|\cdot\|_F$ denotes the Frobenius norm. The loss vanishes precisely when $\operatorname{ran} H^{1/2} = \operatorname{ran} H \subseteq W_\mathbb{C}$. Since $\mathbf{1} \in W$, the row-sum vector of the compression is $\Pi H \mathbf{1}$. It is the projection of a solution of $Mz = e$ onto $W_\mathbb{C}$, the orthogonal complement of the complex null space of M, hence equals the least-norm solution q_0. □

4 Two contexts and their atom-intersection graph

Let $(A_i)_{i\in R}$ and $(B_j)_{j\in S}$ be the atoms of two contexts. The atoms of the Boolean algebra they generate are the nonempty intersections

$$a_{ij} = A_i \cap B_j \neq 0.$$

Define $G = (R \sqcup S, E)$ by taking (i, j) as an edge whenever $A_i \cap B_j \neq 0$. This simple bipartite graph has no isolated vertices. We identify X with E and write $N = |E|$.

Let P_R, P_S be the two atom-summing matrices, and define

$$T = \begin{pmatrix} P_R \\ P_S \end{pmatrix}, \qquad B = \begin{pmatrix} P_R \\ -P_S \end{pmatrix}, \qquad L = BB^{\mathsf{T}}. \tag{4.1}$$

The signed incidence matrix B uses $+1$ at the row endpoint and -1 at the column endpoint. Thus L is the unweighted graph Laplacian. Write $p \in \Delta(R)$ and $q \in \Delta(S)$ for the context probabilities.

Lemma 4.1 (Intersection algebra and signed extensions)**.** *The atoms of the intersection of the two contexts are the edge sets of the connected components of G. If these components are indexed by ℓ, agreement on the intersection algebra is equivalent to*

$$\sum_{i\in R_\ell} p_i = \sum_{j\in S_\ell} q_j = w_\ell \qquad \text{for every } \ell. \tag{4.2}$$

Under this condition, the set

$$\mathcal{Q}_G(p, q) = \{x \in \mathbb{R}^E : P_R x = p,\ P_S x = q\} \tag{4.3}$$

is a nonempty affine space of dimension $|E| - |R| - |S| + c$, where c is the number of connected components.

Proof. An event in the row context is a union of entire row stars, and similarly for columns. If it belongs to both contexts, the two endpoint indicators of each edge agree. They are therefore constant on each component, and any union of components belongs to both contexts. This proves the description of the common algebra and (4.2). The marginal equations are $Bx = (p, -q)$. On each connected component the image of the incidence matrix is the subspace of vectors with coordinate sum zero, as one sees by successively solving at the leaves of a spanning tree. Component balance is precisely this condition on $(p, -q)$. Since $\operatorname{rank} B = |R| + |S| - c$, the dimension follows. □

4.1 Existence and the smallest realization rank

Theorem 4.2 (Two-context realizations)**.** *For $p \in \Delta(R)$ and $q \in \Delta(S)$, the following conditions are equivalent.*

(i) The component balances (4.2) hold.

(ii) The signed extension space (4.3) is nonempty.

(iii) The two contexts admit a common strongly positive realization with medium decoherence and probabilities p, q.

Let $r_+ = |\operatorname{supp} p|$ and $s_+ = |\operatorname{supp} q|$. Under these conditions the minimum rank is the same over real symmetric and complex Hermitian realizations, namely

$$\min_H \operatorname{rank} H = \max(r_+, s_+). \tag{4.4}$$

For every specified $x \in \mathcal{Q}_G(p, q)$, this rank is attained by a real realization satisfying $H\mathbf{1} = x$.

Proof. By Lemma 4.1, (i) and (ii) are equivalent. A common realization agrees on shared events, so (iii) implies (i). Assume (i) and put $d = \max(r_+, s_+)$. In $\mathbb{R}^d$, choose mutually orthogonal row vectors u_i with $\|u_i\|^2 = p_i$. Their component sums

$$h_\ell = \sum_{i \in R_\ell} u_i$$

are mutually orthogonal with squared norms w_ℓ. Choose a second orthogonal family $\widetilde{v}_j \in \mathbb{R}^d$ with squared norms q_j. Its component sums have the same Gram matrix as the h_ℓ. The isometry sending these sums to h_ℓ extends to an orthogonal operator on $\mathbb{R}^d$. Applying this operator to the $\widetilde{v}_j$ gives mutually orthogonal column vectors v_j satisfying

$$\sum_{j \in S_\ell} v_j = h_\ell = \sum_{i \in R_\ell} u_i. \tag{4.5}$$

Solve the vector-valued incidence equation

$$\sum_{e \ni i} b_e = u_i, \qquad \sum_{e \ni j} b_e = v_j. \tag{4.6}$$

By Lemma 4.1, equality of the component sums in (4.5) is the solvability condition in each coordinate. The Gram matrix $H_{ef} = \langle b_e, b_f \rangle$ then has the prescribed context matrices. Its total vector $h = \sum_\ell h_\ell$ satisfies $\|h\|^2 = \sum_\ell w_\ell = 1$. This gives a normalized real realization of rank at most d.

Conversely, a Gram representation of any realization contains r_+ nonzero mutually orthogonal row-sum vectors and s_+ nonzero mutually orthogonal column-sum vectors in the span of the edge vectors. This forces rank at least d over either field. Finally, Theorem 3.1 allows us to prescribe $H\mathbf{1} = x$ while retaining the rank of the real realization just constructed. □

For disconnected G, the construction places both orthogonal families in the same $\mathbb{R}^d$. Row vectors and column vectors from different components may share directions orthogonal to the component sums. Taking independent spaces for each component could therefore overestimate the required dimension.

4.2 The largest rank and positive definite realizations

Let $r_0 = |R| - r_+$ and $s_0 = |S| - s_+$, and let c_0 count the components of zero mass.

Proposition 4.3 (Maximum realization rank)**.** *Under* (4.2)*,*

$$\max_H \operatorname{rank} H = N - r_0 - s_0 + c_0. \tag{4.7}$$

This maximum can be attained by a real realization with any prescribed real signed extension.

Proof. If a context atom has probability zero, its indicator has zero quadratic form under $H \succeq 0$ and hence lies in $\ker H$. A linear relation among the row and column indicators has coefficients α_i, β_j satisfying $\alpha_i + \beta_j = 0$ on every edge. Thus on each component all row coefficients equal a common constant and all column coefficients equal its negative. A relation involving only zero-probability indicators can use a component only if its mass is zero. Their span therefore has dimension $r_0 + s_0 - c_0$, which proves the upper bound.

For attainment, define the $|R| \times |S|$ matrix

$$F_{ij} = \begin{cases} p_i q_j / w_\ell, & i, j \text{ in component } \ell,\ w_\ell > 0, \\ 0, & \text{otherwise.} \end{cases}$$

This definition includes nonadjacent vertex pairs. Consider the block matrix

$$J = \begin{pmatrix} \operatorname{diag} p & F \\ F^{\mathsf{T}} & \operatorname{diag} q \end{pmatrix}.$$

On each component of mass $w > 0$, write p', q' for the restricted marginals. The corresponding block of J is the sum of $\binom{p'}{q'}\binom{p'}{q'}^{\mathsf{T}}/w$ and the block diagonal matrix with blocks $\operatorname{diag}(p') - p'p'^{\mathsf{T}}/w$ and $\operatorname{diag}(q') - q'q'^{\mathsf{T}}/w$. Both covariance blocks are positive semidefinite, with ranks $|\operatorname{supp} p'| - 1$ and $|\operatorname{supp} q'| - 1$. Their ranges have zero coordinate sum on each side, so the first summand adds one dimension. Zero-mass components contribute nothing. Hence $J \succeq 0$ and $\operatorname{rank} J = r_+ + s_+ - c_+$, where $c_+ = c - c_0$. The row-minus-column relation on each component lies in $\ker J$, so $\operatorname{ran} J \subseteq \operatorname{ran} T$. Let Π_0 project orthogonally onto $\ker T$. For $\varepsilon > 0$ the matrix

$$H = T^+ J (T^+)^{\mathsf{T}} + \varepsilon \Pi_0$$

satisfies $THT^{\mathsf{T}} = J$, since TT^+ projects onto $\operatorname{ran} T$. It therefore realizes both contexts. The first term has rank $\operatorname{rank} J$ on $(\ker T)^{\perp}$, and the second has range $\ker T$. Their ranks add to

$$r_+ + s_+ - c_+ + (N - |R| - |S| + c) = N - r_0 - s_0 + c_0.$$

Keeping the covariance of this real H and applying Theorem 3.1 gives the same rank for every prescribed real signed extension. □

Corollary 4.4. *Under the component balance conditions (4.2), a positive definite realization exists if and only if all row and column probabilities are strictly positive.*

Proof. A zero-probability context atom supplies a nonzero kernel vector. If all probabilities are positive, Proposition 4.3 gives rank N. □

Thus strict positivity of both marginals guarantees a positive definite realization whenever they agree on overlaps. A classical extension has the additional constraints below.

4.3 The additional condition for a classical measure

For $U \subseteq R$, let $N_G(U) \subseteq S$ be its neighbor set.

Proposition 4.5 (Weighted transport criterion)**.** *A classical extension exists if and only if*

$$p(U) \le q(N_G(U)) \qquad \text{for every } U \subseteq R. \tag{4.8}$$

Proof. Any nonnegative edge measure must send the mass of U into $N_G(U)$, which proves necessity. For sufficiency, join a source to row i with capacity p_i, and column j to a sink with capacity q_j. Give each edge of G capacity 2 from row to column. A cut crossing such an edge has capacity at least 2. Any other cut with source-side row set U must also place $N_G(U)$ on the source side, and thus has capacity at least $1 - p(U) + q(N_G(U)) \ge 1$. The max-flow min-cut theorem gives a unit flow. It saturates all source and sink arcs, so its edge flows have the required marginals. □

Criterion (4.8) is the finite supported-coupling condition in classical marginal theory [25]. For the same graph, Theorem 4.2 requires only the component balances for a strongly positive realization.

Corollary 4.6 (Universal classical extension)**.** *For two fixed contexts, every overlap-compatible pair of probabilities extends classically if and only if every component of G is complete bipartite. This is also equivalent to surjectivity of the map j in Proposition 2.4.*

Proof. A row atom and a column atom are compatible on the intersection algebra exactly when their vertices lie in the same component. They come from an ambient atom exactly when they form an edge, giving the criterion for surjectivity. On a complete bipartite component of mass $w_\ell > 0$, set $x_{ij} = p_i q_j / w_\ell$, and assign zero mass to every edge of a zero-mass component. This gives a classical extension. A missing edge in a component yields compatible point masses at its two endpoints with no classical extension. □

For two finite Boolean subalgebras this surjectivity condition is precisely commutation. In a full product measurement cover, compatible coordinate assignments always glue, so any two overlap-compatible context marginals extend classically.

5 Signed extensions, transport, and interference

5.1 Minimum negative mass and classical noise

For a real vector x of total mass 1, define

$$N_-(x) = \sum_a (-x_a)_+ = \frac{\|x\|_1 - 1}{2}.$$

Whenever $\mathcal{Q}(e) \neq \varnothing$, put

$$\eta(e) = \min_{x \in \mathcal{Q}(e)} N_-(x). \tag{5.1}$$

The sublevel sets of N_- in the closed affine space $\mathcal{Q}(e)$ are compact, since $\|x\|_1 = 1 + 2N_-(x)$. Thus the minimum exists, and $\eta(e) = 0$ exactly when $e \in \mathsf{NC}_{\mathfrak{G}}$. Minimum ℓ^1 norm has also been used to quantify contextuality [13]. Here the choice of history algebra fixes the atoms on which the signed measure is defined.

Proposition 5.1 (Duality, realization, and noise)**.** *Suppose $\mathcal{Q}(e) \neq \varnothing$. Then*

$$1 + 2\eta(e) = \max_{\|M^{\mathsf{T}} y\|_\infty \leq 1} e^{\mathsf{T}} y, \tag{5.2}$$

$$\eta(e) = \min\{t \geq 0 : e + t n_- = (1+t) n_+, \quad n_-, n_+ \in \mathsf{NC}_{\mathfrak{G}}\}. \tag{5.3}$$

If e admits a strongly positive realization, then

$$\eta(e) = \min_{H \in \mathcal{H}(e)} N_-(\mathrm{Re}(H\mathbf{1})), \tag{5.4}$$

and each rank occurring in $\mathcal{H}(e)$ occurs among the minimizers in (5.4).

Proof. Linear programming duality for $\min_{Mx=e} \|x\|_1$ gives (5.2). To prove (5.3), let $x = x^+ - x^-$ be a minimizing extension and write $\eta = \eta(e)$. Its positive and negative parts have masses $1 + \eta$ and η. If $\eta > 0$, normalizing these parts gives distributions $g_+ = x^+/(1+\eta)$ and $g_- = x^-/\eta$. Then $n_\pm = M g_\pm$ satisfy $e + \eta n_- = (1+\eta) n_+$. If $\eta = 0$, take $n_+ = e$ and any $n_- \in \mathsf{NC}_{\mathfrak{G}}$. Conversely, any feasible $n_\pm$ lift to distributions $g_\pm$ on X. The vector $(1+t)g_+ - t g_-$ extends e and has negative mass at most t.

For (5.4), Lemma 2.3 gives the lower bound. Given any $H \in \mathcal{H}(e)$, Theorem 3.1 produces a realization of the same rank whose canonical extension is x. □

Measure the distance of e from classical data by the largest total variation distance over the contexts, minimized over global distributions,

$$\delta(e) = \min_{g \in \Delta(X)} \max_{C \in \mathcal{M}} \frac{1}{2} \|e_C - P_C g\|_1. \tag{5.5}$$

Taking $Mg = n_+$ in this definition yields

$$\delta(e) \le \frac{\eta(e)}{1+\eta(e)}. \tag{5.6}$$

Indeed, $e - n_+ = \eta(e - n_-)/(1+\eta)$, and the total variation distance between e_C and $(n_-)_C$ is at most 1 for each C. Equation (5.3) identifies $\eta/(1+\eta)$ as the least λ for which $(1-\lambda)e + \lambda n_-$ is classical for some classical n_-. Both optimizations use the polytope $\mathsf{NC}_{\mathfrak{G}}$. The contextual fraction uses a different decomposition [26].

5.2 An exact transport formula for two contexts

On the bipartite graph of Section 4, let d_G be the unweighted shortest-path distance within each component. Regard p and q as measures supported on R and S, respectively. Their Wasserstein distance is

$$W_1^G(p,q) = \min_{\substack{\pi \ge 0 \\ \pi \mathbf{1} = p,\ \pi^{\mathsf{T}} \mathbf{1} = q}} \sum_{\substack{i \in R,\, j \in S \\ d_G(i,j) < \infty}} d_G(i,j)\pi_{ij}, \tag{5.7}$$

where $d_G(i,j) = \infty$ across components and the minimum is over plans with $\pi_{ij} = 0$ for all such pairs. Component balance guarantees a feasible plan. Within a component, the plan may transport mass between nonadjacent vertices. Write $\eta_G(p,q)$ for the minimum in (5.1) on this graph.

Theorem 5.2 (Graph transport formula)**.** *For every component-balanced pair p, q,*

$$\eta_G(p,q) = \frac{W_1^G(p,q) - 1}{2}. \tag{5.8}$$

Equivalently,

$$1 + 2\eta_G(p,q) = \max_{\substack{f : R \sqcup S \to \mathbb{R} \\ |f(i) - f(j)| \le 1 \ \ ((i,j) \in E)}} \left(\sum_i p_i f(i) - \sum_j q_j f(j) \right). \tag{5.9}$$

Every signed edge measure attaining $\eta_G(p,q)$ is the canonical extension of a strongly positive realization of rank $\max(r_+, s_+)$.

Proof. Since the extension equations are $Bx = (p, -q)$, it suffices for (5.8) to prove

$$\min_{Bx = (p,-q)} \|x\|_1 = W_1^G(p,q).$$

For any feasible plan π, send the mass π_{ij} along a shortest path from i to j whenever $\pi_{ij} > 0$. Count flow along each edge with its sign relative to the orientation of B. The resulting flow has divergence $(p, -q)$, and its ℓ^1 norm is at most the transport cost by the triangle inequality.

Conversely, direct each edge with nonzero flow according to the sign of x_e and give it weight $|x_e|$. Removing directed cycles preserves the divergence and can only reduce the total weight. Decompose the remaining flow into paths from row vertices with positive divergence to column vertices with negative divergence. Their weights define a plan π. Each path has length at least the distance between its endpoints, so the cost of π is at most $\|x\|_1$. This proves the identity. As every extension has total mass 1, (5.8) follows. Duality for the ℓ^1 flow problem gives (5.9), and Theorem 4.2 gives the rank assertion. □

The flow identity is the graph Beckmann formulation of W_1 [27, Sec. 1.4]. Here it gives the minimum negative mass on the history atoms. A single Hall deficit can be smaller than this minimum. For example, on a path of five edges, Dirac marginals at opposite endpoints determine the edge weights

$$(1, -1, 1, -1, 1).$$

Their negative mass is 2, although every individual Hall deficit is at most 1.

Corollary 5.3 (Deterministic marginals and sharp size dependence)**.** *For a row vertex s and column vertex t in the same component,*

$$\eta_G(\delta_s, \delta_t) = \frac{d_G(s,t) - 1}{2}. \tag{5.10}$$

Every signed extension x of these Dirac marginals gives a rank-one realization $H = xx^{\mathsf{T}}$. For any component-balanced pair p, q on a graph with N edges,

$$\eta_G(p, q) \le \left\lfloor \frac{N-1}{2} \right\rfloor. \tag{5.11}$$

For each $N \ge 1$ there are a graph and marginals attaining the bound. The smallest two-context model with compatible data and no classical extension has three atoms.

Proof. The only plan for the Dirac marginals transports one unit from s to t, giving (5.10). For $H = xx^{\mathsf{T}}$, the compressed matrix on either context is the outer product of its Dirac marginal with itself, hence equals the required diagonal matrix.

Every shortest path from a row to a column in the same component has odd length at most N. A feasible plan has mass 1, so its cost is at most the largest such length. Equation (5.8) gives (5.11). For odd N, take the two endpoints of a path of length N. For even N, take the endpoints of a path of length $N-1$ and attach a pendant edge. Their distance remains $N-1$. With at most two edges, every component is complete bipartite and hence admits a classical coupling for any balanced marginals. Dirac data at the endpoints of the three-edge path have no such coupling. □

5.3 Sharp bounds on real interference and diagonal mass

For a Hermitian realization H, define the total negative real interference on the atoms by

$$I_-(H) = \sum_{a<b} (-\operatorname{Re} H_{ab})_+, \tag{5.12}$$

where the sum counts each unordered pair once.

Theorem 5.4 (Interference and trace bounds)**.** *For every normalized strongly positive realization on N atoms,*

$$I_-(H) \ge N_-(q_H)(1 + N_-(q_H)) \ge \eta(e)(1 + \eta(e)), \tag{5.13}$$

$$\operatorname{tr} H \ge \frac{(1 + 2N_-(q_H))^2}{N} \ge \frac{(1 + 2\eta(e))^2}{N}, \qquad q_H = \operatorname{Re}(H\mathbf{1}). \tag{5.14}$$

If every local probability vector is deterministic and a signed extension exists, then

$$\min_{H \in \mathcal{H}(e)} I_-(H) = \eta(e)(1 + \eta(e)). \tag{5.15}$$

Proof. For events S, T, write $H(S,T) = \mathbf{1}_S^{\mathsf{T}} H \mathbf{1}_T$ and denote the full event by 1. Set $S = \{a : (q_H)_a < 0\}$ and $n_- = N_-(q_H)$. Since $H(1,1) = 1$, Cauchy–Schwarz for the form defined by H gives

$$H(S,S) \geq |H(S,1)|^2 \geq (\operatorname{Re} H(S,1))^2 = n_-^2.$$

Consequently,

$$\operatorname{Re} H(S,S^c) = -n_- - H(S,S) \leq -n_-(1+n_-).$$

The sum of the negative parts of the real entries crossing the cut is at least $-\operatorname{Re} H(S,S^c)$ and at most $I_-(H)$. This gives the first inequality of (5.13). For (5.14), put $z = H\mathbf{1}$. The covariance decomposition gives $\|z\|_2^2 \leq \operatorname{tr} H$, so

$$1 + 2n_- = \|q_H\|_1 \leq \sqrt{N}\|q_H\|_2 \leq \sqrt{N \operatorname{tr} H}.$$

Finally, $q_H \in \mathcal{Q}(e)$ implies $n_- \geq \eta(e)$.

For deterministic local data, let x be an optimal signed extension. The rank-one matrix $H = xx^{\mathsf{T}}$ realizes the data. Its negative off-diagonal entries occur exactly between positive and negative coordinates, giving

$$I_-(xx^{\mathsf{T}}) = \left(\sum_a x_a^+\right)\left(\sum_a x_a^-\right) = (1+\eta)\eta.$$

This attains the lower bound in (5.13). □

For two Dirac marginals at odd distance d, (5.10) and (5.15) give minimum negative real interference $(d^2-1)/4$. The parallel paths in Proposition 8.2 attain the dimension-dependent trace bound. The interference in these inequalities occurs on the common atomic refinement. Distinct atoms within either prescribed context have zero interference.

6 Minimum trace and graph geometry

6.1 An exact Laplacian formula

Let G be connected. To treat its components separately later, allow the marginals p and q to have a common mass $w > 0$. Their local Gram equations then give $\mathbf{1}^{\mathsf{T}} H \mathbf{1} = w$. Let L^+ denote the Moore–Penrose inverse of the graph Laplacian and set

$$A = (L^+)_{R,S}, \qquad R_p = \operatorname{diag} p - \frac{pp^{\mathsf{T}}}{w}, \qquad S_q = \operatorname{diag} q - \frac{qq^{\mathsf{T}}}{w}. \tag{6.1}$$

Write $\|\cdot\|_*$ for the nuclear norm, the sum of the singular values.

Theorem 6.1 (Minimum-trace formula)**.** *For a connected atom-intersection graph and nonnegative marginals of equal mass $w > 0$, the minimum trace among all strongly positive realizations is*

$$\tau_G(p,q) = \sum_{i\in R}(L^+)_{ii}p_i + \sum_{j\in S}(L^+)_{jj}q_j - \frac{2}{w}p^{\mathsf{T}}Aq - 2\|R_p^{1/2}AS_q^{1/2}\|_*. \tag{6.2}$$

A real positive semidefinite matrix attains the minimum. On a disconnected graph with balanced component masses, $\tau_G(p,q)$ is the sum of (6.2) over the positive-mass components. Every zero-mass component contributes zero.

Proof. Choose edge vectors b_e with Gram matrix H. Their row sums u_i are mutually orthogonal and satisfy $\|u_i\|^2 = p_i$. Their column sums v_j are also mutually orthogonal and satisfy $\|v_j\|^2 = q_j$. Both families sum to the same vector h, with $\|h\|^2 = w$.

For fixed vertex vectors, form Z with rows u_i and $-v_j$, and form Y with rows b_e. Then $BY = Z$. Its solution of least Frobenius norm is $Y = B^+Z$, with energy

$$\min_{BY=Z} \sum_e \|b_e\|^2 = \operatorname{tr}(Z^* L^+ Z). \tag{6.3}$$

Indeed, $\sum_i u_i = \sum_j v_j$ puts each column of Z in $\operatorname{ran} B$, and $(B^+)^* B^+ = (BB^*)^+ = L^+$ gives (6.3).

Write

$$u_i = \frac{p_i}{w} h + \alpha_i, \qquad v_j = \frac{q_j}{w} h + \beta_j.$$

The vectors α_i and β_j lie in $h^\perp$. Their two Gram matrices are R_p and S_q, respectively. It follows that the possible cross Gram matrices F of the original families (u_i) and (v_j) are exactly

$$F = \frac{pq^{\mathsf{T}}}{w} + R_p^{1/2} W S_q^{1/2}, \qquad \|W\|_{\mathrm{op}} \le 1. \tag{6.4}$$

To see this, consider the joint Gram matrix of the residuals. Positivity forces the off-diagonal block to map $\operatorname{ran} S_q$ into $\operatorname{ran} R_p$ and to vanish on $\ker S_q$. Conjugating by the inverse square roots on these ranges gives $\left(\begin{smallmatrix} I & W \\ W^* & I \end{smallmatrix}\right)$, which is positive semidefinite exactly when $\|W\|_{\mathrm{op}} \le 1$. Conversely, any such contraction yields residual vectors with the prescribed Gram matrices after multiplication by the square roots. Their sums vanish since $R_p \mathbf{1} = S_q \mathbf{1} = 0$. Adding the indicated multiples of h therefore recovers both families of vertex vectors.

Orthogonality within each context reduces (6.3) to

$$\sum_i (L^+)_{ii} p_i + \sum_j (L^+)_{jj} q_j - 2 \operatorname{Re} \operatorname{tr}(A^{\mathsf{T}} F). \tag{6.5}$$

The singular-value identity

$$\max_{\|W\|_{\mathrm{op}} \le 1} \operatorname{Re} \operatorname{tr}\big((R_p^{1/2} A S_q^{1/2})^* W\big) = \|R_p^{1/2} A S_q^{1/2}\|_*$$

now gives (6.2). A real singular-value decomposition gives a real maximizing contraction. The resulting real Gram matrix has a real factorization, so the edge vectors from $Y = B^+ Z$ give a real minimizer.

For a disconnected graph, the principal block on the edges of each component satisfies its local Gram equations. Summing the corresponding trace bounds gives a lower bound on the full trace. Taking the direct sum of minimizing blocks attains it. On a zero-mass component the zero block satisfies all local constraints and has the least possible trace. □

For fixed vertex vectors, (6.3) is the vector-valued form of the electrical minimum-energy principle [14]. Optimizing over their cross Gram matrix gives the singular-value maximization familiar from the orthogonal Procrustes problem [15, 16].

Corollary 6.2 (Effective resistance)**.** *For Dirac marginals at a row vertex s and a column vertex t in the same component,*

$$\tau_G(\delta_s, \delta_t) = \mathrm{R}_{\mathrm{eff}\,G}(s, t) = (\delta_s - \delta_t)^{\mathsf{T}} L^+ (\delta_s - \delta_t). \tag{6.6}$$

An attaining realization is the rank-one matrix xx^{T}, where $x = B^+(\delta_s - \delta_t)$ is the minimum-energy unit flow.

Proof. The covariance blocks R_p, S_q are zero for Dirac marginals, so (6.2) reduces to (6.6). The stated signed edge measure has squared norm equal to this value and gives a realization by Corollary 5.3. □

For the same Dirac marginals we thus have

$$\eta_G(\delta_s, \delta_t) = \frac{d_G(s,t) - 1}{2}, \qquad \tau_G(\delta_s, \delta_t) = \mathrm{R}_{\mathrm{eff}\,G}(s,t).$$

The first formula minimizes the ℓ^1 norm of a unit flow, and the second minimizes its squared ℓ^2 norm. On a tree the flow is unique and $\mathrm{R}_{\mathrm{eff}\,G}(s,t) = d_G(s,t)$. On graphs with cycles the minimizing flows can differ.

6.2 The geometry of minimum-trace realizations

Although the minimum trace is additive over components, a minimizing realization may have nonzero correlations between residual vectors from different components. To describe all minimizers, let c_+ be the number of positive-mass components and put

$$a = r_+ - c_+, \qquad b = s_+ - c_+.$$

Regard p_ℓ, q_ℓ as vectors extended by zero outside component ℓ, and define

$$\mathsf{R} = \operatorname{diag} p - \sum_{w_\ell > 0} \frac{p_\ell p_\ell^{\mathsf{T}}}{w_\ell}, \qquad \mathsf{S} = \operatorname{diag} q - \sum_{w_\ell > 0} \frac{q_\ell q_\ell^{\mathsf{T}}}{w_\ell}. \tag{6.7}$$

Choose real factors of full column rank, $\mathsf{R} = F_R F_R^{\mathsf{T}}$ and $\mathsf{S} = F_S F_S^{\mathsf{T}}$. They have a and b columns, respectively, and either number may be zero. For the Laplacian of the whole graph, set

$$\mathsf{C} = F_R^{\mathsf{T}} (L^+)_{R,S} F_S, \qquad k = \operatorname{rank} \mathsf{C}. \tag{6.8}$$

Changing the full factors only changes the orthonormal coordinates in which C is expressed, so its rank k is well defined. Choosing factors separately on each component makes C block diagonal.

Theorem 6.3 (The minimum-trace face)**.** *Let p, q satisfy component balance and choose a full real singular-value decomposition $\mathsf{C} = U \Sigma V^{\mathsf{T}}$ with the k positive singular values first. Define*

$$C_0 = \sum_{w_\ell > 0} \frac{p_\ell q_\ell^{\mathsf{T}}}{w_\ell}, \qquad C(W) = C_0 + F_R W F_S^{\mathsf{T}}, \tag{6.9}$$

$$J(W) = \begin{pmatrix} \operatorname{diag} p & -C(W) \\ -C(W)^* & \operatorname{diag} q \end{pmatrix}, \qquad H(W) = B^+ J(W) (B^+)^*. \tag{6.10}$$

The trace-minimizing realizations are exactly $H(W)$ with

$$W = U \begin{pmatrix} I_k & 0 \\ 0 & W_0 \end{pmatrix} V^{\mathsf{T}}, \qquad W_0 \in \mathbb{C}^{(a-k)\times(b-k)}, \qquad \|W_0\|_{\mathrm{op}} \le 1. \tag{6.11}$$

The map from W_0 to $H(W)$ is injective and affine over $\mathbb{R}$. In particular,

(i) a real realization attains both minimum trace and minimum rank,

(ii) the ranks of the trace minimizers are all integers in the interval

$$\max(r_+, s_+) \le \operatorname{rank} H \le r_+ + s_+ - c_+ - k, \tag{6.12}$$

and each rank in this interval occurs for a real realization,

(iii) the trace minimizer is unique if and only if $k = \min(a, b)$.

Proof. Let u_i and v_j be the row and column sums of the edge vectors of a realization. The two families have the same total h_ℓ on each component. These totals are mutually orthogonal, with $\|h_\ell\|^2 = w_\ell$. On each positive-mass component put

$$\alpha_i = u_i - \frac{p_i}{w_\ell} h_\ell, \qquad \beta_j = v_j - \frac{q_j}{w_\ell} h_\ell \quad (i \in R_\ell,\ j \in S_\ell,\ w_\ell > 0).$$

Each residual vector is orthogonal to every component total. For row residuals this follows by writing each h_ℓ as a sum of row vectors, and for column residuals by writing it as a sum of column vectors. A vertex of zero probability has zero branch sum. The row and column residuals therefore have Gram matrices R and S. As in (6.4), their cross Gram matrix is $F_R W F_S^{\mathsf{T}}$ for some W satisfying

$$\begin{pmatrix} I_a & W \\ W^* & I_b \end{pmatrix} \succeq 0, \quad \text{equivalently} \quad \|W\|_{\mathrm{op}} \le 1. \tag{6.13}$$

The full column ranks of F_R and F_S determine W uniquely from the cross Gram matrix. Conversely, every such W gives residual vectors whose sums on each component vanish, since the columns of both factors sum to zero there. Adjoining the orthogonal totals recovers row and column vectors with matching component sums. Their signed Gram matrix $J(W)$ is positive semidefinite and has range in $\operatorname{ran} B$. Thus $BH(W)B^{\mathsf{T}} = J(W)$, and (6.10) gives a realization.

By Proposition 3.4, every trace minimizer has range in $\operatorname{ran} B^{\mathsf{T}}$. The vertex Gram $J = BHB^{\mathsf{T}}$ then determines $H = B^+ J (B^+)^*$. Since B^+ is injective on $\operatorname{ran} B$, this lift preserves the rank of J. Applying (6.5) on the residual spaces gives

$$\operatorname{tr} H(W) = D_0 - 2 \operatorname{Re} \operatorname{tr}(\mathsf{C}^* W),$$

where D_0 is independent of W. In singular-vector coordinates the quantity to be maximized is $\sum_{i=1}^k \Sigma_{ii} \operatorname{Re}(U^{\mathsf{T}} W V)_{ii}$. A contraction attains its upper bound $\sum_{i=1}^k \Sigma_{ii}$ precisely when each of the first k diagonal entries equals 1. Every row and column of a contraction has norm at most 1, so the other entries in those rows and columns must vanish. The remaining block is an arbitrary contraction. This proves (6.11).

The vertex vectors span the orthogonal direct sum of the c_+ nonzero total directions and the residual span, since each total is itself a sum of vertex vectors. By the full column ranks of the factors, the dimension of the residual span equals the rank of the block matrix in (6.13). Its kernel consists of pairs $(-Wy, y)$ with $W^*Wy = y$. If W_0 has t singular values equal to 1, counted with multiplicity, that kernel has dimension $k + t$. Hence

$$\operatorname{rank} H(W) = c_+ + a + b - k - t. \tag{6.14}$$

For any integer $0 \le t \le \min(a - k, b - k)$, take W_0 real and rectangular diagonal with t ones and all other entries zero. These choices give every rank in (6.12). The lower endpoint is $c_+ + \max(a, b) = \max(r_+, s_+)$, the minimum over all realizations by Theorem 4.2. Uniqueness holds exactly when the parameter ball is a singleton, that is, when $a - k = 0$ or $b - k = 0$. The lift gives zero edge vectors on zero-mass components. □

The minimizing face has real affine dimension $2(a-k)(b-k)$, and its real subfamily has dimension $(a-k)(b-k)$. Its extreme points are exactly its minimum-rank realizations. Indeed, if a singular value of W_0 is less than 1, perturbing that entry in both directions in singular-vector coordinates expresses W_0 as the midpoint of two distinct contractions. If all singular values equal 1, then W_0 or its adjoint is an isometry. An isometry is extreme because it takes each unit vector to an extreme point of the Hilbert unit ball. Equation (6.14) gives the rank assertion.

The supports of p and q determine k. The range of R consists of the vectors supported on $\operatorname{supp} p$ whose entries sum to zero on each component, and S has the analogous range for q. Changing the positive probabilities on fixed supports therefore changes the column bases F_R, F_S by invertible transformations and preserves the rank in (6.8). When every vertex has positive probability, k can be read from the adjacency matrix.

Corollary 6.4 (Adjacency criterion for full-support data)**.** *Suppose p, q satisfy component balance and have full support. Let N_G be the $|R| \times |S|$ bipartite adjacency matrix and let c be the number of components. Then*

$$k = \operatorname{rank} \mathsf{N}_G - c. \tag{6.15}$$

The ranks of the trace minimizers are exactly the integers in

$$\max(|R|, |S|) \;\le\; \operatorname{rank} H \;\le\; |R| + |S| - \operatorname{rank} \mathsf{N}_G. \tag{6.16}$$

The trace minimizer is unique if and only if N_G has full row rank or full column rank.

Proof. Choose componentwise factors in (6.8) so that the ranks add. On one connected component write N for the adjacency matrix and D_R, D_S for the diagonal degree matrices. Use R, S temporarily for the two vertex sets of this component. Full support gives the residual spaces $E_R = \mathbf{1}_R^{\perp}$ and $E_S = \mathbf{1}_S^{\perp}$. The rank of the corresponding block of C is the rank of

$$K : E_S \longrightarrow E_R, \qquad Kg = \Pi_{E_R} (L^{+})_{R,S} g.$$

For $g \in E_S$, let $f = L^{+}(0, g)$, so $Lf = (0, g)$. The condition $Kg = 0$ says $f_R = d\mathbf{1}_R$ for some d. Writing $f_S = d\mathbf{1}_S + h$, the two block equations for $L = \left(\begin{smallmatrix} D_R & -\mathsf{N} \\ -\mathsf{N}^{\mathsf{T}} & D_S \end{smallmatrix}\right)$ become

$$\mathsf{N}h = 0, \qquad g = D_S h.$$

Conversely, $h \in \ker \mathsf{N}$ gives $D_S h \in E_S$, because $\mathbf{1}_S^{\mathsf{T}} D_S h = \mathbf{1}_R^{\mathsf{T}} \mathsf{N} h = 0$. Choose d so that $(d\mathbf{1}_R, d\mathbf{1}_S + h)$ has total sum zero. It then equals $L^{+}(0, D_S h)$. Thus $\ker K = D_S \ker \mathsf{N}$. Since D_S is invertible, the component contributes rank

$$\operatorname{rank} K = (|S| - 1) - (|S| - \operatorname{rank} \mathsf{N}) = \operatorname{rank} \mathsf{N} - 1.$$

Summing over components gives (6.15), and Theorem 6.3 gives the remaining assertions. □

For $K_{r,s}$ the adjacency matrix has rank one. With strictly positive marginals, the face has real affine dimension $2(r-1)(s-1)$ and contains realizations of every rank from $\max(r, s)$ to $r + s - 1$. For instance, $K_{2,2}$ has trace minimizers of ranks two and three. For a cycle C_{2m} with $m \ge 2$, the adjacency matrix is I plus a cyclic permutation matrix. Its rank is m when m is odd and $m - 1$ when m is even. Thus full-support data on C_{4n+2} have a unique trace minimizer of rank $2n + 1$. On C_{4n} the minimizing face is an affine image of the closed complex unit disk. The boundary points have rank $2n$ and the interior points rank $2n + 1$.

6.3 An explicit trace–negativity tradeoff

Let G consist of two internally disjoint paths from $s \in R$ to $t \in S$ of odd lengths $a < b$, allowing $a = 1$. For Dirac marginals at s and t, flow conservation gives a constant flow α from s to t along the first path and $1 - \alpha$ along the second. Relative to the row-to-column edge orientation, the coefficients alternate in sign on each path. Thus

$$\|x(\alpha)\|_1 = a|\alpha| + b|1 - \alpha|, \tag{6.17}$$

$$\|x(\alpha)\|_2^2 = a\alpha^2 + b(1 - \alpha)^2. \tag{6.18}$$

Routing the unit mass along the shorter path gives minimum negative mass $\eta = (a - 1)/2$. Minimizing the squared norm instead gives

$$\tau = \frac{ab}{a + b}, \qquad \alpha_0 = \frac{b}{a + b}.$$

The canonical extension of every minimum-trace realization is $x(\alpha_0)$ and has negative mass

$$N_-(x(\alpha_0)) = \frac{ab}{a + b} - \frac{1}{2}.$$

For $\eta \le \nu \le ab/(a + b) - 1/2$, the least trace subject to $N_-(\mathrm{Re}(H\mathbf{1})) \le \nu$ is

$$\frac{ab}{a + b} + (a + b)\left(\frac{b - 1 - 2\nu}{b - a} - \frac{b}{a + b}\right)^2. \tag{6.19}$$

For larger ν the unconstrained minimizer is feasible and the least trace is $ab/(a + b)$. To see (6.19), first project α onto $[0, 1]$. This can only decrease both expressions in (6.17) and (6.18). On this interval the negative-mass constraint reads $\alpha \ge (b - 1 - 2\nu)/(b - a)$. In the stated range of ν the lower endpoint lies between α_0 and 1, so it minimizes the quadratic $\tau + (a+b)(\alpha - \alpha_0)^2$. Since $K = 0$ is feasible for deterministic marginals, Theorem 3.3 identifies this flow minimum with the least trace over all realizations.

For $a = 3, b = 5$, the two unrestricted optima are

$$\eta = 1, \qquad \tau = \frac{15}{8},$$

and fixing the negative mass at 1 raises the least trace to 3. For $a = 1, b = 3$, all flow along the one-edge path gives a classical extension, while $x(\alpha_0)$ has negative mass $1/4$.

6.4 Full product covers and visible Fourier characters

Let Y be a finite set of measurements and identify the d_y outcomes of each measurement y with $\mathbb{Z}_{d_y}$. Take

$$X = \prod_{y \in Y} \mathbb{Z}_{d_y}, \qquad N = |X|.$$

Let $\mathcal{M}$ be a family of coordinate contexts covering Y. The support of a product character χ_α is the set of coordinates on which its factor is nonconstant. Call χ_α *visible* if its support lies in a single context. Let $v(\mathcal{M})$ count all visible characters, including the constant character.

Theorem 6.5 (Minimum trace for product covers)**.** *For every empirical model on X that admits a strongly positive realization,*

$$\tau(e) = \frac{v(\mathcal{M})}{N}. \tag{6.20}$$

Compressing any realization to the span of the visible characters attains this trace. Thus the same minimum holds for all realizable data on a fixed product cover with fixed outcome cardinalities.

Proof. The normalized characters $f_\alpha = \chi_\alpha/\sqrt{N}$ form an orthonormal basis of $\mathbb{C}^X$. Functions of the coordinates in C are spanned both by its atom indicators and by the characters supported in C. Taking the span over all C identifies the observable subspace $W_{\mathbb{C}}$ of Proposition 3.4 with the span of the visible characters. A character visible in C has modulus one and takes a single value $\chi_\alpha(b)$ on each atom $b \in X_C$. Hence the local Gram equation gives

$$\chi_\alpha^* H \chi_\alpha = \sum_{b \in X_C} e_C(b)|\chi_\alpha(b)|^2 = 1.$$

Thus $f_\alpha^* H f_\alpha = 1/N$ for each visible character. Every other diagonal entry in the character basis is nonnegative, so $\operatorname{tr} H \geq v(\mathcal{M})/N$. By Proposition 3.4, compression to $W_{\mathbb{C}}$ preserves the local constraints. The compressed matrix has zero diagonal entries on all remaining characters, attaining the bound. □

For binary outcomes, $v(\mathcal{M})$ counts the subsets $U \subseteq Y$ contained in some context. For general cardinalities,

$$v(\mathcal{M}) = \sum_{\substack{U \subseteq Y \\ U \subseteq C \text{ for some } C \in \mathcal{M}}} \prod_{y \in U} (d_y - 1).$$

Two separate measurements with r and s outcomes therefore give $\tau = (r + s - 1)/(rs)$. For the three pair contexts on three binary measurements, only the character supported on all three measurements is invisible. Hence $v = 7$, $N = 8$, and $\tau = 7/8$.

7 Projective histories with postselection

A final projection can realize any normalized strongly positive history matrix after conditioning on a successful outcome. We use a pure initial state, one intermediate projective decomposition, and one final projection. This is a finite instance of the two-boundary histories formalism [17]. A Gram-matrix construction of the type used in [18] determines which success probabilities are possible. The largest is given by a diagonal-majorant problem familiar from the theory of coherence [20].

7.1 An exact projective realization criterion

Let ψ be a unit vector in a finite-dimensional Hilbert space. Let $(P_a)_{a \in X}$ be mutually orthogonal projections summing to I, and let F be an orthogonal projection. Write $P_A = \sum_{a \in A} P_a$. The probability of the final event when the intermediate alternatives are summed coherently is

$$s = \langle \psi | F | \psi \rangle.$$

Assume $s > 0$. Conditioning on F gives the decoherence functional

$$D_F(A, B) = \frac{\operatorname{tr}(F P_A |\psi\rangle\langle\psi| P_B F)}{s}. \tag{7.1}$$

The underlying histories have intermediate alternatives $a \in X$ and final alternatives F and $I - F$. We use the Heisenberg picture, with time evolution absorbed into the projections.

For $H \succeq 0$ with history normalization $\mathbf{1}^* H \mathbf{1} = 1$, define

$$\Gamma(H) = \min \left\{ \mathbf{1}^{\mathsf{T}} d : \operatorname{diag}(d) - H \succeq 0,\ d \in \mathbb{R}^X \right\}. \tag{7.2}$$

A feasible vector satisfies $d_a \geq H_{aa} \geq 0$ for each a. The dual program is

$$\Gamma(H) = \max \left\{ \operatorname{tr}(HZ) : Z \succeq 0,\ Z_{aa} = 1\ (a \in X) \right\}. \tag{7.3}$$

The Lagrangian is $\sum_a (1 - Z_{aa})d_a + \operatorname{tr}(HZ)$. A sufficiently large multiple of I is strictly feasible in the primal program, and $Z = I$ is strictly feasible in the dual. Thus both optima are attained and agree. The diagonal-majorant formula for the robustness of coherence C_R [20], applied to the density matrix $H/\operatorname{tr} H$, gives

$$\Gamma(H) = \operatorname{tr} H \left[1 + C_R\left(\frac{H}{\operatorname{tr} H}\right)\right]. \tag{7.4}$$

Theorem 7.1 (Projective realization and optimal success)**.** *Let $H \succeq 0$ satisfy $\mathbf{1}^* H \mathbf{1} = 1$, and fix $s > 0$ and $d \in \mathbb{R}^X$. There exists a realization of the form (7.1) with atomic matrix H, success probability s, and intermediate probabilities $\|P_a\psi\|^2 = d_a$ if and only if*

$$d \geq 0, \qquad \mathbf{1}^{\mathsf{T}} d = 1, \qquad R := \operatorname{diag}(d) - sH \succeq 0. \tag{7.5}$$

Consequently, the largest success probability for H is

$$p_{\max}(H) = \Gamma(H)^{-1}. \tag{7.6}$$

For fixed feasible s, d, the minimum dimension of a pure-state realization, with zero projections allowed for outcomes of probability zero, is

$$\operatorname{rank} H + \operatorname{rank} R. \tag{7.7}$$

An optimal-success realization has dimension at most $\operatorname{rank} H + N - 1$.

Proof. Use the history Gram convention $G_{ab} = \langle v_b | v_a \rangle$. The vectors $v_a = P_a\psi$ are mutually orthogonal, with squared norms d_a. Their projections $b_a = Fv_a$ have Gram matrix sH. The residual vectors $c_a = (I - F)v_a$ therefore have Gram matrix R, which proves necessity. The spans of (b_a) and (c_a) are orthogonal and have dimensions $\operatorname{rank} H$ and $\operatorname{rank} R$, respectively.

Conversely, choose vector families (b_a) and (c_a) with Gram matrices sH and R in Hilbert spaces of their respective ranks. In their orthogonal direct sum put

$$v_a = b_a \oplus c_a, \qquad \psi = \sum_a v_a.$$

The Gram matrix of (v_a) is $\operatorname{diag}(d)$, so the nonzero v_a are orthogonal and $\|\psi\| = 1$. Project onto the individual spans of the v_a and add their common orthogonal complement to one nonzero span. The resulting projections P_a sum to I and satisfy $P_a\psi = v_a$. Let F project onto the first summand. Then $FP_a\psi = b_a$ (identified with $b_a \oplus 0$) and

$$\|F\psi\|^2 = s\mathbf{1}^* H \mathbf{1} = s.$$

This realizes H in dimension (7.7). If every Boolean atom must have a nonzero projection, one may adjoin an unused direction for each zero projection.

If (7.5) holds, then $\operatorname{diag}(d/s) \succeq H$, so $1/s \geq \Gamma(H)$. Conversely, for a minimizing diagonal majorant $\operatorname{diag}(d_0)$ in (7.2), take $s = 1/\Gamma(H)$ and $d = d_0/\Gamma(H)$. Normalization ensures $\Gamma(H) \geq 1$, since

$$\mathbf{1}^{\mathsf{T}} d_0 \geq \mathbf{1}^* H \mathbf{1} = 1.$$

Thus $0 < s \leq 1$. At a minimizing majorant, $\operatorname{diag}(d_0) - H$ must be singular, since otherwise one could subtract a positive multiple of I and reduce the objective. The residual $R = s(\operatorname{diag}(d_0) - H)$ therefore has rank at most $N - 1$. □

The dimension bound can be sharp even at success probability one. For $m \geq 2$ and $0 < \theta < 1$, take

$$H_\theta = \frac{1-\theta}{m} I_m + \frac{\theta}{m^2} \mathbf{1}\mathbf{1}^{\mathsf{T}}.$$

This matrix has rank m and $\mathbf{1}^* H_\theta \mathbf{1} = 1$. The majorant I_m/m attains the lower bound $\Gamma(H_\theta) = 1$. For any optimal majorant D, positivity and $\mathbf{1}^*(D - H_\theta)\mathbf{1} = 0$ imply $(D - H_\theta)\mathbf{1} = 0$. Since D is diagonal, this forces $D = I_m/m$. Its residual has rank $m - 1$, so the minimum dimension at success probability one is $2m - 1$.

When $H \in \mathcal{H}(e)$, the probability s has a direct meaning for each specified context. For $C \in \mathcal{M}$, measure the intermediate projections $P_b = \sum_{a \leq b} P_a$, $b \in X_C$, followed by F. The joint and marginal success probabilities are

$$\Pr(b, F \mid C) = \|F P_b \psi\|^2 = s\, e_C(b), \qquad \Pr(F \mid C) = s.$$

The local Gram equations make the successful branch vectors mutually orthogonal. The unsuccessful branches may be combined into a single event with branch vector $(I - F)\psi$, orthogonal to each successful vector. This partition is medium decoherent on the unconditioned history space. If the intermediate measurement instead resolves every atom $a \in X$, summing the individual success probabilities gives $\sum_a \|F P_a \psi\|^2 = s \operatorname{tr} H$.

7.2 Success probability for fixed context data

For context data with $\mathcal{H}(e) \neq \varnothing$, define

$$\Gamma(e) = \min_{H \in \mathcal{H}(e)} \Gamma(H), \qquad p_{\max}(e) = \Gamma(e)^{-1}. \tag{7.8}$$

Equivalently, one minimizes $\mathbf{1}^{\mathsf{T}} d$ jointly over $H \in \mathcal{H}(e)$ and $\operatorname{diag}(d) \succeq H$. Piani et al. use the same diagonal-majorant objective for density matrices subject to linear observable constraints [20, Sec. III.D, Eq. (27)]. Here the context constraints fix $\mathbf{1}^* H \mathbf{1} = 1$ and allow $\operatorname{tr} H$ to vary. Since $d \geq 0$ and $\mathbf{1}^{\mathsf{T}} d \geq \operatorname{tr} H$, each objective sublevel set is compact. A minimizer exists, and averaging H with its complex conjugate gives a real minimizer.

Theorem 7.2 (Bounds on postselection success probability)**.** *For every normalized strongly positive realization,*

$$\max\{\operatorname{tr} H, \|H\mathbf{1}\|_1^2\} \leq \Gamma(H) \leq \left(\sum_a \sqrt{H_{aa}} \right)^2 \leq N \operatorname{tr} H. \tag{7.9}$$

For fixed realizable context probabilities on N atoms,

$$(1 + 2\eta(e))^2 \leq \Gamma(e) \leq N\tau(e), \qquad \frac{1}{N\tau(e)} \leq p_{\max}(e) \leq \frac{1}{(1 + 2\eta(e))^2}. \tag{7.10}$$

Moreover, $p_{\max}(e) = 1$ if and only if e has a classical extension. When every local distribution is deterministic and a signed extension exists,

$$p_{\max}(e) = \frac{1}{(1 + 2\eta(e))^2}. \tag{7.11}$$

Proof. Put $z = H\mathbf{1}$ and choose phases u_a such that $\overline{u_a} z_a = |z_a|$. Cauchy–Schwarz for H and the normalization $\mathbf{1}^* H \mathbf{1} = 1$ give

$$\|z\|_1^2 = |u^* H \mathbf{1}|^2 \leq u^* H u.$$

For any majorant $\operatorname{diag}(d) \succeq H$, the right side is at most $u^* \operatorname{diag}(d) u = \mathbf{1}^{\mathsf{T}} d$. Taking traces gives the other lower bound, $\operatorname{tr} H \le \mathbf{1}^{\mathsf{T}} d$. For the upper bound, let $h_a = \sqrt{H_{aa}}$. Since $|H_{ab}| \le h_a h_b$, weighted Cauchy–Schwarz gives, for every vector v,

$$v^* H v \le \left(\sum_a h_a |v_a| \right)^2 \le \left(\sum_a h_a \right) \sum_a h_a |v_a|^2.$$

Hence $\operatorname{diag}((\sum_a h_a) h)$ majorizes H. The bound $(\sum_a h_a)^2 \le N \sum_a h_a^2$ completes (7.9). For every realization, $\operatorname{Re} z$ is a signed extension, so $\|z\|_1 \ge \|\operatorname{Re} z\|_1 \ge 1 + 2\eta(e)$. For a trace minimizer the upper bound reads $\Gamma(H) \le N\tau(e)$. Together these give (7.10). If $p_{\max}(e) = 1$, then $\eta(e) = 0$ and a classical extension exists. Conversely, a classical extension g gives $H = \operatorname{diag}(g)$ with $\Gamma(H) = 1$.

The same diagonal-majorant objective, evaluated on an arbitrary rank-one matrix, satisfies

$$\Gamma(xx^*) = \|x\|_1^2. \tag{7.12}$$

Weighted Cauchy–Schwarz gives the majorant $\operatorname{diag}(\|x\|_1 |x|)$. Taking $Z = uu^*$ in (7.3), with $u_a = x_a/|x_a|$ at the nonzero coordinates and $u_a = 1$ elsewhere, gives the matching objective $\|x\|_1^2$. For deterministic data, choose a real signed extension x of minimum negative mass. Then xx^{T} realizes the data and $\|x\|_1 = 1 + 2\eta(e)$, which proves (7.11). □

For a pure state, (7.12) recovers the coherence formula of [20]. An inverse-square normalization also occurs in linear combinations of unitaries [28, Appendix A]. To realize the history matrix directly, take a real signed extension x of total mass one, set $L = \|x\|_1$, and index an orthonormal basis by the nonzero entries of x. Put

$$|\psi\rangle = \sum_a \sqrt{\frac{|x_a|}{L}} |a\rangle, \qquad |\phi\rangle = \sum_a \operatorname{sgn}(x_a) \sqrt{\frac{|x_a|}{L}} |a\rangle. \tag{7.13}$$

Take $P_a = |a\rangle\langle a|$ on the support of x, $P_a = 0$ elsewhere, and $F = |\phi\rangle\langle\phi|$. The transition amplitude through a is x_a/L. Summing the amplitudes gives success probability $1/L^2$, and conditioning on success gives the matrix xx^{T}. For the parallel odd paths of Proposition 8.2, the optimal extension has $L = \ell$. Both bounds in (7.10) then give $p_{\max} = \ell^{-2}$.

8 Explicit models and separating examples

8.1 Exact costs for an asymmetric three-atom model

Let $X = \{1, 2, 3\}$. The first context distinguishes $\{1\}$ from its complement and the second distinguishes $\{2\}$ from its complement. Assign the singleton probabilities $p, q \in [0, 1]$, respectively. Denote the realization set by $\mathcal{H}(p, q)$ and the cost (7.8) by $\Gamma(p, q)$. Put

$$\sigma_p = \sqrt{p(1-p)}, \qquad \sigma_q = \sqrt{q(1-q)}, \qquad a = pq - \sigma_p \sigma_q.$$

Proposition 8.1 (Three-atom optimization and variance mismatch)**.** *For these two contexts,*

$$\eta(p, q) = (p + q - 1)_+, \tag{8.1}$$

$$\tau(p, q) = 1 + 2a, \tag{8.2}$$

$$\min_{H \in \mathcal{H}(p,q)} I_-(H) = 2a_+, \tag{8.3}$$

$$\Gamma(p, q) = 1 + 8a_+. \tag{8.4}$$

The minimum-trace matrix is unique and real. When $p+q>1$, it also minimizes I_- and Γ, and the gaps in their lower bounds are

$$\min_H I_-(H) - \eta(1+\eta) = (\sigma_p - \sigma_q)^2, \tag{8.5}$$

$$\Gamma(p,q) - (1+2\eta)^2 = 4(\sigma_p - \sigma_q)^2. \tag{8.6}$$

Thus for $p+q>1$ each bound is attained exactly when $p=q$.

Proof. The unique signed extension is $x = (p, q, 1-p-q)^{\mathsf{T}}$, proving (8.1). The context equations and normalization force every Hermitian realization to have the form

$$H(z) = \begin{pmatrix} p & z & -z \\ \overline{z} & q & -\overline{z} \\ -\overline{z} & -z & 1-p-q+2\operatorname{Re} z \end{pmatrix}. \tag{8.7}$$

In particular, $H(z)\mathbf{1} = x$ is real. Let

$$J = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{pmatrix}.$$

The covariance factors as

$$H(z) - xx^{\mathsf{T}} = J \begin{pmatrix} \sigma_p^2 & z-pq \\ \overline{z}-pq & \sigma_q^2 \end{pmatrix} J^{\mathsf{T}}.$$

Since J has full column rank, Theorem 3.1 gives $H(z) \succeq 0$ precisely when the middle matrix is positive semidefinite, or equivalently

$$|z - pq| \le \sigma_p \sigma_q. \tag{8.8}$$

If either variance vanishes, this disk reduces to a point. In every case its unique leftmost point is $z = a$. Since $\operatorname{tr} H(z) = 1 + 2\operatorname{Re} z$, this proves (8.2) and uniqueness.

The three off-diagonal pairs give

$$I_-(H(z)) = (-\operatorname{Re} z)_+ + 2(\operatorname{Re} z)_+.$$

The disk (8.8) intersects the imaginary axis if and only if $p+q \le 1$. In that case $H(0) = \operatorname{diag}(p, q, 1-p-q)$ is feasible and has zero negative interference. If $p+q>1$, then $a>0$ and the minimum is $2a$, proving (8.3).

We next evaluate (7.2) on the real diameter of the disk, using the dual (7.3). Put $s = (1,1,-1)^{\mathsf{T}}$. When $z \ge 0$, the choices

$$Z_+ = ss^{\mathsf{T}}, \qquad D_+ = \operatorname{diag}(H_{11}+2z, H_{22}+2z, H_{33}+2z)$$

are dual and primal feasible, respectively, because

$$D_+ - H(z) = z(3I - ss^{\mathsf{T}}) \succeq 0.$$

Both objectives equal $1+8z$. Conjugating $H(z)$ by $\operatorname{diag}(s)$ makes it entrywise nonnegative, so the same value also follows from [20, Theorem 6].

When $z = -b \le 0$, take

$$Z_- = \frac{3I - ss^{\mathsf{T}}}{2}, \qquad D_- = \operatorname{diag}(H_{11}+b, H_{22}+b, H_{33}+b).$$

These are feasible, since $D_- - H(z) = bss^\mathsf{T} \succeq 0$, and both objectives equal $1 + b = 1 - z$. Hence for every feasible real z,

$$\Gamma(H(z)) = \begin{cases} 1 - z, & z \le 0, \\ 1 + 8z, & z \ge 0. \end{cases} \tag{8.9}$$

For complex z, a real diagonal majorant of $H(z)$ also dominates $H(\overline{z})$ and therefore their average $H(\operatorname{Re} z)$. Hence the minimum over the disk is attained on the real diameter. Minimizing (8.9) there gives (8.4).

Finally, for $p + q > 1$ we have $\eta = p + q - 1$, and direct expansion gives

$$2(pq - \sigma_p\sigma_q) - (p + q - 1)(p + q) = (\sigma_p - \sigma_q)^2.$$

This proves (8.5). The identity $(1 + 2\eta)^2 = 1 + 4\eta(1 + \eta)$ gives (8.6). The standard deviations agree exactly when $(p - q)(1 - p - q) = 0$, which for $p + q > 1$ reduces to $p = q$. □

For example, $p = 1$, $q = 1/2$ gives the unique realization

$$H = \begin{pmatrix} 1 & 1/2 & -1/2 \\ 1/2 & 1/2 & -1/2 \\ -1/2 & -1/2 & 1/2 \end{pmatrix}, \qquad \eta = \frac{1}{2}, \quad \tau = 2, \quad \min I_- = 1, \quad \Gamma = 5.$$

The majorant $\operatorname{diag}(2, 3/2, 3/2)$ and the dual matrix ss^T both give $\Gamma = 5$. The maximum common success probability is $1/5$, strictly below the bound $(1 + 2\eta)^{-2} = 1/4$, even on this three-atom tree.

8.2 The symmetric three-atom family

Set $p = q = t$ with $0 \le t \le 1$ in the preceding model. The unique signed extension is

$$x_t = (t, t, 1 - 2t)^\mathsf{T}. \tag{8.10}$$

With $v = (1, -1, 0)^\mathsf{T}$, the minimum-trace matrix from Proposition 8.1 is

$$H_t = x_t x_t^\mathsf{T} + t(1 - t)vv^\mathsf{T} = \begin{pmatrix} t & t(2t - 1) & t(1 - 2t) \\ t(2t - 1) & t & t(1 - 2t) \\ t(1 - 2t) & t(1 - 2t) & (1 - 2t)^2 \end{pmatrix}. \tag{8.11}$$

It is positive semidefinite and normalized, with zero interference between distinct atoms of either context. Its rank is 2 for $0 < t < 1$ and 1 at the endpoints. These are the minimum ranks in Theorem 4.2.

The vector x_t is nonnegative exactly when $t \le 1/2$. Write $\eta(t) = \eta(t, t)$ and $\tau(t) = \tau(t, t)$. Then Proposition 8.1 gives

$$\eta(t) = (2t - 1)_+, \qquad \tau(t) = 1 - 2t + 4t^2, \qquad \min_H I_-(H) = \eta(t)(1 + \eta(t)). \tag{8.12}$$

Figure 1 plots the three minima. For $0 < t < 1/2$, the parameter of H_t is $z = 2t^2 - t < 0$, so (8.9) gives

$$\Gamma(H_t) = 1 + t - 2t^2 > 1 = \Gamma(t, t).$$

For $1/2 \le t \le 1$, the matrix H_t attains both minimum trace and optimal success, with $\Gamma(H_t) = (4t - 1)^2 = \Gamma(t, t)$.

Every local outcome has positive probability for $0 < t < 1$. Hence for $1/2 < t < 1$ the classical extension fails despite full local support. At $t = 1$, the contexts assign probability

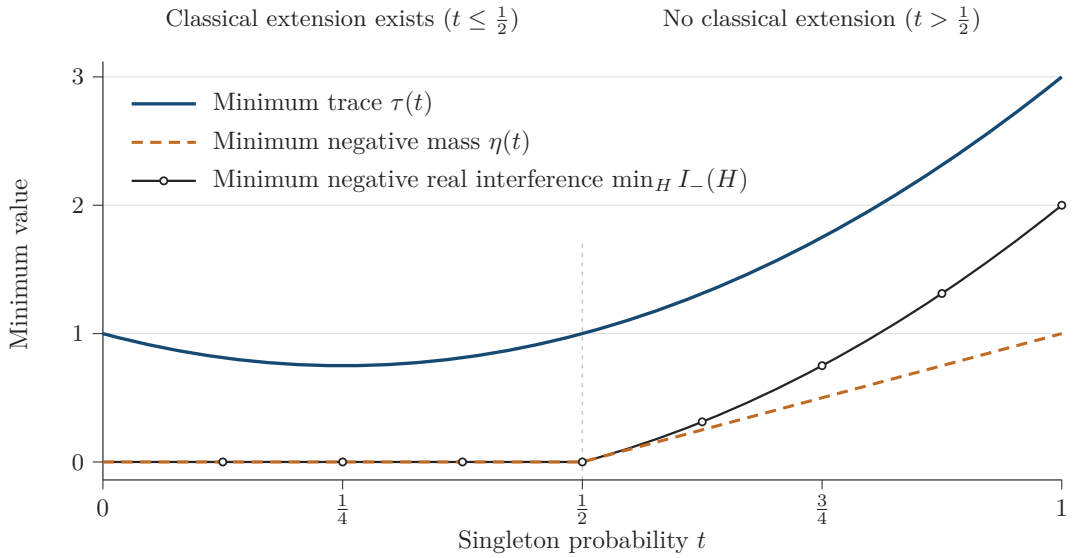


Figure 1: Minimum trace $\tau(t)$, minimum negative mass $\eta(t)$, and minimum negative real interference $\min_H I_-(H)$ for the symmetric three-atom family in (8.12). For $0 < t < 1/2$, the diagonal classical realization has zero negative real interference and the trace minimizer H_t has $I_-(H_t) > 0$. For $1/2 \le t \le 1$, H_t minimizes both trace and negative real interference.

one to the disjoint events $\{1\}$ and $\{2\}$, so $S_e = \varnothing$. This is the finite pattern of contrary inferences [23, 24]. At $t = 3/4$, for example,

$$H_{3/4} = \frac{1}{8}\begin{pmatrix} 6 & 3 & -3 \\ 3 & 6 & -3 \\ -3 & -3 & 2 \end{pmatrix}, \qquad \eta = \frac{1}{2}.$$

The overlap diagram permits all four assignments to the two binary measurements. The history algebra excludes the assignment in which both singleton events occur. Thus its classical polytope is strictly smaller, as in Proposition 2.4. For every $0 < t < 1$ the center $z = t^2$ of the feasible disk gives a positive definite realization.

8.3 Parallel odd paths

Proposition 8.2 (Simultaneous sharpness)**.** *Let G consist of $k \ge 1$ internally disjoint paths of the same odd length $\ell = 2m + 1 \ge 3$ between a row vertex s and a column vertex t. For the Dirac endpoint marginals,*

$$N = k\ell, \qquad \eta = m, \qquad \tau = \frac{\ell}{k}, \qquad \min_H I_-(H) = m(m+1). \tag{8.13}$$

There is a signed extension x attaining η for which xx^{T} attains the trace and interference minima, with equality in (5.14).

Proof. The distance between the endpoints is ℓ, so Corollary 5.3 gives $\eta = m$. Each path consists of ℓ unit resistors in series, and the k paths are connected in parallel. Hence $\mathrm{R}_{\mathrm{eff}}(s,t) = \ell/k$. Send flow $1/k$ along each path and let x be its signed coordinates in the fixed bipartite orientation. Internal vertex sums vanish and each endpoint sum is one. The $k(m+1)$ positive and km negative entries all have absolute value $1/k$, so

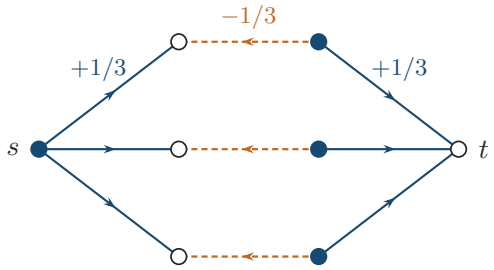
$$\|x\|_1 = \ell, \qquad \|x\|_2^2 = \ell/k, \qquad N_-(x) = m.$$

The rank-one realization xx^{T} has trace ℓ/k and negative real interference $N_-(x)(1+N_-(x)) = m(m+1)$. These attain the bounds in Theorem 5.4, since

$$\frac{(1+2\eta)^2}{N} = \frac{\ell^2}{k\ell} = \frac{\ell}{k} = \tau.$$

□

(a) Signed edge weights x_e



Arrows mark the orientation $R \to S$.

(b) Minima for $N = 3k$ atoms

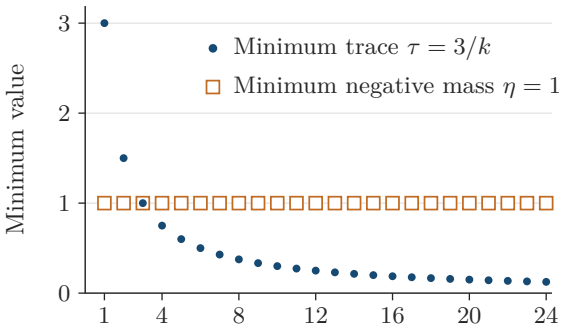


Figure 2: A unit flow from s to t on k parallel paths of length three. In (a), the arrows give the reference orientation from row to column. For $k = 3$, the successive signed edge weights x_e on each path from s to t are $+1/3, -1/3, +1/3$. In (b), the minimum trace is $\tau = 3/k$ and the minimum negative mass is $\eta = 1$ for each integer $k \geq 1$.

For fixed m, the minimum trace tends to zero as k increases, while the minimum negative mass stays at m. Figure 2 shows the case $m = 1$. Taking $k = \ell = 2m+1$ instead keeps $\tau = 1$ while η grows without bound. Thus the factor N in (5.14) cannot be replaced by an absolute constant. In each case $S_e = \varnothing$ because no edge joins the two positive-probability context atoms.

8.4 Trace minimization and postselection success probability

Proposition 8.3 (Vanishing success probability for trace minimizers of classical data)**.** *Let G consist of one edge from $s \in R$ to $t \in S$, together with $k \geq 1$ further internally vertex-disjoint paths from s to t, all of the same odd length $L \geq 3$. For the deterministic marginals (δ_s, δ_t),*

$$\eta_G = 0, \qquad \Gamma_G = 1. \tag{8.14}$$

The minimum-trace realization H_{tr} is unique and satisfies

$$\tau_G = \frac{L}{L+k}, \qquad \Gamma(H_{\mathrm{tr}}) = \left(\frac{L(k+1)}{L+k}\right)^2. \tag{8.15}$$

In particular, set $k = L^2$ and let L tend to infinity through odd integers. Then

$$\tau_G = \frac{1}{L+1} \longrightarrow 0, \qquad p_{\max}(H_{\mathrm{tr}}) = \frac{1}{\Gamma(H_{\mathrm{tr}})} = \left(\frac{L+1}{L^2+1}\right)^2 \longrightarrow 0, \tag{8.16}$$

although a realization of the same classical data succeeds with probability one.

Proof. Unit mass on the direct edge gives a classical extension. Its diagonal realization attains $\Gamma = 1$, the lower bound for every normalized H.

For deterministic context data the covariance $K = 0$ is feasible. In the decomposition $H = zz^* + K$, the trace is $\|z\|_2^2 + \operatorname{tr} K$. Its minimum requires $K = 0$ and $z = f$, the unique real signed extension of least Euclidean norm. Thus $H_{\mathrm{tr}} = ff^{\mathsf{T}}$, where f is the electrical unit flow from s to t.

Let α be its flow on the direct edge and β its flow along each long path. Equal potential drops and unit total flow give

$$\alpha = L\beta, \qquad \alpha + k\beta = 1, \qquad \alpha = \frac{L}{L+k}, \quad \beta = \frac{1}{L+k}.$$

Along a long path the signed edge coordinates alternate relative to the fixed row-to-column orientation. Consequently,

$$\|f\|_2^2 = \alpha^2 + kL\beta^2 = \frac{L}{L+k}, \qquad \|f\|_1 = \alpha + kL\beta = \frac{L(k+1)}{L+k}.$$

The first equality gives the minimum trace, while (7.12) proves the second formula in (8.15). Substitution of $k = L^2$ proves (8.16). □

Unit mass on the direct edge gives $p_{\max}(e) = 1$. Splitting the flow reduces its squared Euclidean norm, hence its trace, but increases $\|f\|_1$ above one. For the trace minimizer the optimal success probability therefore falls to $p_{\max}(H_{\mathrm{tr}}) = \|f\|_1^{-2}$, as shown in Figure 3(b).

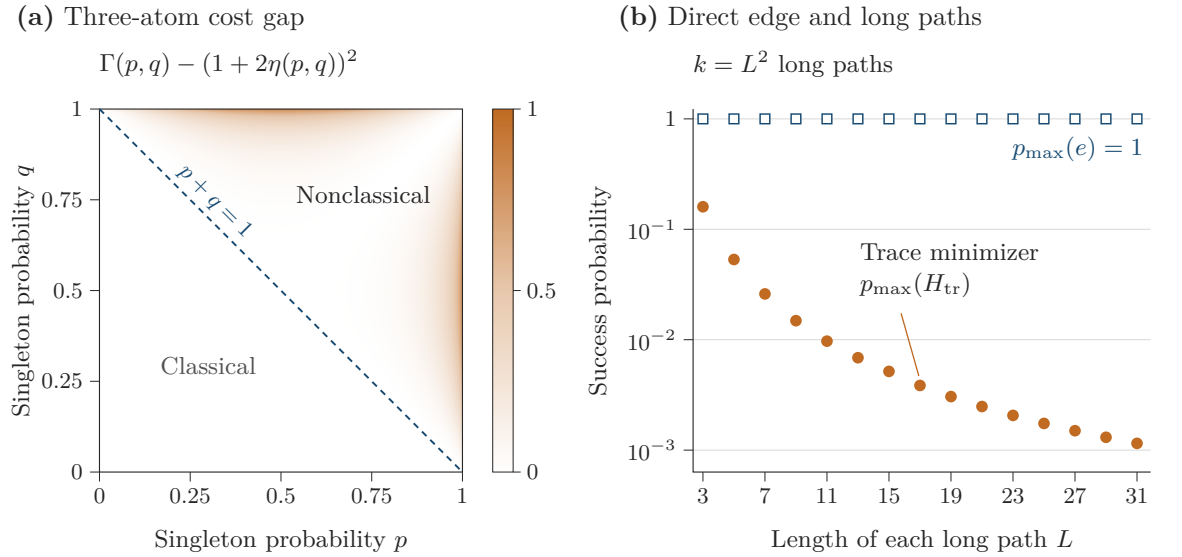


Figure 3: Postselection cost and success probability. In (a), the gap $\Gamma(p,q) - (1 + 2\eta(p,q))^2$ vanishes throughout the classical region $p + q \le 1$. For $p + q > 1$ it equals $4(\sigma_p - \sigma_q)^2$ and vanishes exactly on $p = q$. In (b), a direct edge joins the endpoints of $k = L^2$ parallel paths of odd length L. The data admit $p_{\max}(e) = 1$, while the unique trace minimizer has $p_{\max}(H_{\mathrm{tr}}) = ((L+1)/(L^2+1))^2$.

8.5 Three-time qubit histories on a common event algebra

Consider Halliwell's three-time qubit model [4, Sec. IV.C]. A rotation about the z axis takes his Hamiltonian, proportional to σ_x, to the equivalent σ_y convention used here. With the

Heisenberg observable $Q(t)$, set

$$\widehat{H}_{\mathrm{sys}} = \frac{\hbar\omega}{2}\sigma_y, \qquad Q = \sigma_z, \qquad \rho = I/2, \qquad P_q(t) = \frac{1}{2}(I + qQ(t)).$$

Choose times $t_1 < t_2 < t_3$, and set

$$X = \{\pm 1\}^3, \qquad C_{q_1q_2q_3} = P_{q_3}(t_3)P_{q_2}(t_2)P_{q_1}(t_1).$$

The atomic matrix

$$H_{\alpha\beta} = \mathrm{tr}(C_\alpha \rho C_\beta^*) \tag{8.17}$$

is positive semidefinite as a Hilbert–Schmidt Gram matrix of the operators $C_\alpha\rho^{1/2}$. Since $\sum_\alpha C_\alpha = I$, it is normalized. The contexts are the three coordinate-pair algebras. Summing over an omitted coordinate uses the identity resolution at that time. For example,

$$\sum_{q_2} C_{q_1q_2q_3} = P_{q_3}(t_3)P_{q_1}(t_1).$$

For $i < j$ the entry of the pair-context Gram matrix indexed by (a, b) and (a', b') is

$$\frac{1}{2}\,\mathrm{tr}\big(P_b(t_j)P_a(t_i)P_{a'}(t_i)P_{b'}(t_j)\big).$$

Orthogonality at t_i makes this zero if $a \neq a'$. If $b \neq b'$, cyclicity of the trace brings the two orthogonal projectors at t_j together. Thus each pair context is medium decoherent. Its diagonal entries are

$$e_{ij}(q_i, q_j) = \frac{1}{4}(1 + c_{ij}q_iq_j), \qquad c_{ij} = \cos(\omega(t_j - t_i)). \tag{8.18}$$

The one-time marginals are uniform.

More generally, let $c = (c_{12}, c_{23}, c_{13})$ be the correlations of three unbiased binary pair distributions. Define

$$F_1 = 1 - c_{12} - c_{23} + c_{13}, \qquad F_2 = 1 - c_{12} + c_{23} - c_{13},$$
$$F_3 = 1 + c_{12} - c_{23} - c_{13}, \qquad F_4 = 1 + c_{12} + c_{23} + c_{13}.$$

These are the three-time Leggett–Garg facets [29, 30]. A classical extension exists exactly when all $F_i \geq 0$. Indeed, the signed extension

$$x_0(q_1, q_2, q_3) = \frac{1}{8}(1 + c_{12}q_1q_2 + c_{23}q_2q_3 + c_{13}q_1q_3) \tag{8.19}$$

takes the four values $F_i/8$, one for each pair of opposite triples. Conversely, each deterministic triple satisfies the four inequalities, so every classical mixture does as well.

For these data, the minimum-negative-mass formula is the minimum ℓ^1-norm formula of de Barros et al. [13, Sec. 4], in the normalization $\eta = (\min_{x\in\mathcal{Q}(e)}\|x\|_1 - 1)/2$. The same facets also determine the distance to the classical polytope and its noise thresholds.

Proposition 8.4 (Exact quantities for unbiased pair data)**.** *Let $c_{ij} \in [-1, 1]$ specify the unbiased pair distributions $e_{ij}(a, b) = (1 + c_{ij}ab)/4$. Put*

$$v = \max\{0, -F_1, -F_2, -F_3, -F_4\}.$$

Then

$$\eta = \frac{v}{4}, \qquad \delta = \frac{v}{6}. \tag{8.20}$$

The smallest fraction of classical noise that makes the mixture classically extendable is $v/(4+v)$. If the noise is uniform and independent, the required fraction is $v/(1 + v)$. Whenever the data admit a strongly positive realization, the minimum trace is $7/8$.

Proof. When $v = 0$, (8.19) is a classical extension, so both distances and both noise fractions vanish. Suppose $v > 0$. Since $|c_{ij}| \le 1$, every two distinct facets sum to a nonnegative number. Exactly one facet is negative, and the two corresponding opposite triples have total mass $-v/4$ under x_0. This total is fixed by the pair marginals because the indicator of the two triples is $(1 + s_{12}q_1q_2 + s_{23}q_2q_3 + s_{13}q_1q_3)/4$, where s_{ij} are the signs in that facet. Every signed extension therefore has negative mass at least $v/4$, and x_0 attains it.

If every pair distribution is within total variation d of the corresponding marginal of one classical model, each correlation changes by at most $2d$. Correcting the negative facet thus requires $6d \ge v$. Write that facet as $1 + s \cdot c = -v$ and set $c' = c + (v/3)s$. Its new value is zero. The other facet sign vectors have inner product -1 with s, so each other value decreases by $v/3$ from at least v. Thus c' defines classical unbiased marginals, each at total variation distance $v/6$ from its original value.

Equation (5.3) gives the optimal classical-noise fraction $\eta/(1+\eta) = v/(4+v)$. Uniform independent noise with fraction λ replaces c by $(1-\lambda)c$. The negative facet becomes $-v + \lambda(1+v)$, while all other facets remain nonnegative. Finally, for any v, the constant character, three one-coordinate characters, and three pair characters are visible. Hence Theorem 6.5 gives minimum trace $7/8$ whenever a realization exists. □

Return to the qubit dynamics and take equal angular increments $\theta = \omega(t_2 - t_1) = \omega(t_3 - t_2)$. At $\theta = \pi/3$,

$$(c_{12}, c_{23}, c_{13}) = (1/2, 1/2, -1/2), \qquad v = 1/2.$$

Hence

$$\eta = \frac{1}{8}, \quad \delta = \frac{1}{12}, \quad \lambda_{\text{classical}} = \frac{1}{9}, \quad \lambda_{\text{uniform}} = \frac{1}{3}, \quad \tau = \frac{7}{8}. \tag{8.21}$$

Every pair atom has positive probability, so $S_e = X$. The four branch sums of any one pair context are nonzero and mutually orthogonal, giving rank at least four. The class operators belong to the four-dimensional Hilbert–Schmidt space of qubit operators, so their Gram matrix has rank exactly four. Compression to the seven visible characters preserves feasibility and cannot increase rank. The compressed matrix therefore has rank four and minimum trace $7/8$. The original class-operator matrix has trace one, the sum of the probabilities for a sequence of three projective measurements.

For the same physical model at $\theta = \pi/2$, the pair correlations are $(0, 0, -1)$, and

$$g(q_1, q_2, q_3) = \frac{1}{8}(1 - q_1q_3)$$

is a classical extension. For the original class-operator functional, taking $t_1 = 0$ gives

$$D((+,+,+),(+,-,+)) = -1/8.$$

For the union of these two disjoint atoms the diagonal value differs from the sum of the two diagonal values by $2(-1/8) = -1/4$. Thus the diagonal assignment fails additivity on the generated algebra. A common medium-decoherent refinement for this functional would give an additive diagonal there by restriction.

8.6 Boundary cases for general context families

With three contexts, overlap agreement can fail to give even a signed extension. On $X = \{a, b, c\}$ take the three binary contexts generated by the individual singletons. In each context assign probability one to its singleton. All pairwise intersection algebras are trivial, yet a signed extension would have $x_a = x_b = x_c = 1$ and total mass one. Hence $\mathcal{Q}(e) = \varnothing$, and Lemma 2.3 rules out a strongly positive realization.

A different obstruction can arise when $\mathcal{Q}(e)$ is nonempty. On the full four-variable binary product, take the four pair contexts $\{A_i, B_j\}$, $i, j \in \{0, 1\}$, with zero means and correlations $c_{00} = c_{01} = c_{10} = 1$, $c_{11} = -1$. The signed vector

$$x(a_0, a_1, b_0, b_1) = \frac{1}{16}(1 + a_0 b_0 + a_0 b_1 + a_1 b_0 - a_1 b_1)$$

gives these marginals. In a Gram representation, let w_x denote the branch vector of atom x and set $u_i = \sum_x a_i w_x$, $v_j = \sum_x b_j w_x$. Within each pair context its four branch sums are orthogonal. Grouping the vectors by their two outcomes therefore gives $\|u_i\| = \|v_j\| = 1$ and $\langle u_i, v_j \rangle = c_{ij}$. Cauchy–Schwarz and the parallelogram identity would imply

$$4 = \mathrm{Re}\big(\langle u_0, v_0 + v_1 \rangle + \langle u_1, v_0 - v_1 \rangle\big) \le \|v_0 + v_1\| + \|v_0 - v_1\| \le 2\sqrt{2},$$

which is impossible. This is the strong-positivity derivation of the Tsirelson bound [8, 9]. Here the affine equations admit a signed solution but the covariance equations of Corollary 3.2 are infeasible.

8.7 A strict hierarchy of extension conditions

For a fixed strongly positive realization and a generating family of medium-decoherent contexts, we have

$$\begin{aligned} &\mathfrak{G} \text{ is medium decoherent} \\ \Longrightarrow\ &(A \mapsto D(A, A)) \text{ is a probability measure on } \mathfrak{G} \\ \Longrightarrow\ &e \in \mathsf{NC}_{\mathfrak{G}} \\ \Longrightarrow\ &S_e \neq \varnothing \\ \Longrightarrow\ &X \neq \varnothing \end{aligned} \tag{8.22}$$

Each converse fails. For the first, take the two single-coordinate contexts on $\{\pm 1\}^2$. In lexicographic order write $f_0 = (1, 1, 1, 1)^{\mathsf{T}}/2$ and $f_{12} = (1, -1, -1, 1)^{\mathsf{T}}/2$. For $0 < \varepsilon < 1/4$, the matrix

$$H = \frac{1}{4} I + i\varepsilon(f_{12} f_0^{\mathsf{T}} - f_0 f_{12}^{\mathsf{T}})$$

has eigenvalues $1/4 \pm \varepsilon$ and $1/4$ (twice), so it is positive definite. Each coordinate summing matrix annihilates f_{12}, and hence each coordinate context is medium decoherent. For every event A the imaginary antisymmetric part cancels in $D(A, A)$, leaving $|A|/4$. The matrix H has nonzero imaginary off-diagonal entries, which prevent medium decoherence on the full algebra.

For the second converse, the qubit model at $\theta = \pi/2$ has a classical extension although its original diagonal assignment fails additivity. At $\theta = \pi/3$ the same model has $S_e = X$ and $\eta > 0$, disproving the third converse. Finally, the three-atom model at $t = 1$ has $X \neq \varnothing$ and $S_e = \varnothing$.

9 Discussion

The distinction between the ambient atoms X and the compatible local assignments L matters even for two contexts. In that case an element of L chooses a row and a column in the same connected component of the atom-intersection graph. It represents an element of X precisely when those vertices are adjacent. Consequently, agreement on the intersection algebra suffices for a strongly positive realization on every such graph, and suffices for a classical extension of every compatible pair of marginals exactly when each component is complete bipartite.

For unit masses at vertices s and t on opposite sides of a component, the minimum negative mass is $(d_G(s,t)-1)/2$, while the minimum trace is the effective resistance between s and t. A direct edge therefore suffices for a classical extension. Adding long paths in parallel leaves the graph distance equal to one but changes the minimum-energy flow. In the family of Proposition 8.3, the unique trace minimizer then has success probability tending to zero, even though the realization supported on the direct edge succeeds with probability one. The asymmetric three-atom example gives a different comparison. In its nonclassical region, the gap between the optimal postselection cost and its lower bound from negative mass is four times the squared difference of the Bernoulli standard deviations (Proposition 8.1).

For more than two contexts, a structural question remains. Which families of Boolean subalgebras admit a strongly positive realization for every collection of probabilities that agree on overlaps? For coordinate covers of full product spaces there are structural criteria for universal classical extendability [31, 32]. In a prescribed history algebra, a criterion for strong positivity must account for the ambient Boolean relations. The examples in Section 8 exhibit compatible data with no signed extension and data with a signed extension but no feasible covariance.

A second question concerns fixed dynamics. The projectors and state in Theorem 7.1 may be chosen to realize the prescribed matrix. If the state, Hamiltonian, or projectors are specified in advance, which trace or success optimizers can those dynamics realize? At $\theta = \pi/3$ in the qubit model of Section 8.5, the specified projector chain gives trace one. Orthogonal compression preserves all three pair distributions and lowers the trace to $7/8$. Determining when an optimizing matrix is itself the matrix of a prescribed class of histories is a separate inverse problem.

A Class operators and the endpoint trace

Let $U(t,s)$ be the unitary propagator in the Schrödinger picture. At time t_k, let $\{P_{a_k}^{(k)}\}_{a_k}$ resolve the identity into mutually orthogonal projections. For an atomic history $\alpha = (a_1,\dots,a_n)$ and an event A, the class operators are

$$K_\alpha = P_{a_n}^{(n)} U(t_n,t_{n-1}) \cdots P_{a_1}^{(1)} U(t_1,t_0), \qquad K_A = \sum_{\alpha\in A} K_\alpha.$$

Summing over all alternatives at each time gives $K_X = U(t_n,t_0)$. For an initial density operator ρ, the decoherence functional is

$$D(A,B) = \mathrm{tr}(K_A \rho K_B^*). \tag{A.1}$$

The operators $K_\alpha \rho^{1/2}$ form a Hilbert–Schmidt Gram representation of the atomic matrix, which proves strong positivity. Unitarity gives $D(X,X) = \mathrm{tr}\,\rho = 1$. In the Heisenberg picture used in Section 8.5, left multiplication by $U(t_n,t_0)^*$ removes the common final unitary and gives $C_X = I$.

When the operators have configuration-space kernels, (A.1) becomes

$$D(A,B) = \int \mathrm{d}q_f \; \mathrm{d}q_i \; \mathrm{d}q_i' K_A(q_f,q_i) \rho(q_i,q_i') \overline{K_B(q_f,q_i')}. \tag{A.2}$$

For configuration-space coarse grainings one may write the kernel formally as

$$K_A(q_f,q_i) = \int_{q(t_0)=q_i}^{q(t_n)=q_f} \mathbf{1}_A[q] e^{iS[q]/\hbar}\, \mathcal{D}q.$$

The trace in (A.2) integrates the common final configuration q_f. For finite projective histories the operator formula (A.1) already defines D. Summing class operators within the atoms of C gives the coarse-grained Gram matrix $P_C H P_C^{\mathsf{T}}$. Its off-diagonal entries vanish exactly when C is medium decoherent [1, 2].

B Finite optimization and numerical checks

For data admitting a signed extension, splitting that extension into positive and negative parts gives the linear program

$$1 + 2\eta(e) = \min_{u,v \geq 0} \mathbf{1}^{\mathsf{T}}(u + v) \quad \text{subject to} \quad M(u - v) = e. \tag{B.1}$$

Its dual is (5.2). For an arbitrary context family, minimizing $\operatorname{tr} H$ over positive semidefinite matrices satisfying (1.1) gives the minimum trace, provided the data are realizable. For two contexts, one can instead evaluate (6.2) on each positive-mass component using the graph Laplacian pseudoinverse and a singular-value decomposition. Factoring the residual covariances on their ranges gives the parametrization of the minimizing matrices in Theorem 6.3.

The classical extension test also yields a certificate for data given with error bounds. Let $\widehat{e}_C$ be probability estimates and suppose $\mathrm{TV}(e_C, \widehat{e}_C) \leq \varepsilon_C$. For y with $M^{\mathsf{T}} y \geq 0$, define

$$\operatorname{osc}(y_C) = \max_b (y_C)_b - \min_b (y_C)_b.$$

Since $e_C - \widehat{e}_C$ has total mass zero, its positive and negative parts have the same mass, at most ε_C. Hence

$$|y_C^{\mathsf{T}}(e_C - \widehat{e}_C)| \leq \varepsilon_C \operatorname{osc}(y_C).$$

If

$$y^{\mathsf{T}}\widehat{e} + \sum_C \varepsilon_C \operatorname{osc}(y_C) < 0, \tag{B.2}$$

then $y^{\mathsf{T}} e < 0$ for every probability family within these bounds. By (2.7), none has a classical extension on the prescribed algebra.

The numerical calculations compared the signed-extension and shortest-path transport programs on 80 connected bipartite graphs and checked the minimum-trace construction on the same graphs. Further checks used 101 parameter values for the three-atom family, 15 parallel-path networks, and 25 angles in the qubit model. The random seed for this group of calculations was 20261103.

To check the ranks on the minimizing face, we used seven fixed examples and 240 random graph models (seed 813749), including disconnected graphs, zero marginal probabilities, and components of zero mass. The calculation constructed 412 trace minimizers across the predicted rank intervals and tested the adjacency-rank formula on the 154 models with full support.

The postselection calculations (seed 20261104) constructed 17 projective realizations, including examples with complex atomic matrices. We checked the primal and dual certificates for 26 asymmetric three-atom parameter pairs, 31 symmetric classical cases, and 18 path networks. In the reference calculation with NumPy 2.3.5, the largest primal–dual gap was below 1.8×10^{-14} and the largest class-operator reconstruction error was below 2.2×10^{-15}.

References

[1] H. F. Dowker and J. J. Halliwell. Quantum mechanics of history: The decoherence functional in quantum mechanics. *Physical Review D*, 46(4):1580–1609, 1992. doi: 10.1103/PhysRevD.46.1580.

[2] Robert B. Griffiths. *Consistent Quantum Theory*. Cambridge University Press, Cambridge, 2002. doi: 10.1017/CBO9780511606052.

[3] C. J. Isham. Topos theory and consistent histories: The internal logic of the set of all consistent sets. *International Journal of Theoretical Physics*, 36:785–814, 1997. doi: 10.1007/BF02435786.

[4] J. J. Halliwell. Incompatible multiple consistent sets of histories and measures of quantumness. *Physical Review A*, 96(1):012123, 2017. doi: 10.1103/PhysRevA.96.012123.

[5] Samson Abramsky and Adam Brandenburger. The sheaf-theoretic structure of non-locality and contextuality. *New Journal of Physics*, 13(11):113036, 2011. doi: 10.1088/1367-2630/13/11/113036.

[6] Lutz Heindorf and Leonid B. Shapiro. *Nearly Projective Boolean Algebras*, volume 1596 of *Lecture Notes in Mathematics*. Springer, Berlin, Heidelberg, 1994. ISBN 978-3-540-58787-3. doi: 10.1007/BFb0094103. With an appendix by Sakaé Fuchino.

[7] David Milovich. Amalgamating many overlapping Boolean algebras, 2016. URL https://arxiv.org/abs/1607.07944v1. arXiv:1607.07944v1.

[8] Fay Dowker, Joe Henson, and Petros Wallden. A histories perspective on characterizing quantum non-locality. *New Journal of Physics*, 16(3):033033, 2014. doi: 10.1088/1367-2630/16/3/033033.

[9] David Craig, Fay Dowker, Joe Henson, Seth Major, David Rideout, and Rafael D. Sorkin. A Bell inequality analog in quantum measure theory. *Journal of Physics A: Mathematical and Theoretical*, 40(3):501–523, 2007. doi: 10.1088/1751-8113/40/3/010.

[10] Stan Gudder. Hilbert space representations of decoherence functionals and quantum measures. *Mathematica Slovaca*, 62(6):1209–1230, 2012. doi: 10.2478/s12175-012-0074-1.

[11] Fay Dowker and Henry Wilkes. An argument for strong positivity of the decoherence functional in the path integral approach to the foundations of quantum theory. *AVS Quantum Science*, 4 (1):012601, 2022. doi: 10.1116/5.0073587.

[12] Li Zhou, Shenggang Ying, Nengkun Yu, and Mingsheng Ying. Strassen's theorem for quantum couplings. *Theoretical Computer Science*, 802:67–76, 2020. doi: 10.1016/j.tcs.2019.08.026.

[13] J. Acacio de Barros, Ehtibar N. Dzhafarov, Janne V. Kujala, and Gary Oas. Measuring observable quantum contextuality. In Harald Atmanspacher, Thomas Filk, and Emmanuel Pothos, editors, *Quantum Interaction*, volume 9535 of *Lecture Notes in Computer Science*, pages 36–47. Springer, Cham, 2016. doi: 10.1007/978-3-319-28675-4_4.

[14] Russell Lyons and Yuval Peres. *Probability on Trees and Networks*, volume 42 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, 2016. doi: 10.1017/9781316672815. URL https://rdlyons.pages.iu.edu/prbtree/.

[15] Peter H. Schönemann. A generalized solution of the orthogonal Procrustes problem. *Psychometrika*, 31(1):1–10, 1966. doi: 10.1007/BF02289451.

[16] Rajendra Bhatia, Tanvi Jain, and Yongdo Lim. On the Bures–Wasserstein distance between positive definite matrices. *Expositiones Mathematicae*, 37(2):165–191, 2019. doi: 10.1016/j.exmath.2018.01.002.

[17] Murray Gell-Mann and James B. Hartle. Time symmetry and asymmetry in quantum mechanics and quantum cosmology. In J. J. Halliwell, J. Pérez-Mercader, and W. H. Zurek, editors, *Physical Origins of Time Asymmetry*, pages 311–345. Cambridge University Press, Cambridge, 1994. URL https://arxiv.org/abs/gr-qc/9304023.

[18] Anthony Chefles, Richard Jozsa, and Andreas Winter. On the existence of physical transformations between sets of quantum states. *International Journal of Quantum Information*, 2(1):11–21, 2004. doi: 10.1142/S0219749904000031.

[19] Carmine Napoli, Thomas R. Bromley, Marco Cianciaruso, Marco Piani, Nathaniel Johnston, and Gerardo Adesso. Robustness of coherence: An operational and observable measure of quantum coherence. *Physical Review Letters*, 116(15):150502, 2016. doi: 10.1103/PhysRevLett.116.150502.

[20] Marco Piani, Marco Cianciaruso, Thomas R. Bromley, Carmine Napoli, Nathaniel Johnston, and Gerardo Adesso. Robustness of asymmetry and coherence of quantum states. *Physical Review A*, 93(4):042107, 2016. doi: 10.1103/PhysRevA.93.042107.

[21] Sheldon Goldstein and Don N. Page. Linearly positive histories: Probabilities for a robust family of sequences of quantum events. *Physical Review Letters*, 74(19):3715–3719, 1995. doi: 10.1103/PhysRevLett.74.3715.

[22] J. J. Halliwell and J. M. Yearsley. Negative probabilities, Fine's theorem, and linear positivity. *Physical Review A*, 87(2):022114, 2013. doi: 10.1103/PhysRevA.87.022114.

[23] Adrian Kent. Consistent sets yield contrary inferences in quantum theory. *Physical Review Letters*, 78(15):2874–2877, 1997. doi: 10.1103/PhysRevLett.78.2874.

[24] Petros Wallden. Contrary inferences in consistent histories and a set selection criterion. *Foundations of Physics*, 44:1195–1215, 2014. doi: 10.1007/s10701-014-9837-6.

[25] Volker Strassen. The existence of probability measures with given marginals. *The Annals of Mathematical Statistics*, 36(2):423–439, 1965. doi: 10.1214/aoms/1177700153.

[26] Samson Abramsky, Rui Soares Barbosa, and Shane Mansfield. Contextual fraction as a measure of contextuality. *Physical Review Letters*, 119(5):050504, 2017. doi: 10.1103/PhysRevLett.119.050504.

[27] Sawyer Jack Robertson, Dhruv Kohli, Gal Mishne, and Alexander Cloninger. On a generalization of Wasserstein distance and the Beckmann problem to connection graphs. *SIAM Journal on Scientific Computing*, 47(5):A2774–A2800, 2025. doi: 10.1137/23M1624828.

[28] Andrew M. Childs and Nathan Wiebe. Hamiltonian simulation using linear combinations of unitary operations. *Quantum Information and Computation*, 12(11–12):901–924, 2012. doi: 10.26421/QIC12.11-12-1.

[29] A. J. Leggett and Anupam Garg. Quantum mechanics versus macroscopic realism: Is the flux there when nobody looks? *Physical Review Letters*, 54(9):857–860, 1985. doi: 10.1103/PhysRevLett.54.857.

[30] J. J. Halliwell and C. Mawby. Fine's theorem for Leggett–Garg tests with an arbitrary number of measurement times. *Physical Review A*, 100(4):042103, 2019. doi: 10.1103/PhysRevA.100.042103.

[31] N. N. Vorob'ev. Consistent families of measures and their extensions. *Theory of Probability and Its Applications*, 7(2):147–163, 1962. doi: 10.1137/1107014.

[32] Albert Atserias and Phokion G. Kolaitis. Consistency, acyclicity, and positive semirings. In Alessandra Palmigiano and Mehrnoosh Sadrzadeh, editors, *Samson Abramsky on Logic and Structure in Computer Science and Beyond*, volume 25 of *Outstanding Contributions to Logic*, pages 623–668. Springer, Cham, 2023. doi: 10.1007/978-3-031-24117-8_17.