Square-Root Log-Rank Standard Deviations for the Higher-Rank Kadison–Singer Problem: Polynomial-Time Algorithms via Schatten-Norm Potentials*

Zhao Song† Song Yue

September 20, 2026

Abstract

We study the matrix signing problem for positive semidefinite matrices, assigning one sign to each original matrix regardless of its rank. This includes the higher-rank Kadison–Singer partition problem [KS59] when the matrices sum to the identity. For arbitrary positive semidefinite matrices $A_1, \dots, A_m \in \mathbb{C}^{n\times n}$, we prove that there are signs $\sigma_i \in \{-1, 1\}$ such that $\|\sum_i \sigma_i A_i\| \le 3.34963 \|\sum_i \operatorname{tr}[A_i] A_i\|^{1/2}$. We give a deterministic algorithm achieving this bound in $\widetilde{O}(mn^2 + n^{4.75})$ real arithmetic operations. Inspired by the Lee–Sidford barrier, we use a potential function built from Schatten norms to obtain a square-root logarithmic dependence on the rank. Under the additional assumptions $\sum_i A_i = I$, $\|A_i\| \le \alpha$, and $\operatorname{rank}(A_i) \le r$, we obtain a discrepancy of at most $\min\{1, 3.48745\sqrt{\alpha \log(2r)}\}$. We also give a deterministic algorithm achieving a discrepancy of at most $\min\{1, 4.171\sqrt{\alpha \log(2r)}\}$ in $\widetilde{O}(mn^2 + n^{6.38})$ real arithmetic operations. This dependence on α and r is optimal up to an absolute constant when $0 < \alpha \le 1/256$ and $r \ge 256\alpha^{-6}$: even diagonal matrices can force every signing to have a discrepancy of at least $\min\{1, \sqrt{\alpha \log(2r)}\}$.

For a zero-one incidence matrix with at most d ones in each row and each column, our deterministic algorithm achieves a discrepancy of $O(\sqrt{d \log(d)})$, recovering Harvey's bound [Har15] by a different method. We give a deterministic construction of a single spanning tree that is simultaneously spectrally thin with respect to any given collection of $k \ge 1$ positive edge weightings of the same graph, assuming small edge leverage in every weighting. The tree retains the original edge weights, and its thinness bound grows only logarithmically with k.

*This paper extends the recent polynomial-time algorithm for the rank-one Kadison–Singer problem [SY26a] to the high-rank setting. Several arguments in the current draft could be replaced by references to [SY26a], but we retain them for completeness. We plan to improve the presentation in a future version of the paper.

†magic.linuxkde@gmail.com. The author would like to thank Josh Alman, Victor Reis, Omri Weinstein, and Lichen Zhang for very helpful discussions.

1 Introduction

Given Hermitian positive semidefinite matrices $A_1, \dots, A_m \in \mathbb{C}^{n\times n}$, the matrix signing problem asks for signs $\epsilon_i \in \{-1, 1\}$ that make $\|\sum_{i=1}^m \epsilon_i A_i\|$ small, where $\|\cdot\|$ denotes the operator norm. When $\sum_i A_i = I_n$, such a signing partitions the matrices into two sums close to $I_n/2$. Each original matrix receives one sign, including when its rank exceeds one.

Weaver [Wea04] connected rank-one matrix discrepancy to the Kadison–Singer problem [KS59]. Marcus, Spielman, and Srivastava [MSS15] resolved that problem and obtained dimension-independent partitions of rank-one positive semidefinite matrices. Bownik, Casazza, Marcus, and Speegle [BCMS19] sharpened the constant in Weaver's rank-one partition bound. A variance formulation was established by Kyng, Luh, and Song [KLS20]: for rank-one inputs, signs can be chosen with discrepancy at most 4σ, where $\sigma := \|\sum_i A_i^2\|^{1/2}$. Xie, Xu, and Zhu [XXZ22] improved this constant to 3. Song and Yue [SY26b] further improved this constant to 2.176 and proved that no universal constant smaller than 2 is possible, even for isotropic rank-one inputs.[1]

Very recently, three independent and concurrent works [EJ26, Kat26, SY26a] gave polynomial-time algorithms for the Kadison–Singer problem, each with a different constant.

Higher-rank bounds distinguish trace control from operator-norm control. Cohen [Coh16] and Brändén [Brä18] independently extended the partition theorem to arbitrary ranks under a bound on each input trace. In particular, when $\sum_i A_i = I_n$ and $\operatorname{tr}[A_i] \le \delta$, their results give a signing bound of $2\sqrt{2\delta} + 2\delta$. Under the assumptions $\|A_i\| \le \alpha$ and $\operatorname{rank}(A_i) \le r$, the estimate $\operatorname{tr}[A_i] \le r\alpha$ makes this bound depend on $r\alpha$.

Bownik [Bow26] extended the variance formulation to arbitrary-rank positive semidefinite matrices. For symmetric signs, his bound is $4\max\{\max_i \operatorname{tr}[A_i], \|\sum_i \operatorname{tr}[A_i]A_i\|^{1/2}\}$.

Spectrally thin spanning tree A spanning tree F is θ_{thin}-spectrally thin in a graph G if $Q_F \preceq \theta_{\text{thin}} Q_G$, where Q_F and Q_G are the corresponding Laplacians. For an unweighted graph on N vertices in which every edge has effective resistance at most α, Harvey and Olver [HO14] used pipage rounding and matrix concentration to obtain a deterministic polynomial-time construction with thinness $O(\alpha \log N/\log\log N)$. They also showed that the Kadison–Singer theorem of Marcus, Spielman, and Srivastava [MSS15], applied recursively, implies the existence of an $O(\alpha)$-spectrally thin spanning tree under the same assumption.

Anari and Oveis Gharan [AO14] extended the Kadison–Singer method to strongly Rayleigh measures. Their result gives an $O(\alpha + 1/q)$-spectrally thin spanning tree when the eligible edges form a q-edge-connected spanning subgraph and each eligible edge has effective resistance at most α in the ambient unweighted graph. In subsequent work, Anari and Oveis Gharan [AO15] introduced effective-resistance-reducing flows. For a q-edge-connected graph with $q \ge 7\log N$, they obtained a polylog$(q)/q$ spectral bound relative to a modified reference matrix whose cut quadratic forms are within constant factors of those of G. This yields combinatorial thinness in G; the spectral comparison uses the modified matrix.

Anari, Oveis Gharan, Saberi, and Srivastava [AOSS18] made these interlacing-family arguments constructive in subexponential time by approximating the largest root of a real-rooted polynomial from its leading coefficients. Applied to the strongly Rayleigh spanning-tree distribution, their method gives an $O(\alpha)$-spectrally thin tree when every edge leverage is at most α.

Strongly Rayleigh distributions and random spanning trees Strongly Rayleigh measures were introduced by Borcea, Brändén, and Liggett [BBL09] through real stable generating polynomi-

[1] [SZ26] also proves a lower bound for constant 2 for Matrix Spencer problem.

als. They established strong negative dependence properties for these measures. This class includes weighted spanning-tree distributions, in which the probability of a tree is proportional to the product of its edge weights. Kyng and Song [KS18] proved a matrix Chernoff bound for homogeneous strongly Rayleigh distributions. For an N-vertex graph and $0 < \eta < 1$, they showed that averaging $O(\eta^{-2} \log^2 N)$ independent random spanning trees, with each sampled edge reweighted by the inverse of its leverage score, gives a $(1 \pm \eta)$ spectral approximation with high probability. Kaufman, Kyng, and Soldá [KKS22] strengthened this concentration result through ℓ_∞-independence, reducing the number of trees to $O(\eta^{-2} \log N)$. These results concern spectral approximation by averages of reweighted trees. Theorem 1.8 concerns a common spanning tree that retains the original edge weights and satisfies a one-sided spectral bound simultaneously for several weightings.

Discrepancy Column sparsity is the setting of the classical Beck–Fiala theorem [BF81], which gives discrepancy $O(d_{\text{col}})$ for a zero-one matrix with at most d_{col} nonzero entries in each column, without any row-size restriction. Banaszczyk's vector-balancing theorem [Ban98] implies the bound $O(\sqrt{d_{\text{col}} \log(2m_{\text{row}})})$ for m_{row} rows. Bansal, Dadush, Garg, and Lovett [BDGL18] gave the Gram–Schmidt walk, a polynomial-time randomized algorithm achieving the same order of discrepancy. More recently, Bansal and Jiang [BJ26] obtained an algorithmic $O(\sqrt{d_{\text{col}}})$ bound when $d_{\text{col}} \geq \log^2 n_{\text{col}}$, where n_{col} is the number of columns.

When both row sizes and column degrees are bounded, the Lovász local lemma gives a bound in terms of these two parameters. Harvey [Har15] extended this approach to real matrices with entries bounded in absolute value by one, row ℓ_1-norms at most R, and column ℓ_1-norms at most Δ. For $R \geq 4$ and $\Delta \geq 2$, he proved discrepancy $O(\sqrt{R \log(R\Delta)})$. Thus the $O(\sqrt{d_0 \log(2d_0)})$ scale for zero-one matrices with at most d_0 ones in every row and column is already known. Harvey's column assumption bounds the ℓ_1-norm and allows arbitrarily many small nonzero entries.

1.1 Our results

We state our result as follows.

1.1.1 Trace-variance signing

We first control the discrepancy by the trace variance of the original matrices.

Theorem 1.1 (Informal version of Theorem 3.7)**.** *Let $A_1, \dots, A_m \in \mathbb{C}^{n \times n}$ be Hermitian positive semidefinite matrices of arbitrary ranks, and define $\sigma_{\text{tr}} := \| \sum_{i=1}^m \operatorname{tr}[A_i] A_i \|^{1/2}$. There are signs $\epsilon_1, \dots, \epsilon_m \in \{-1, 1\}$ such that $\| \sum_{i=1}^m \epsilon_i A_i \| \leq 3.34963 \sigma_{\text{tr}}$.*

A deterministic algorithm attains the same trace-variance bound.

Theorem 1.2 (Informal version of Theorem 9.1, deterministic algorithm)**.** *For the dense inputs of Theorem 1.1, a deterministic algorithm finds signs satisfying $\| \sum_{i=1}^m \epsilon_i A_i \| \leq 3.34963 \sigma_{\text{tr}}$ in $\widetilde{O}(mn^2 + n^{4.75})$ arithmetic operations.*

Randomization reduces the expected running time while preserving the discrepancy guarantee.

Theorem 1.3 (Informal version of Theorem 10.4, Las Vegas algorithm)**.** *For the dense inputs of Theorem 1.1, a Las Vegas algorithm finds signs satisfying $\| \sum_{i=1}^m \epsilon_i A_i \| \leq 3.34963 \sigma_{\text{tr}}$ in expected $\widetilde{O}(mn^2 + n^{4.50})$ arithmetic operations. Every returned signing satisfies the stated discrepancy bound.*

The input size of this problem is $O(mn^2)$. When $m = n$, we believe this problem should be solvable in almost cubic time, matching the running time for the Matrix Spencer problem [SZ26].

1.1.2 Square root log rank Discrepacny

Our second bound assumes $\sum_i A_i = I_n$ and controls the discrepancy through the operator norms and ranks of the inputs.

Theorem 1.4 (Upper bound, informal version of Theorem 11.23)**.** *Let $A_1, \dots, A_m \in \mathbb{C}^{n\times n}$ be Hermitian positive semidefinite matrices satisfying $\sum_{i=1}^m A_i = I_n$, $\|A_i\| \le \alpha$, $\operatorname{rank}(A_i) \le r$, where $r \ge 1$. There are signs $\sigma_1, \dots, \sigma_m \in \{-1, 1\}$ such that $\|\sum_{i=1}^m \sigma_i A_i\| \le \min\{1, 3.48745\sqrt{\alpha \log(2r)}\}$.*

The following lower bound shows that the dependence on α and r in Theorem 1.4 is necessary, up to an absolute constant, in the stated parameter range. It already holds for diagonal inputs.

Theorem 1.5 (Lower bound, informal version of Theorem 15.3)**.** *For every $0 < \alpha \le 1/256$ and all integers $n \ge r \ge 256\alpha^{-6}$, there are diagonal positive semidefinite matrices $A_1, \dots, A_m \in \mathbb{R}^{n\times n}$ satisfying $\sum_{i=1}^m A_i = I_n$, $\|A_i\| \le \alpha$, and $\operatorname{rank}(A_i) \le r$, such that every $\sigma \in \{-1, 1\}^m$ satisfies $\|\sum_{i=1}^m \sigma_i A_i\| \ge \min\{1, \sqrt{\alpha \log(2r)}\}$.*

Table 1: Rank-dependent discrepancy bounds under $\sum_{i=1}^m A_i = I$, $\|A_i\| \le \alpha$, and $\operatorname{rank}(A_i) \le r$. The lower bound applies in every dimension $n \ge r$ for $0 < \alpha \le 1/256$ and $r \ge 256\alpha^{-6}$.

(a) Upper bounds

Reference	Discrepancy bound
Theorem 1.4	$\min\{1, 3.48745\sqrt{\alpha \log(2r)}\}$

(b) Lower bounds

Reference	Discrepancy bound
Theorem 1.5	$\min\{1, \sqrt{\alpha \log(2r)}\}$

We next give a deterministic construction with logarithmic rank dependence.

Theorem 1.6 (Informal version of Theorem 12.15, deterministic algorithm)**.** *Under the hypotheses of Theorem 1.4, a deterministic algorithm returns signs satisfying $\|\sum_{i=1}^m \sigma_i A_i\| \le \min\{1, 4.171\sqrt{\alpha \log(2r)}\}$ using $\widetilde{O}(mn^2 + n^{6.38})$ real arithmetic operations.*

The input size of this problem is $O(mnr)$. Currently, the running time includes an mn^2 term because we explicitly form $n \times n$ matrices to evaluate the potential functions. One possible direction for future work is to avoid forming all m such matrices, perhaps by using TensorSRHT [AKK+20, SWYZ21] to maintain an approximation of the relevant matrix.

1.1.3 Spectrally thin spanning trees

The rank-dependent bound yields a common spectrally thin spanning tree for several edge weightings.

Definition 1.7 (Spectrally thin spanning tree)**.** *Let $G = (V, E)$ be a connected undirected loopless multigraph with positive edge weights τ_e, and let $\theta_{\text{thin}} \ge 0$. A spanning tree $F \subseteq E$, with every retained edge keeping its original weight, is θ_{thin}-spectrally thin in G if $Q_F \preceq \theta_{\text{thin}} Q_G$, where Q_G and Q_F are the weighted graph Laplacians. Equivalently, $x^* Q_F x \le \theta_{\text{thin}} x^* Q_G x$ for every $x \in \mathbb{R}^V$. For several weightings $\tau_e^{(\ell)}$, $\ell \in [k]$, the tree is simultaneously θ_{thin}-spectrally thin if $Q_F^{(\ell)} \preceq \theta_{\text{thin}} Q^{(\ell)}$ for every $\ell \in [k]$, where $Q^{(\ell)}$ and $Q_F^{(\ell)}$ are the Laplacians under the ℓth weighting.*

We obtain one tree for all weightings, with a thinness bound that grows logarithmically in their number.

Theorem 1.8 (Informal version of Theorem 13.18, common spectrally thin trees)**.** *Let $G = (V, E)$ be a connected undirected loopless multigraph with $N := |V| \geq 2$ vertices. For each $\ell \in [k]$, assign positive weights $\tau_e^{(\ell)} > 0$ to every edge of the same edge set, and put $Q^{(\ell)} := \sum_{e \in E} \tau_e^{(\ell)} u_e u_e^\top$, $Q_{\mathcal{E}}^{(\ell)} := \sum_{e \in \mathcal{E}} \tau_e^{(\ell)} u_e u_e^\top$. Here $\mathcal{E} \subseteq E$ is an arbitrary edge subset, u_e is a signed incidence vector, and $(Q^{(\ell)})^\dagger$ denotes the Moore–Penrose pseudoinverse. If every edge leverage satisfies $\tau_e^{(\ell)} u_e^\top (Q^{(\ell)})^\dagger u_e \leq \alpha$, there is one spanning tree $F \subseteq E$ such that, simultaneously for all $\ell \in [k]$,*

$$Q_F^{(\ell)} \preceq \min\{1, 31.6\alpha \log(2k), 4(k+1)\alpha\} Q^{(\ell)}.$$

In particular, the same tree is $31.6\alpha \log(2k)$-spectrally thin for all k weightings, with every retained edge weight unchanged.

Writing $m_{\mathrm{edge}} = |E|$ and $d_{\mathrm{blk}} = k(N-1)$, a deterministic algorithm constructs a common spanning tree satisfying $Q_F^{(\ell)} \preceq \min\{1, 268\alpha \log(2k)\} Q^{(\ell)}$ for all ℓ, with the original edge weights, in $\widetilde{O}(m_{\mathrm{edge}} d_{\mathrm{blk}}^2 + d_{\mathrm{blk}}^{6.38})$ real arithmetic operations.

The next lower bound shows that logarithmic dependence on the number of weightings is necessary up to a $\log \log k$ factor.

Theorem 1.9 (Lower bound, informal version of Corollary 16.2)**.** *There are sequences of instances satisfying the hypotheses of Theorem 1.8, with $k \to \infty$ and $\alpha \to 0$, such that every simultaneously θ_{thin}-spectrally thin spanning tree must satisfy $\theta_{\mathrm{thin}} \geq (1 - o(1))\alpha \log k / \log \log k$. Here $o(1)$ tends to zero along the sequence. These instances can additionally satisfy $\alpha \log k \to 0$.*

1.1.4 Discrepancy via fractional-coloring walks

For diagonal inputs, the rank-dependent theorem gives the following weighted discrepancy bound.

Theorem 1.10 (Informal version of Theorem 14.2, weighted row discrepancy)**.** *Let $\theta, \Lambda > 0$, and let $W \in [0, \theta]^{m_{\mathrm{row}} \times n_{\mathrm{col}}}$ have row sums at most Λ and at most $d_{\mathrm{col}} \geq 1$ nonzero entries in each column. There are signs $\sigma \in \{-1, 1\}^{n_{\mathrm{col}}}$ satisfying $\|W\sigma\|_\infty \leq \min\{\Lambda, 3.48745\sqrt{\theta\Lambda \log(2d_{\mathrm{col}})}\}$. With coefficient 4.171 in place of 3.48745, signs can be found deterministically using $\widetilde{O}(n_{\mathrm{col}} m_{\mathrm{row}} + m_{\mathrm{row}}^{3.38})$ real arithmetic operations. The same discrepancy guarantee holds after restriction to any subset of columns. In particular, a zero-one incidence matrix with at most $d_0 \geq 1$ ones in every row and every column admits discrepancy at most $\min\{d_0, 3.48745\sqrt{d_0 \log(2d_0)}\}$.*

We compare this construction with the earlier local-lemma and Beck–Fiala approaches.

Remark 1.11 (Comparison with earlier bounds)**.** The asymptotic estimate $O(\sqrt{\theta\Lambda \log(2d_{\mathrm{col}})})$ in Theorem 1.10 also follows by combining earlier results. Put $\chi := \Lambda/\theta$. For $\chi \geq 4$ and $d_{\mathrm{col}} \geq 2$, Harvey's theorem [Har15], applied to W/θ, gives $O(\theta\sqrt{\chi \log(\chi d_{\mathrm{col}})})$. The weighted Beck–Fiala bound [BF81] gives discrepancy at most $2\theta d_{\mathrm{col}}$, since every column has ℓ_1-norm at most θd_{col}. If $\chi \leq d_{\mathrm{col}}^2$, then $\log(\chi d_{\mathrm{col}}) \leq 3 \log d_{\mathrm{col}}$, so Harvey's estimate has the desired order. If $\chi > d_{\mathrm{col}}^2$, then $2\theta d_{\mathrm{col}} < 2\theta\sqrt{\chi} = 2\sqrt{\theta\Lambda}$. The case $\chi < 4$ follows from the trivial bound $\Lambda \leq 2\sqrt{\theta\Lambda}$, and when $1 \leq d_{\mathrm{col}} < 2$, each column meets at most one row, so balancing each row separately gives discrepancy at most $\min\{\theta, \Lambda\}$. These estimates also apply to every restriction to a subset of columns.

Table 2: Spectral thinness bounds for general graphs, where $N = |V|$, α bounds edge leverage, and k counts weightings. In the w/u column, w denotes weighted graphs and u denotes unweighted graphs. The 2014 Anari–Oveis Gharan bound assumes a q-edge-connected eligible subgraph with edge effective resistances at most α in the ambient unweighted graph. Their 2015 bound assumes $q \geq 7 \log N$ and compares the tree to a cut-preserving matrix D, rather than Q_G. The first five rows concern $k = 1$. The last three rows give two upper bounds and a lower bound for a common tree under k positive weightings on the same edge set, each with leverage at most α. The lower-bound Time entry is not applicable. Here lb denotes a lower bound. "Poly" and "Subexp" denote polynomial and subexponential construction time; "Existential" asserts no construction-time bound. For this paper, time counts real arithmetic operations.

Year	**Authors**	**Reference**	**w/u**	k	**Result**	**Time**
2014	Harvey and Olver	[HO14]	u	1	$O(\alpha \log N / \log\log N)$	Poly
2014	Harvey and Olver, using Marcus, Spielman, and Srivastava	[HO14] [MSS15]	u	1	$O(\alpha)$	Existential
2014	Anari and Oveis Gharan	[AO14]	u	1	$O(\alpha + 1/q)$	Existential
2015	Anari and Oveis Gharan	[AO15]	u	1	$\mathrm{polylog}(q)/q$	Existential
2018	Anari, Oveis Gharan, Saberi, and Srivastava	[AOSS18]	w	1	$O(\alpha)$	Subexp
2026	This paper	Theorem 1.8	w	k	$O(\alpha \log k)$	Poly
2026	This paper	Theorem 13.19	w	k	$O(\alpha \frac{\log(2Nk)}{\log\log(4Nk)})$	Poly
2026	This paper	Theorem 1.9	w	k	$(1 - o(1))\alpha \frac{\log k}{\log\log k}$	(lb)

Theorem 1.10 therefore provides the existence coefficient 3.48745 and a deterministic construction with coefficient 4.171 and the stated running time, through the higher-rank signing framework. We do not claim a new asymptotic discrepancy bound or optimal coefficients. Harvey's result already admits a randomized polynomial-time construction [Har15].

For zero-one matrices with at most $d_0 \geq 2$ ones in every row and column, the classical local-lemma estimate is $2\sqrt{2d_0 \log d_0}$ [Har15, Theorem 1]. Its coefficient is smaller than the one obtained here. The zero-one specialization thus illustrates the general theorem rather than improving the classical bound.

Organization. Section 2 explains the main proof ideas and algorithmic constructions. Section 3 proves the higher-rank trace-variance bound. Section 4 gives the supporting potential identities and scalar certificates. Section 5 develops the shared deterministic numerical tools. Section 6 develops root continuation and randomized numerical tools. Section 7 gives finite-precision implementations of these numerical tools. Section 8 develops the shared estimates for Schatten-norm functionals and response covariances. Section 9 gives a deterministic algorithm for trace-variance signing and proves its running-time guarantee. Section 10 gives a Las Vegas algorithm for trace-variance signing and bounds its expected running time. Section 11 proves the signing bound with logarithmic rank dependence. Section 12 gives a deterministic algorithm with rank-dependent coefficient 4.171. Section 13 derives bounds and a deterministic construction for simultaneously spectrally thin spanning trees. Section 14 applies the rank-dependent signing bound to row-constrained discrepancy and describes the resulting fractional-coloring walks. Section 15 proves rank-dependent signing lower bounds using diagonal matrices. Section 16 proves lower bounds for simultaneous spectral thinness.

2 Technical overview

2.1 Upper Bound

Section 2.1.1 explains the proof of the trace-variance bound in Theorem 1.1 and the deterministic and Las Vegas algorithms in Theorems 1.2 and 1.3. Section 2.1.2 explains the rank-dependent existence bound in Theorem 1.4 and the deterministic construction in Theorem 1.6. Section 2.1.3 explains the application to common thin spanning trees in Theorem 1.8. Section 2.1.4 explains the application to weighted row discrepancy in Theorem 1.10.

2.1.1 Trace-variance signing

Theorem 1.1 bounds discrepancy by the trace variance of the original matrices. At a minimum, its potential admits an endpoint move or a direction of negative curvature unless all signs are fixed. The deterministic algorithm in Theorem 1.2 turns this alternative into finite progress; Theorem 1.3 accelerates the direction search by random sampling.

Couple each input through its trace. Following the rank-one potential method of Song and Yue [SY26a], keep one fractional sign $y_i \in [-1, 1]$ per input and write $M(y) := \sum_{i=1}^m y_i A_i$, including signs already fixed at endpoints. Choose a nonnegative budget β with $\beta(0) = 1$ and $\beta(\pm 1) = 0$, and define

$$\mathcal{K}_y(D) := \iota \sum_{i=1}^m \beta(y_i) \operatorname{tr}[A_i D] A_i.$$

For $\rho > 0$, the potential $P_\rho(y)$ in Section 3.1 minimizes $b + \rho \operatorname{tr}[U + V]$ over $b \in \mathbb{R}$ and $U, V \succ 0$ subject to

$$U^{-1} + M(y) + \mathcal{K}_y(V) \preceq bI, \qquad V^{-1} - M(y) + \mathcal{K}_y(U) \preceq bI.$$

It dominates $\|M(y)\|$. Since $\mathcal{K}_0(I) = \iota \sum_i \operatorname{tr}[A_i] A_i$, its initial scale is $\sigma_{\mathrm{tr}}^2 := \|\sum_i \operatorname{tr}[A_i] A_i\|$. Lemma 3.1 gives a unique smooth optimizer on each open face and bounds the initial value by $2\sqrt{\iota \sigma_{\mathrm{tr}}^2 + 2n\rho}$.

Convert local matrix responses into a scalar certificate. At an optimizer, the move $y_i \to 1$ preserves feasibility if $\iota \beta(y_i) \operatorname{tr}[A_i V] \geq 1 - y_i$; the move $y_i \to -1$ does so if $\iota \beta(y_i) \operatorname{tr}[A_i U] \geq 1 + y_i$. Call a coordinate light when neither test holds. At a stationary light state, differentiate the tight matrix constraints. Sections 3.2–3.3 compress the primal and dual matrices to each input's support and normalize their traces. The resulting trace-one matrices determine local response blocks through their purities $\operatorname{tr}[S^2]$ and overlaps $\operatorname{tr}[ST]$. Lemma 3.2 converts weights on these blocks into a covariance of feasible responses.

The response cost separates into its rank-one value and corrections from these purities and overlaps. The bounds $0 \leq \operatorname{tr}[S^2] \leq 1$ and $\operatorname{tr}[ST] \leq \sqrt{\operatorname{tr}[S^2]}$ reduce the corrections to scalar maximizations. Section 3.5 evaluates them without a rank factor. We then choose

$$\beta(y) := (1 - y^2)^{3/4} f_{\mathrm{hr}}(|y|), \qquad \iota := 561/200,$$

where f_{hr} is an even polynomial of degree 20. Its slope controls the response equations, while its second derivative supplies negative curvature. Section 3.10 chooses this polynomial and the row weights jointly. Lemma 3.6 verifies the scalar comparison on the entire light domain, together with the endpoint and positivity conditions. The covariance therefore has strictly negative expected second variation.

Force a minimum to a signing. Choose a global minimizer of P_ρ with the fewest fractional coordinates. An endpoint move would reduce their number without increasing the potential. If every remaining coordinate is light, negative curvature contradicts minimality. Thus the minimizer is a signing. Sending $\rho \to 0$ along a sequence on which this signing is fixed gives

$$\|\sum_{i=1}^{m} \sigma_i A_i\| \le 2\sqrt{561/200}\,\sigma_{\mathrm{tr}} \le 3.34963\,\sigma_{\mathrm{tr}},$$

as proved in Theorem 3.7.

Construct the signing deterministically. Grouped reduction leaves at most n^2 fractional inputs, with the fixed signed contribution retained in the potential. It smooths the potential and adds a small budget reservoir to preserve curvature away from stationarity. Let F be this maintained potential and $g := \nabla F$. Section 3.17 chooses coefficients a_i so that the quadratic path $y_i(t) = y_i + th_i - t^2 a_i h_i^2$ has second derivative

$$h^\top H^\sharp h, \qquad H^\sharp := \nabla^2 F - 2\,\mathrm{diag}(a_i g_i).$$

After scaling by the response metric, an inverse-polylogarithmic fraction of the eigenvalues lie below a negative inverse-polylogarithmic threshold. A deterministic spectral filter selects a negative direction while controlling its primal and dual responses. Lemma 3.12 turns this direction into a finite step with decrease $\widetilde{\Omega}(k/n^2)$ when k coordinates remain movable.

A phase with $K/2 < k \le K$ therefore uses $\widetilde{O}(n^2/K)$ such moves. Section 3.19 deletes near-endpoint coordinates in batches, and Lemma 3.11 completes the final small instance. Charging direction selection and every optimizer computation to these phases gives $\widetilde{O}(mn^2 + n^{\max\{2\omega,9/2\}})$ operations, where ω is a matrix-multiplication exponent. The current matrix-multiplication bound of Dupont et al. [DEK+26] gives $\widetilde{O}(mn^2 + n^{4.75})$. The numerical and rounding allowances fit within the coefficient 3.34963 of Theorem 9.1.

Randomized refinement. The Las Vegas algorithm uses the same budget and accelerated curvature argument. It samples candidate directions through Gibbs features, then checks their true Rayleigh quotients and response norms. Finite trial caps and a deterministic fallback ensure correctness for every returned signing. The phase analysis in Section 10.2 reduces the expected total work to $\widetilde{O}(mn^2 + n^{4.50})$, while preserving the discrepancy guarantee of Theorem 10.4.

2.1.2 Rank-dependent signing via Schatten potentials

Function, $X \succeq 0$	**Exact formula**	**Reference**
Trace functional $f_i(X)$	$\mathrm{tr}[A_i X]$	[SY26a]
Schatten-norm functional $f_i(X)$	$(\mathrm{tr}[(A_i^{1/2} X A_i^{1/2})^s])^{1/s}$, $s \ge 2$	This paper

Theorem 1.4 replaces linear rank dependence in the trace scale by logarithmic dependence. The proof uses a Schatten-norm functional whose order balances initialization against curvature. Theorem 1.6 implements this argument with a quantitatively stable choice of budget.

Replace the trace by a Schatten-norm functional. Suppose $\sum_i A_i = I_n$, $\|A_i\| \le \alpha$, and $\mathrm{rank}(A_i) \le r$. For an integer $s \ge 2$, Section 8.1 uses

$$f_i(X) := \|A_i^{1/2} X A_i^{1/2}\|_{S_s}, \qquad \|C\|_{S_s} := (\mathrm{tr}[C^s])^{1/s} \quad (C \succeq 0).$$

Replace the trace coupling by $\mathcal{F}_y(X) := \iota \sum_i \beta(y_i) f_i(X) A_i$ in the two potential constraints. Since $f_i(I_n) \le r^{1/s}\alpha$, Section 8.2 gives

$$P_\rho(0) \le 2\sqrt{\iota r^{1/s}\alpha + 2n\rho}.$$

The gain from increasing s is useful only if the curvature argument needs coupling ι proportional to s.

Control nonlinear responses without a rank factor. After discarding zero inputs, write $Df_i(X)[H] = \operatorname{tr}[B_i(X)H]$ and define the response energy $\mathcal{E}_i(X,H) := \operatorname{tr}[B_i(X)HX^{-1}H]$. Lemmas 8.1 and 8.2 show that $0 \preceq B_i(X) \preceq A_i$, $\operatorname{tr}[XB_i(X)] = f_i(X)$, and

$$(Df_i(X)[H])^2 \le f_i(X)\mathcal{E}_i(X,H),$$
$$0 \le D^2 f_i(X)[H,H] \le (s-1)(\mathcal{E}_i(X,H) - \frac{(Df_i(X)[H])^2}{f_i(X)}).$$

These inequalities charge both derivatives of f_i to a common energy. Also $A_i^{1/2} X A_i^{1/2} \preceq f_i(X) I$ implies $\operatorname{tr}[WA_iXA_i] \le f_i(X)\operatorname{tr}[WA_i]$ for $W \succeq 0$, bounding diagonal response costs independently of r. Section 8.5 combines these estimates with the projection inequality to construct feasible response covariances.

The negative squared derivative above offsets the mixed budget–functional terms. For example, completing squares for $\beta(y) = \sqrt{1-y^2}$ leaves the budget curvature

$$-\beta''(y) - \frac{\beta'(y)^2}{(s-1)\beta(y)} = \frac{1 - y^2/(s-1)}{(1-y^2)^{3/2}}.$$

Section 8.7 proves this comparison. Section 11.6 sharpens it by retaining the actual energy diagonals and off-diagonal entries of the local response blocks. Choosing the two response weights jointly cancels mixed terms in the row-square comparison, reducing negative expected curvature to a scalar local cost below one.

Certify the budget and select the order. The remaining scalar cost depends on moments of normalized compressed matrices. For $R, S \succeq 0$ with $\operatorname{tr}[R^s] = \operatorname{tr}[S] = 1$, put $\kappa := \operatorname{tr}[R^{s+1}]$. The bounds $0 \le \kappa \le 1$ and $\operatorname{tr}[RS] \le (s+\kappa)/(s+1)$ reduce the moment comparison to endpoint values. Section 11.7 combines these bounds with budgets of the form $\sqrt{1-y^2} f(|y|)$ and verifies the required inequalities by exact rational Bernstein bounds. A positive row margin gives strict negative curvature at every stationary light state.

The compact-minimum argument from Section 11.1 now forces a signing and gives discrepancy at most $2\sqrt{\iota r^{1/s}\alpha}$. Because $\iota = O(s)$, taking s proportional to $\log(2r)$ keeps $r^{1/s}$ bounded. Section 11.9 obtains the final coefficient by sharpening the rank-three moment envelope, extending the higher-order certificate to real orders, and selecting the budget and order in each rank range. Rank one uses the variance bound of Song and Yue [SY26b], which gives discrepancy at most $2.176\sqrt{\alpha}$. Together with the trivial bound $-I_n \preceq \sum_i \sigma_i A_i \preceq I_n$, these comparisons prove $\|\sum_{i=1}^m \sigma_i A_i\| \le \min\{1, 3.48745\sqrt{\alpha \log(2r)}\}$.

Relation to Lee–Sidford barriers. A similar balance appears in the Lee–Sidford construction of barriers using Lewis weights. The ordinary logarithmic barrier has a parameter proportional to the number of constraints, whereas the Lee–Sidford construction achieves a parameter proportional

to the dimension up to logarithmic factors. Its analysis chooses an order $q = \Theta(\log m)$, keeping the factor $m^{1/(q+2)}$ bounded while incurring derivative costs polynomial in q [LS19, Section 5 and Theorem 30]. Our choice of Schatten order follows an analogous principle: increasing s reduces the initialization factor $r^{1/s}$, while the functional-Hessian estimate costs only a factor proportional to s. Here the derivative estimates establish the curvature needed for signing; the Lee–Sidford analysis establishes self-concordance of a barrier.

Construct the rank-dependent signing. Section 12.1 reduces to at most n^2 fractional matrices and normalizes the potential. Ranks one and two use the trace algorithm. For larger ranks, the construction combines the order-three budget of Section 11.5 with higher-order square-root budgets. Entropy smoothing makes the potential differentiable, and a small budget reservoir absorbs the curvature error away from stationarity. Lemma 12.2 then gives either a large weighted coordinate gradient or a negative direction in the scaled Hessian.

The walk takes a coordinate step in the first case. In the second, a spectral filter selects a direction whose relative responses permit a decrease of $\widetilde{\Omega}(k/n^2)$ when $k > n$ coordinates remain. Short steps reuse a nearby optimizer; larger curvature steps receive a fresh solve. The phase count $\widetilde{O}(n^2/K)$ for $K/2 < k \le K$, together with the batched derivative costs in Section 12.5, gives $\widetilde{O}(mn^2 + n^{\omega+4})$. The current matrix-multiplication bound of Dupont et al. [DEK+26] gives $\widetilde{O}(mn^2 + n^{6.38})$. The budget comparisons leave room for finite-accuracy computation and final rounding. Theorem 12.15 consequently attains coefficient 4.171 in place of 3.48745.

2.1.3 Spectrally thin spanning trees

Theorem 1.8 requires one tree to be thin for every weighting. A block-diagonal encoding turns the number of weightings into the matrix rank, so the signing framework yields logarithmic dependence on that number.

Encode all weightings in one matrix family. Let $N := |V|$. Restrict each Laplacian $Q^{(\ell)}$ to the orthogonal complement of the constant vector, and put

$$\psi_e^{(\ell)} := \sqrt{\tau_e^{(\ell)}}\,(Q^{(\ell)})^{-1/2} u_e, \qquad X_e := \bigoplus_{\ell=1}^{k} \psi_e^{(\ell)} \psi_e^{(\ell)*}.$$

Then $\sum_e X_e = I$, $\operatorname{rank}(X_e) \le k$, and the leverage assumption gives $\|X_e\| \le \alpha$. Selecting a subfamily of these matrices chooses the same edges in every weighting. The block dimension is $d_{\mathrm{blk}} := k(N-1)$.

Keep the support connected while reducing its upper bound. A positive lower bound in the first block certifies connectivity; upper bounds in all blocks certify simultaneous thinness. Lemma 11.12 extends the signing argument to two fixed matrix targets and a biased starting point. Their derivatives vanish, so the same response identities apply. Lemmas 13.12 and 13.14 use this freedom to reduce the upper bounds while retaining a strictly positive first-block lower bound. The Schatten order controls the factor $k^{1/\nu}$ in this comparison, and the budgets at orders three and $\nu \ge 5$ provide the required curvature.

The proof temporarily splits each X_e into M identical copies X_e/M to enter a prescribed rounding sequence. If the retained copies sum to S and $\mathcal{J}$ is their set of original indices, then

$$S \preceq \sum_{e \in \mathcal{J}} X_e \preceq MS.$$

Thus restoring each original edge once preserves connectivity and gives the required upper bound. Any spanning tree F of this support has a smaller Laplacian in every weighting and keeps the original edge weights.

Lemma 13.15 controls the initial part of the rounding sequence. The recursive interval certificate in Lemma 13.17 and the higher-order analytic estimate yield coefficient 31.6. The strongly Rayleigh comparison in Lemma 13.11 also gives $4(k+1)\alpha$. Choosing the better construction and using $Q_F^{(\ell)} \preceq Q^{(\ell)}$ proves

$$Q_F^{(\ell)} \preceq \min\{1, 31.6\alpha \log(2k), 4(k+1)\alpha\} Q^{(\ell)} \qquad (\ell \in [k]).$$

Construct a common tree deterministically. The algorithm in Theorem 13.18 repeatedly applies the deterministic rank-dependent signing to halve an edge family while maintaining connectivity. The accumulated spectral error gives

$$Q_F^{(\ell)} \preceq \min\{1, 268\alpha \log(2k)\} Q^{(\ell)}.$$

With $m_{\text{edge}} := |E|$ and $d_{\text{blk}} = k(N-1)$, the total work is $\widetilde{O}(m_{\text{edge}} d_{\text{blk}}^2 + d_{\text{blk}}^{6.38})$ using the current matrix-multiplication bound of Dupont et al. [DEK+26].

2.1.4 Discrepancy via fractional-coloring walks

Theorem 1.10 applies the rank-dependent bound to diagonal matrices. Each column becomes one input matrix, and its support size becomes the rank. The diagonal potential then gives a walk directly on fractional column colorings.

Encode column sparsity as matrix rank. For nonzero $W \in [0,\theta]^{m_{\text{row}} \times n_{\text{col}}}$, let Λ_{act} be its largest row sum and θ_0 its largest entry. Set

$$A_i := \Lambda_{\text{act}}^{-1} \operatorname{diag}(W_{1i}, \dots, W_{m_{\text{row}} i}), \qquad S := \sum_{i=1}^{n_{\text{col}}} A_i \preceq I.$$

The normalization gives $\|S\| = 1$, $\|A_i\| \le \theta_0/\Lambda_{\text{act}}$, and $\operatorname{rank}(A_i) \le d_{\text{col}}$, where d_{col} bounds the number of nonzeros per column. Lemma 14.1 extends the potential argument to $\sum_i A_i \preceq I$: the same upper bound suffices for initialization, and the local response identities are unchanged. Every signing also lies between $-S$ and S.

A sign on A_i is exactly a sign on column i, and

$$\|W\sigma\|_\infty = \Lambda_{\text{act}} \| \sum_{i=1}^{n_{\text{col}}} \sigma_i A_i \|.$$

Corollary 11.22 therefore gives discrepancy at most $\min\{\Lambda_{\text{act}}, 3.48745\sqrt{\theta_0 \Lambda_{\text{act}} \log(2d_{\text{col}})}\}$. Using $\Lambda_{\text{act}} \le \Lambda$ and $\theta_0 \le \theta$ gives the existence bound in Theorem 1.10.

Reduce the potential to scalar row variables. Averaging a feasible primal pair over diagonal sign conjugations preserves feasibility and the objective. Hence take $U = \operatorname{diag}(u)$ and $V = \operatorname{diag}(v)$, with two positive variables per row. For $\overline{W}_{ji} := W_{ji}/\Lambda_{\text{act}}$, the order-$s$ functional is

$$f_i(u) := \Big(\sum_{j=1}^{m_{\text{row}}} (\overline{W}_{ji} u_j)^s \Big)^{1/s}.$$

The two matrix constraints become two scalar inequalities per row, and entropy smoothing becomes a log-sum-exponential over these constraints. Section 14 uses endpoint moves and the gradient-or-curvature alternative to round the original column variables.

Exploit the diagonal structure in the walk. For $d_{\mathrm{col}} \geq m_{\mathrm{row}}^{1/4}$, the even-moment derandomization in Lemma 14.3 already gives the deterministic guarantee in $\widetilde{O}(n_{\mathrm{col}} m_{\mathrm{row}})$ work. For $d_{\mathrm{col}} < m_{\mathrm{row}}^{1/4}$, grouped reduction preserves $Wy = 0$ and leaves at most m_{row} fractional columns. The fixed signed contribution stays in the potential, and $0 \leq \beta \leq 1$ preserves its initial bound.

The completion in Section 14.5 uses a reservoir of order $1/(m_{\mathrm{row}} \operatorname{polylog}(m_{\mathrm{row}}))$, giving $\widetilde{O}(m_{\mathrm{row}})$ outer steps. Each functional depends on at most d_{col} rows, and its Hessian is a diagonal matrix minus a rank-one matrix. These two facts reduce the total completion cost to $\widetilde{O}(m_{\mathrm{row}}^{11/4} d_{\mathrm{col}} + m_{\mathrm{row}}^{\omega+1})$. Since $d_{\mathrm{col}} < m_{\mathrm{row}}^{1/4}$, preprocessing, completion, and final rounding together cost $\widetilde{O}(n_{\mathrm{col}} m_{\mathrm{row}} + m_{\mathrm{row}}^{\omega+1})$. The current matrix-multiplication bound of Dupont et al. [DEK+26] gives $\widetilde{O}(n_{\mathrm{col}} m_{\mathrm{row}} + m_{\mathrm{row}}^{3.38})$. Theorem 14.2 obtains coefficient 4.171 with this running time.

Deleting columns preserves the entry, row-sum, and column-support bounds, so the same construction applies to every column restriction. For a zero-one matrix with at most d_0 ones in each row and column, take $\theta = 1$, $\Lambda = d_0$, and $d_{\mathrm{col}} = d_0$. This gives existence discrepancy $\min\{d_0, 3.48745\sqrt{d_0 \log(2d_0)}\}$ and the deterministic counterpart with coefficient 4.171.

2.2 Lower Bound

Section 2.2.1 explains the rank-dependent signing lower bound in Theorem 1.5. Section 2.2.2 explains the simultaneous thin-tree lower bound in Theorem 1.9.

2.2.1 Rank-dependent signing

Theorem 1.5 uses diagonal matrices, so the obstruction already occurs in scalar row discrepancy. For an integer $k \geq 2$, choose independent uniformly random k-element subsets $S_1, \ldots, S_r$ of $[16k]$ and set

$$A_i := \frac{1}{k} \operatorname{diag}(\mathbf{1}_{\{i \in S_1\}}, \ldots, \mathbf{1}_{\{i \in S_r\}}) \qquad (1 \leq i \leq 16k).$$

The matrices sum to I_r, have norms at most $1/k$, and have ranks at most r. Their signed sum has norm $k^{-1} \max_j |\sum_{i \in S_j} \sigma_i|$.

For any fixed signing, a random row has a sufficiently large tail. For $1 \leq h \leq k$ of the same parity as k, Lemma 15.2 gives probability at least $2p_{k,h}$ for an absolute row sum of at least h, where

$$p_{k,h} := \frac{\binom{8k}{(k+h)/2}\binom{8k}{(k-h)/2}}{\binom{16k}{k}} \geq \frac{\exp(-73h^2/(100k))}{2\sqrt{k}}.$$

Independence across rows and a union bound over all 2^{16k} signings bound the probability that some signing has discrepancy below h/k by $\exp(16k \log 2 - 2rp_{k,h})$. Thus $rp_{k,h} > 8k \log 2$ guarantees a family that defeats every signing. Section 15 takes $k = \lceil 1/\alpha \rceil$ and chooses the threshold from r. Theorem 15.3, using a sharper tail estimate for small h/k, obtains $\min\{1, \sqrt{\alpha \log(2r)}\}$ for $0 < \alpha \leq 1/256$ and $r \geq 256\alpha^{-6}$.

Theorem 15.4 also isolates an asymptotic coefficient barrier. Its variant uses $2k^2$ matrices and the tail exponent $k\psi(h/k)$, where $\psi(t) = t^2/2 + O(t^4)$. Choosing $\log r$ much larger than $\log k$ but much smaller than k gives lower bound $c\sqrt{\alpha \log(2r)}$ for every fixed $c < \sqrt{2}$, along a sequence with $\alpha \log(2r) \to 0$.

2.2.2 Simultaneous spectral thinness

Theorem 1.9 fixes a family of weightings that contains a cut obstruction for every spanning tree. Section 16 starts with K_{4h} for an integer $h \geq 2$. For each balanced cut and each h-element set F of crossing edges, assign weight $4h$ inside the two sides, weight one to F, and weight $1/(4h^2)$ to the other crossing edges. Counting each unordered cut once gives

$$k_h := \binom{4h-1}{2h-1}\binom{4h^2}{h}$$

positive weightings. Each has total crossing weight $B := h + 1 - 1/(4h)$.

The dense internal cliques and a flow spread across the cut give leverage at most $1/h$ for every edge, as proved in Theorem 16.1. Every tree has a balanced cut crossed by at least $2h$ tree edges: this is their expected number for a uniformly random balanced cut. Select F from these edges. In its corresponding weighting, the tree has crossing weight at least h. Testing the spectral inequality on the cut indicator forces thinness at least h/B.

Finally, replace each edge by M parallel copies of the same weight. The graph Laplacians increase by a factor M, reducing every copy's leverage to at most $\alpha := 1/(hM)$. A tree uses at most one copy of each original edge, so its forced cut ratio is at least $h/(MB)$. Consequently $\theta_{\text{thin}}/\alpha \geq h^2/B = (1 - o(1))h$. Since $\log k_h = h \log h + O(h)$, Corollary 16.2 gives

$$\theta_{\text{thin}} \geq (1 - o(1))\alpha \frac{\log k_h}{\log \log k_h}.$$

Taking $M = h^2$ also ensures $\alpha \log k_h \to 0$. Thus the obstruction persists even when the upper bound tends to zero, leaving a $\log \log k$ gap in the dependence on the number of weightings.

3 Higher-rank signing in terms of trace variance

Let $A_1, \ldots, A_m$ be positive semidefinite Hermitian matrices of arbitrary ranks. Define

$$t_i := \operatorname{tr}[A_i], \qquad T := \sum_{i=1}^{m} t_i A_i, \qquad \sigma_{\text{tr}} := \|T\|^{1/2}. \tag{1}$$

The rank-one potential argument is due to Song and Yue [SY26a]; its statements and proofs are reproduced in Section 4. Here the trace coupling controls the variance in Eq. (1) for inputs of arbitrary rank.

3.1 The trace coupling and the regularized potential

For the preliminary scalar construction set

$$\iota := \frac{699}{250}, \qquad \eta := 1 + 10^{-6}, \qquad \beta(y) := (1 - y^2)^{3/4} f(|y|), \tag{2}$$

where the exact rank-one polynomial, reproduced in Section 4.8, is

$$\begin{aligned} f(z) := 1 &+ 0.1915505z^2 + 0.07716335z^4 + 0.03917979z^6 + 0.04234609z^8 - 0.08891557z^{10} \\ &+ 0.40792039z^{12} - 0.91331458z^{14} + 1.30793628z^{16} - 1.02145746z^{18} + 0.35713938z^{20}. \end{aligned} \tag{3}$$

These decimals denote rational numbers. The scalar budget satisfies $\beta(0) = 1$, $\beta(\pm 1) = 0$, $1 - y^2 \le \beta(y) \le 1$, $-\beta'' > 1/2$, and $0 \le -\beta'(z)/\beta(z) \le 3/[4(1-z)]$ for $0 \le z < 1$. These properties control the regularity of the potential and its response scales.

For rank one, $A_i D A_i = \operatorname{tr}[A_i D] A_i$. For arbitrary ranks, define the trace coupling by

$$M(y) := F_0 + \sum_{i \in I} y_i A_i, \qquad \mathcal{K}_y(D) := \iota \sum_{i \in I} \beta(y_i) \operatorname{tr}[A_i D] A_i. \tag{4}$$

Here I is the active set and F_0 is the fixed contribution on a face. The map $\mathcal{K}_y$ is positive and self-adjoint for the real trace inner product on Hermitian matrices. For $\rho > 0$ define

$$\begin{aligned} P(y) := \min_{b \in \mathbb{R},\ U,V \succ 0} \ & b + \rho \operatorname{tr}[U + V] \\ \text{subject to} \quad & U^{-1} + M(y) + \mathcal{K}_y(V) \preceq bI, \\ & V^{-1} - M(y) + \mathcal{K}_y(U) \preceq bI. \end{aligned} \tag{5}$$

The regularized potential has a unique optimizer on every face, which permits the response calculations below.

Lemma 3.1 (Regularity and initial value)**.** *On every open face, the optimizer in Eq. (5) is unique, is twice continuously differentiable, and has locally Lipschitz second derivatives. The positive definite dual multipliers W, Z satisfy*

$$\begin{gathered} W = U\mathcal{K}_y(Z)U + \rho U^2, \qquad Z = V\mathcal{K}_y(W)V + \rho V^2, \\ \operatorname{tr}[W + Z] = 1. \end{gathered} \tag{6}$$

$$\begin{aligned} U^{-1} + M + \mathcal{K}_y(V) &= bI, \\ V^{-1} - M + \mathcal{K}_y(U) &= bI. \end{aligned} \tag{7}$$

The value is continuous on the closed cube. On the initial face, with $F_0 = 0$, it satisfies

$$\|M(y)\| \le P(y), \qquad P(0) \le 2\sqrt{\iota\sigma_{\mathrm{tr}}^2 + 2n\rho}. \tag{8}$$

Proof. The inverse map is operator convex, and $U := V := I$ with large b is strictly feasible. On any bounded objective sublevel, the constraints bound the inverse matrices and the trace penalty bounds the matrices themselves. Thus the minimum exists inside the positive cone. Differentiating the Lagrangian gives Eq. (6), and $W, Z \succ 0$ force Eq. (7) by complementary slackness. For a nonzero Hermitian H,

$$\frac{d^2}{ds^2} \operatorname{tr}[W(U + sH)^{-1}]|_{s=0} = 2 \operatorname{tr}[WU^{-1}HU^{-1}HU^{-1}] > 0.$$

This gives strict convexity in (U, V) and uniqueness.

Let

$$J(H_1, H_2) := (U^{-1}H_1U^{-1} - \mathcal{K}_y(H_2), V^{-1}H_2V^{-1} - \mathcal{K}_y(H_1)).$$

The positive map $N(H_1, H_2) := (U\mathcal{K}_y(H_2)U, V\mathcal{K}_y(H_1)V)$ sends (W, Z) strictly below itself by Eq. (6). Its spectral radius is less than one, so J^{-1} exists and is positive. Let E be the positive definite Hessian of the Lagrangian in (U, V). A kernel vector of the full KKT Jacobian satisfies $JH = -d_b(I, I)$, $J^*Q' = EH$, and $\operatorname{tr}[Q_1' + Q_2'] = 0$. Pairing these equations gives $\langle H, EH \rangle = 0$, hence every component of the kernel vector vanishes. The implicit function theorem proves the asserted regularity. Compact sublevels and a vanishing increase in b prove continuity at a boundary face.

The constraints dominate both M and $-M$. Finally use $U := V := \ell I$ at zero and $\mathcal{K}_0(I) = \iota T$. Minimizing $\ell^{-1} + (\iota\sigma_{\mathrm{tr}}^2 + 2n\rho)\ell$ gives Eq. (8). □

Put $u_i := \operatorname{tr}[A_i U]$, $v_i := \operatorname{tr}[A_i V]$, $w_i := \operatorname{tr}[A_i W]$, and $z_i := \operatorname{tr}[A_i Z]$. Zero inputs can be discarded. A move to 1 is feasible at the same (b, U, V) whenever $\iota\beta_i v_i \geq 1 - y_i$; a move to -1 is feasible whenever $\iota\beta_i u_i \geq 1 + y_i$. Indeed the change in the first constraint for the positive move is $[(1 - y_i) - \iota\beta_i v_i]A_i \preceq 0$, and the second constraint also improves. Consequently the remaining case is the light state

$$\iota\beta_i u_i < 1 + y_i, \qquad \iota\beta_i v_i < 1 - y_i. \tag{9}$$

Define

$$p_i := 1 + \iota\beta_i' v_i, \quad q_i := 1 - \iota\beta_i' u_i, \quad c_i := \sqrt{\beta_i p_i / q_i}, \quad d_i := \sqrt{\beta_i q_i / p_i}, \quad \tau_i := \sqrt{p_i q_i / \beta_i}. \tag{10}$$

These numbers are positive. For example, if $y_i \geq 0$, the logarithmic derivative bound gives $p_i > 1/4$ and $q_i \geq 1$; the roles interchange when $y_i < 0$.

3.2 First and second variations

Hold b fixed and differentiate the tight constraints in a direction $r \in \mathbb{R}^I$. The resulting directional derivatives $\dot{U}$ and $\dot{V}$ are uniquely determined by

$$\begin{aligned} U^{-1}\dot{U}U^{-1} &= \sum_{i \in I} p_i r_i A_i + \mathcal{K}_y(\dot{V}), \\ V^{-1}\dot{V}V^{-1} &= -\sum_{i \in I} q_i r_i A_i + \mathcal{K}_y(\dot{U}). \end{aligned} \tag{11}$$

Define

$$Q(r) := \operatorname{tr}[W U^{-1}\dot{U}U^{-1}\dot{U}U^{-1}] + \operatorname{tr}[Z V^{-1}\dot{V}V^{-1}\dot{V}V^{-1}]. \tag{12}$$

Differentiating the feasible fixed-b comparison path gives

$$\partial_i P = p_i w_i - q_i z_i, \tag{13}$$

$$\frac{1}{2} r^{\mathsf{T}} \nabla^2 P r \leq Q(r) + \sum_{i \in I} \iota\beta_i' r_i (w_i \dot{v}_i + z_i \dot{u}_i) - \frac{\iota}{2} \sum_{i \in I} (-\beta_i'')(v_i w_i + u_i z_i) r_i^2. \tag{14}$$

Here $\dot{u}_i = \operatorname{tr}[A_i \dot{U}]$ and similarly for $\dot{v}_i$. These are equalities for the comparison path and an inequality for the optimized value. The identity $J(W, Z) = \rho(I, I)$ eliminates its second primal derivatives.

We will analyze a stationary light state: $\partial_i P = 0$ for every active i. This is enough to exclude an interior minimum on a face. Define

$$x_i := c_i u_i, \qquad y_i^\flat := d_i v_i, \qquad \omega_i := c_i w_i = d_i z_i, \qquad \delta_i := \iota x_i y_i^\flat, \qquad \Gamma_i := \delta_i^{-1}. \tag{15}$$

The notation $y_i^\flat$ is distinct from the signing coordinate y_i. Below, within one coordinate, abbreviate $x := x_i$ and $v := y_i^\flat$. Lightness implies $\Gamma_i > \iota > 2$.

3.3 Response geometry and local density matrices

Define PSD, entrywise nonnegative real Gram matrices

$$\mathsf{U}_{ij} := \iota c_i c_j \operatorname{tr}[A_i U A_j U], \qquad \mathsf{V}_{ij} := \iota d_i d_j \operatorname{tr}[A_i V A_j V]. \tag{16}$$

For real coefficient vectors a, b, put $M_a := \sum_{i\in I} c_i a_i A_i$ and $M_b := \sum_{i\in I} d_i b_i A_i$. The directional derivatives $\dot U = U M_a U$, $\dot V = V M_b V$ satisfy Eq. (11) precisely when

$$J_c := a - \mathsf{V} b = \mathsf{U} a - b, \qquad r_i = (J_c)_i/\tau_i. \tag{17}$$

The dual equations give $\mathsf{U}(d_i z_i)_i < (c_i w_i)_i$ and $\mathsf{V}(c_i w_i)_i < (d_i z_i)_i$, so $\rho(\mathsf{UV}) < 1$. Thus this response space has dimension $k := |I|$.

Let M_{raw} be the matrix of the PSD quadratic form $\mathrm{tr}[W M_a U M_a] + \mathrm{tr}[Z M_b V M_b]$. Set

$$E_a := \mathrm{diag}(c_i^2 u_i w_i), \quad E_b := \mathrm{diag}(d_i^2 v_i z_i), \quad E := \mathrm{diag}(E_a, E_b),$$

$$D_0 := E^{-1/2} M_{\mathrm{raw}} E^{-1/2}, \qquad \mathsf{T} := \begin{pmatrix} 0 & E_a^{1/2}\mathsf{V}E_b^{-1/2} \\ E_b^{1/2}\mathsf{U}E_a^{-1/2} & 0 \end{pmatrix}. \tag{18}$$

We always have

$$D_0 \succeq \mathsf{T}^{\mathsf{T}} \Gamma \mathsf{T}, \qquad \Gamma := \mathrm{diag}(\Gamma_i, \Gamma_i)_{i\in I}. \tag{19}$$

Indeed, for PSD A, positive definite U, and Hermitian H,

$$|\,\mathrm{tr}[AH]|^2 \le \mathrm{tr}[AU]\,\mathrm{tr}[AHU^{-1}H]. \tag{20}$$

Apply this to $H := U M_a U$ and use Eq. (6) to obtain

$$\mathrm{tr}[W M_a U M_a] \ge \iota \sum_{i\in I} \frac{\beta_i z_i}{u_i}\,\mathrm{tr}[A_i U M_a U]^2 = \sum_{i\in I} \Gamma_i (E_b)_{ii} (\mathsf{U}a)_i^2.$$

The other block is identical, proving Eq. (19).

For each coordinate introduce the trace-one PSD matrices

$$R_U := \frac{A_i^{1/2} U A_i^{1/2}}{u_i}, \quad R_V := \frac{A_i^{1/2} V A_i^{1/2}}{v_i}, \quad R_W := \frac{A_i^{1/2} W A_i^{1/2}}{w_i}, \quad R_Z := \frac{A_i^{1/2} Z A_i^{1/2}}{z_i}, \tag{21}$$

and define

$$\chi_U := \mathrm{tr}[R_U^2], \quad \chi_V := \mathrm{tr}[R_V^2], \quad e_U := \mathrm{tr}[R_U R_W], \quad e_V := \mathrm{tr}[R_V R_Z]. \tag{22}$$

For rank r_i they satisfy $1/r_i \le \chi_U, \chi_V \le 1$ and $0 < e_U \le \sqrt{\chi_U}$, $0 < e_V \le \sqrt{\chi_V}$. The latter inequalities follow from Hilbert–Schmidt Cauchy–Schwarz and $\mathrm{tr}[R_W^2], \mathrm{tr}[R_Z^2] \le 1$.

In the normalized coordinate pair define the orthonormal vectors

$$e_i := \frac{\sqrt{x_i}\, e_{a,i} - \sqrt{y_i^\flat}\, e_{b,i}}{\sqrt{x_i + y_i^\flat}}, \qquad f_i := \frac{\sqrt{y_i^\flat}\, e_{a,i} + \sqrt{x_i}\, e_{b,i}}{\sqrt{x_i + y_i^\flat}}.$$

The response projection Π satisfies $(I - \Pi_e)(I - \mathsf{T})\Pi = 0$, where Π_e projects onto the e_i's. The exact local blocks, in this paired basis, are

$$C_i := \frac{1}{x+v} \begin{pmatrix} x e_U + v e_V & \sqrt{xv}(e_U - e_V) \\ \sqrt{xv}(e_U - e_V) & v e_U + x e_V \end{pmatrix}, \tag{23}$$

$$T_i := \frac{\delta}{x+v} \begin{pmatrix} -(x\chi_U + v\chi_V) & \sqrt{xv}(\chi_V - \chi_U) \\ (x^2\chi_U - v^2\chi_V)/\sqrt{xv} & x\chi_U + v\chi_V \end{pmatrix}. \tag{24}$$

These follow by changing basis in $\mathrm{diag}(e_U, e_V)$ and in the off-diagonal local block of Eq. (18). In rank one, $C_i = I_2$ and $T_i = \left(\begin{smallmatrix} -\delta & 0 \\ L & \delta \end{smallmatrix}\right)$, where $L := \delta(x - v)/\sqrt{xv}$. Both local blocks depend on the density invariants.

3.4 The local row certificate

At a stationary state set

$$\Lambda_i := \frac{\beta_i'/\beta_i}{\tau_i\sqrt{x_i y_i^\flat}}, \qquad k_i := \frac{\iota(-\beta_i'')}{2\eta p_i q_i}, \qquad s_i := 1/k_i. \tag{25}$$

For a normalized response $\xi := E^{1/2}(a,b)$, define $Y_i := e_i^{\mathsf{T}}(I-\mathsf{T})\xi$ and $F_i := f_i^{\mathsf{T}}\xi$. Eq. (17) gives

$$Y_i^2 = (x_i + y_i^\flat)\omega_i\tau_i^2 r_i^2, \qquad \iota\beta_i' r_i(w_i\dot{v}_i + z_i\dot{u}_i) = \Lambda_i Y_i F_i. \tag{26}$$

To check the mixed identity, use $\iota c_i\dot{u}_i = (\mathsf{U}a)_i$ and $\iota d_i\dot{v}_i = (\mathsf{V}b)_i$. At stationarity its left side is $(\beta_i'/\beta_i)\omega_i(J_c)_i(a_i+b_i)/\tau_i$. Substitution of the paired basis proves Eq. (26).

The next lemma converts a local 2-by-2 matrix inequality into a response covariance with the required second-variation bound.

Lemma 3.2 (Local row certificate)**.** *For each active coordinate choose real numbers a,b,r,d and put $v_* := (a,b)^{\mathsf{T}}$, $w_* := (r,d)^{\mathsf{T}}$. With $\delta := \delta_i$, $s := s_i$, and $\Lambda := \Lambda_i$, define*

$$\begin{aligned} Z_{11} &:= 2a - (s-\delta)a^2 + \delta r^2, \\ Z_{12} &:= b - \tfrac{1}{2}\Lambda s a - (s-\delta)ab + r + \delta rd, \\ Z_{22} &:= -\Lambda s b - (s-\delta)b^2 - \tfrac{1}{4}\Lambda^2 s + 2d + \delta d^2. \end{aligned} \tag{27}$$

Suppose $Z_i \succ 0$, define $D_i := Z_i^{-1}$, and assume

$$\mathcal{E}_i := \operatorname{tr}[C_i D_i] + \delta_i(v_*^{\mathsf{T}} D_i v_* + w_*^{\mathsf{T}} D_i w_*) + 2\sum_{j=1}^{2}(T_i)_{j,:}D_i \begin{cases} v_*, & j=1, \\ w_*, & j=2 \end{cases} \le 1. \tag{28}$$

Then there is a centered finite-support response distribution with nonzero movement for which

$$\mathbb{E}[Q(r)] + \sum_{i\in I}\Lambda_i\mathbb{E}[Y_iF_i] \le \sum_{i\in I} k_i\mathbb{E}[Y_i^2]. \tag{29}$$

Proof. Let D be the block diagonal matrix with blocks D_i in the paired basis. If G has orthonormal columns spanning the response space, set

$$A := G(G^{\mathsf{T}}D^{-1}G)^{-1}G^{\mathsf{T}}, \qquad B := D - A.$$

Then $A, B \succeq 0$, $AD^{-1}A = A$, $AD^{-1}B = 0$, and $\operatorname{tr}[D^{-1}A] = k$. Choose a centered finite-support random vector ξ with covariance A. It is supported on the response space. By Eq. (19),

$$\operatorname{tr}[D_0 A] \le \operatorname{tr}[D_0 D] - \operatorname{tr}[\Gamma\mathsf{T}B\mathsf{T}^{\mathsf{T}}]. \tag{30}$$

Fix one pair and define $t_e := e_i^{\mathsf{T}}\mathsf{T}$, $t_f := f_i^{\mathsf{T}}\mathsf{T}$. In the formulas below, v_*, w_* are embedded into this pair. Set $C_* := \Gamma_i B + k_i A \succ 0$. The projection identities give

$$C_*^{-1} = \delta_i D^{-1} + (s_i - \delta_i)D^{-1}AD^{-1}.$$

Completing the square in t_e gives

$$\Gamma_i t_e B t_e^{\mathsf{T}} + k_i(e_i^{\mathsf{T}} - t_e)A(e_i - t_e^{\mathsf{T}}) - \Lambda_i(e_i^{\mathsf{T}} - t_e)Af_i + 2t_e D v_*$$

$$\geq 2(e_i - \tfrac{1}{2}\Lambda_i s_i f_i)^{\mathsf{T}} A v_* - (s_i - \delta_i) v_*^{\mathsf{T}} A v_* - \tfrac{1}{4}\Lambda_i^2 s_i f_i^{\mathsf{T}} A f_i - \delta_i v_*^{\mathsf{T}} D v_*.$$

The response identity $t_f A = f_i^{\mathsf{T}} A$ and a square in the PSD matrix B give

$$\Gamma_i t_f B t_f^{\mathsf{T}} + 2 t_f D w_* \geq 2 f_i^{\mathsf{T}} A w_* + \delta_i w_*^{\mathsf{T}} A w_* - \delta_i w_*^{\mathsf{T}} D w_*.$$

The summed coefficient of the local block of A is exactly Z_i in Eq. (27). Sum over all pairs and use $\operatorname{tr}[D^{-1}A] = k$. Also, $\operatorname{tr}[D_0 D] = \sum_{i \in I} \operatorname{tr}[C_i D_i]$. Combining with Eq. (30) gives

$$\sum_{i \in I} k_i \mathbb{E}[Y_i^2] - \sum_{i \in I} \Lambda_i \mathbb{E}[Y_i F_i] - \mathbb{E}[Q(r)] \geq \sum_{i \in I} (1 - \mathcal{E}_i) \geq 0.$$

Finally $(I - \mathsf{T})$ has no kernel on the response space: zero movement in Eq. (17) would give $a = \mathsf{V}b$ and $b = \mathsf{U}a$, contradicting $\rho(\mathsf{UV}) < 1$ unless $a = b = 0$. Since A is positive definite on that nonzero space, the movement covariance is nonzero. □

3.5 Purity correction

Let the weight block be $D_i := \left(\begin{smallmatrix} w_e & w_{ef} \\ w_{ef} & w_f \end{smallmatrix}\right)$, and define its row multipliers by

$$h := 2(a w_e + b w_{ef}), \quad q := 2(a w_{ef} + b w_f), \quad p := 2(r w_e + d w_{ef}), \quad j := 2(r w_{ef} + d w_f).$$

These auxiliary letters are distinct from p_i, q_i in Eq. (10). Put

$$d_U := \frac{w_e x + w_f v + 2 w_{ef}\sqrt{xv}}{x + v}, \qquad d_V := \frac{w_e v + w_f x - 2 w_{ef}\sqrt{xv}}{x + v}, \tag{31}$$

$$R_U^* := \frac{\delta}{x + v}\left[(h - j)x + q\sqrt{xv} - \frac{p x^2}{\sqrt{xv}}\right], \qquad R_V^* := \frac{\delta}{x + v}\left[(h - j)v - q\sqrt{xv} + \frac{p v^2}{\sqrt{xv}}\right]. \tag{32}$$

Let $\mathcal{E}_0$ denote Eq. (28) evaluated with $\chi_U = \chi_V = e_U = e_V = 1$. Direct substitution of Eq. (23) through Eq. (24) proves the following identity, including the signs of all four corrections:

$$\mathcal{E}_i - \mathcal{E}_0 = R_U^*(1 - \chi_U) + R_V^*(1 - \chi_V) - d_U(1 - e_U) - d_V(1 - e_V). \tag{33}$$

The next lemma bounds this correction uniformly over the density purities, reducing the local row test to a scalar inequality.

Lemma 3.3 (A scalar sufficient test)**.** *For $d > 0$ define*

$$\Phi(R, d) := \begin{cases} 0, & R \leq d/2, \\ (2R - d)^2/(4R), & R > d/2. \end{cases}$$

For fixed weights, the inequality

$$1 - \mathcal{E}_0 \geq \Phi(R_U^*, d_U) + \Phi(R_V^*, d_V) \tag{34}$$

suffices for Eq. (28) for all local PSD density data.

Proof. The coefficients d_U, d_V are positive diagonal entries of a positive definite matrix after an orthogonal change of basis. Since $e_U \leq \sqrt{\chi_U}$, the U contribution to Eq. (33) is at most $R_U^*(1 - \chi_U) - d_U(1 - \sqrt{\chi_U})$. Set $t := \sqrt{\chi_U} \in [0, 1]$. The maximum of $R(1 - t^2) - d(1 - t)$ on this interval is exactly $\Phi(R, d)$: if $R \leq d/2$ its maximum is zero at $t = 1$; otherwise its maximum is at $t = d/(2R)$. The V contribution is identical. □

3.6 Scalar row weights

For a fixed light coordinate, define $z := |y_i|$ and $H := \iota^2\beta(y_i)^2u_iv_i/(1-z^2)$. The row weights use the following helper polynomials from Section 4.8:

$$\begin{aligned}
j_0(z,H) &:= \frac{84047268 - 52938705z + 133367380z^2 + 49643028z^3}{10^8} \\
&\quad + H\frac{22863212 + 66017181z - 147568585z^2}{10^8}, \\
b_0(z,H) &:= \frac{-41550398 + 11091757z - 14435171z^2}{10^8} \\
&\quad + H\frac{10275247 - 10715575z + 21317210z^2}{10^8}, \\
b_1(z,H) &:= \frac{3323951 + 40587513z - 130152459z^2 + 191253811z^3}{10^8} \\
&\quad + H\frac{-2355589 + 49791446z - 117532304z^2}{10^8}.
\end{aligned}$$

For a fixed coordinate, suppress its index. Set $s := 1/k$, $\ell := \Lambda Ls$, and $J := \Lambda^2\delta s/4$, and define

$$\begin{aligned}
F &:= j_0^2 + 2j_0 - J - 4Jsb_0 - 4Js(s-\delta)b_0^2, \\
B &:= 2j_0 - J - \delta(1+2j_0) - 4Jsb_0 - 4Js^2b_0^2, \\
m_0 &:= \frac{1/2 + b_0(s-\delta)}{1+j_0}, \qquad n_0 := -\frac{b_0}{1+j_0}, \\
f_* &:= F + 4Jsn_0^2(F-B) + 2F\ell n_0, \\
u_* &:= B + \delta F + 4Jsm_0n_0(B-F) - F\ell m_0, \\
a &:= f_*/u_*, \qquad \gamma_* := m_0a + n_0 + b_1, \qquad b := \Lambda sb_0, \quad r := \Lambda s\gamma_*, \quad d := j_0/\delta.
\end{aligned}$$

Use these four auxiliaries in Eq. (27). The two scalar certificates below use this construction with their specified budgets and row corrections.

3.7 A scalar certificate for the trace-variance bound

The scalar construction first uses the budget

$$\iota := \frac{29}{10}, \qquad \eta := 1 + 10^{-6}, \qquad \beta(y) := (1-y^2)^{3/4}f(|y|), \tag{35}$$

with the same exact polynomial Eq. (3). The helper polynomials j_0, b_0, b_1 in Section 3.6 are also unchanged. Use this value of ι in the potential, response scales, and local weights throughout this subsection.

At rank one the local invariants satisfy $\chi_U = \chi_V = e_U = e_V = 1$. At higher rank the exact identity Eq. (33) contains nonzero terms even at a stationary point. To prove the local comparison $\mathcal{E}_i \le 1$, retain all four purity terms and use $e_U \le \sqrt{\chi_U}$ and $e_V \le \sqrt{\chi_V}$. Lemma 3.3 then reduces the repair to

$$G := 1 - \mathcal{E}_0 - \Phi(R_U^*, d_U) - \Phi(R_V^*, d_V) \ge 0. \tag{36}$$

This test covers local density matrices of arbitrary rank.

3.8 The compact scalar domain

By exchanging U, V and reversing the sign of the signing coordinate, it suffices to take $z := |y_i| \in [0, 1)$. Define

$$\iota\beta u_i = (1+z)h, \qquad \iota\beta v_i = (1-z)k, \qquad 0 < h, k < 1. \tag{37}$$

Thus $H = hk$ and $\Delta := (1+z)h-(1-z)k$ are respectively the product parameter and the difference of the light variables. To remove the fractional-power singularity at $z = 1$, set

$$z := 1 - t^4, \quad R := 1 - z^2 = t^4(2 - t^4), \quad w := R^{1/4} = t\sqrt{\sqrt{2 - t^4}}, \qquad 0 \le t \le 1.$$

Consequently the complete domain is the cube $(t, h, k) \in [0, 1]^3$. Its boundary is used for enclosing limits; actual nonzero light inputs have $t, h, k > 0$.

Here are formulas without division by w or H for the certificate. They also specify precisely which scalar function is verified. All occurrences of f, f', f'' below are evaluated at z. Define

$$\begin{aligned}
q_c &:= Rf' - \tfrac{3}{2}zf, & d_c &:= \tfrac{3}{2}(1 - \tfrac{1}{2}z^2)f + 3zRf' - R^2f'',\\
P_0 &:= (1+z)f + q_c k, & Q_0 &:= (1-z)f - q_c h,\\
A_0 &:= hP_0/f, & B_0 &:= kQ_0/f,\\
S_0 &:= A_0 + B_0 = (1+z)h + (1-z)k, & \bar{s} &:= \frac{2\eta P_0 Q_0}{\iota d_c f^2}.
\end{aligned} \tag{38}$$

The invariants used in the helper construction become

$$\begin{aligned}
\delta = \frac{wH}{\iota f}, \qquad s = w\bar{s}, \qquad J = \frac{\eta q_c^2}{2 f d_c},\\
\ell = \frac{2\eta w}{\iota d_c}\left(\frac{q_c \Delta}{f} + \frac{2q_c^2 H}{f^2}\right).
\end{aligned} \tag{39}$$

These are exactly $s = 1/k_i$, $J = \Lambda^2\delta s/4$, and $\ell = \Lambda L s$ from Section 3.6. Use its formulas for $F, B, m_0, n_0, f_*, u_*, a, \gamma_*$. For clarity, that helper B is distinct from B_0 in Eq. (38). Put

$$\begin{aligned}
Z_* &:= 2a - (s-\delta)a^2 + 4Js\gamma_*^2,\\
D_* &:= FZ_* - 4Js(1+j_0)^2 b_1^2,\\
W_e &:= F/D_*, \qquad W_f := Z_*/D_*, \qquad W_c := -(1+j_0)b_1/D_*.
\end{aligned} \tag{40}$$

The actual weight entries in Eq. (31) are

$$w_e = W_e, \qquad w_f = \delta W_f, \qquad w_{ef} = \delta\Lambda s W_c.$$

In particular, $F = \delta Z_{22}$ and $D_* = \delta \det Z$. Checking $F, D_* > 0$ proves positive definiteness of the local weight matrix at every actual light state.

Define four rescaled row expressions and two final scalars:

$$\begin{aligned}
H_* &:= 2(aW_e + 4Jsb_0 W_c), & J_* &:= 2(4Js\gamma_* W_c + j_0 W_f),\\
Q_* &:= 2\delta(aW_c + b_0 W_f), & P_* &:= 2(\gamma_* W_e + j_0 W_c),\\
C_* &:= wHq_c\bar{s}/f, & K_* &:= \frac{2\eta w q_c P_*}{\iota d_c f}.
\end{aligned} \tag{41}$$

Direct substitution into Eq. (28) at rank-one local invariants gives the following nonsingular expression:

$$\begin{aligned}\mathcal{E}_0 = {} & W_e + \delta W_f + \delta a^2 W_e + 8aJs\delta b_0 W_c + 4Js\delta b_0^2 W_f \\ & + 4Js\gamma_*^2 W_e + 8Js\gamma_* j_0 W_c + j_0^2 W_f + \delta(-H_* + J_*) + \ell P_*. \end{aligned} \tag{42}$$

For $S_0 > 0$, set $\alpha_{\mathrm{loc}} := A_0/S_0$ and $c_* := C_*/S_0$. Then the four inputs to the purity error bound are

$$\begin{aligned} d_U &= W_e\alpha_{\mathrm{loc}} + \delta W_f(1-\alpha_{\mathrm{loc}}) + 2c_* W_c, \\ d_V &= W_e(1-\alpha_{\mathrm{loc}}) + \delta W_f\alpha_{\mathrm{loc}} - 2c_* W_c, \\ R_U^* &= \delta(H_* - J_*)\alpha_{\mathrm{loc}} + c_* Q_* - K_* A_0\alpha_{\mathrm{loc}}, \\ R_V^* &= \delta(H_* - J_*)(1-\alpha_{\mathrm{loc}}) - c_* Q_* + K_* B_0(1-\alpha_{\mathrm{loc}}). \end{aligned} \tag{43}$$

These formulas are algebraically identical to Eq. (31) through Eq. (32). For boxes touching $S_0 = 0$, use the valid enclosures

$$0 \le \alpha_{\mathrm{loc}} \le 1, \qquad 0 \le \frac{H}{S_0} \le \frac{k}{1+z}, \qquad c_* = \frac{wq_c\bar{s}}{f}\frac{H}{S_0}.$$

These enclosures cover the boundary $S_0 = 0$ and all nonzero light states. The following interval certificate verifies the scalar test on this domain.

Lemma 3.4 (Verified scalar repair at $\iota = 29/10$)**.** *On the complete cube just described the helper denominators stay positive and the formulas above certify Eq.* (36)*. In the interior, the local weight matrix is positive definite.*

Proof. The verification uses outward interval arithmetic and subdivision. Algorithm 5, with the arithmetic rules in Appendix B.3, evaluates the explicit formulas above and their first partial derivatives by the product and quotient rules. For a box Q with center x_0 it may intersect direct interval evaluation of a scalar g with

$$g(x_0) + \sum_{a=1}^{3} (\partial_a g)(Q)(Q_a - x_{0,a}).$$

Both are rigorous enclosures. This is done separately for $1 - \mathcal{E}_0, R_U^*, R_V^*, d_U, d_V$. The derivative enclosure is not used for the last four scalars on boxes where the enclosure of S_0 includes zero. Since $\Phi(R, d)$ is increasing in R and decreasing in $d \ge 0$, its upper bound uses the upper endpoint of R and $\max\{0, \underline{d}\}$ for d. The latter replacement is valid because d_U, d_V are diagonal entries of a positive definite matrix. A box is accepted only when $f, d_c, F, u_*, D_*, 1 + j_0$ have positive lower endpoints and the lower endpoint of Eq. (36) is nonnegative. Otherwise a coordinate interval is bisected.

Use the dyadic rounding, square-root, magnitude-cutoff, and binary64-conversion rules in Appendix B.3. Each stored interval contains the exact rational enclosure of the corresponding operation. The eight initial boxes are the octants of the unit cube; all later boxes are dyadic bisections of these boxes.

Algorithm 5 accepts 1,245,107 terminal boxes after visiting 2,490,206 nodes. Every octant has exact covered volume 1/8, so the covered volume is exactly 1. The largest total bisection depth from the unit cube is 26. The separate octant counts are

Octant	Nodes	Accepted boxes	Depth from octant
0	36,513	18,257	19
1	104,851	52,426	22
2	559,371	279,686	22
3	884,913	442,457	22
4	17,175	8,588	16
5	66,309	33,155	17
6	244,219	122,110	23
7	576,855	288,428	22

The positive interval lower bounds certify the inequality on the complete cube, including its boundary. The arithmetic and deterministic subdivision rules are specified in Appendix B.3.

For the budget properties used in the matrix argument, exact rational Bernstein coefficients on $[0, 1/2]$ and $[1/2, 1]$ verify $f - 1 \geq 0$, $f' \geq 0$, $-q_c \geq 0$, and $d_c - 1/2 \geq 219864451/400000000 > 0$. Consequently $1 - y^2 \leq \beta(y) \leq 1$, $-\beta'' > 1/2$, and the logarithmic derivative bound used for positive response scales holds. These univariate checks are separate from the three-variable repair. □

The scalar certificate supplies the local row inequality at every stationary light state.

Theorem 3.5 (Trace-variance signing)**.** *For positive semidefinite Hermitian matrices $A_1, \ldots, A_m$ of arbitrary ranks there are signs $\varepsilon_i \in \{-1, 1\}$, one per original matrix, such that*

$$\|\sum_{i=1}^{m} \varepsilon_i A_i\| \leq 2\sqrt{\frac{29}{10}}\|\sum_{i=1}^{m} \operatorname{tr}[A_i] A_i\|^{1/2} < 3.4059\, \sigma_{\mathrm{tr}}. \tag{44}$$

The last strict inequality applies when $\sigma_{\mathrm{tr}} > 0$; when it is zero every input is zero.

Proof. Take any stationary light state of any nonvertex face of any PSD input family. Its scalar traces determine a point of Eq. (37). Lemma 3.4, Lemma 3.3, and the exact purity identity imply $\mathcal{E}_i \leq 1$ for every active coordinate, for the actual local density matrices and their arbitrary ranks. Lemma 3.2 therefore supplies a nonzero response covariance with

$$\mathbb{E}[Q(r)] + \sum_{i \in I} \Lambda_i \mathbb{E}[Y_i F_i] \leq \sum_{i \in I} k_i \mathbb{E}[Y_i^2], \qquad k_i := \frac{\iota(-\beta_i'')}{2\eta p_i q_i}.$$

At stationarity $Y_i^2 = p_i q_i (v_i w_i + u_i z_i) r_i^2$. Combining with Eq. (14) yields

$$\tfrac{1}{2}\mathbb{E}[r^{\mathsf{T}} \nabla^2 P r] \leq -(\eta - 1) \sum_{i \in I} k_i \mathbb{E}[Y_i^2] < 0.$$

Here the second-variation formula and all preceding matrix lemmas use the new coupling $\iota = 29/10$; their proofs do not require the old numerical value. This excludes a stationary light minimum. A global minimizer of P with the fewest active coordinates therefore must be a vertex: an available endpoint otherwise produces another minimizer with fewer active coordinates. The initial trial $U := V := \ell I$ gives $P(0) \leq 2\sqrt{(29/10)\sigma_{\mathrm{tr}}^2 + 2n\rho}$. Send $\rho \downarrow 0$ and pass to a vertex occurring along a subsequence. Finally

$$(3.4059)^2 - 4(29/10) = \frac{15481}{10^8} > 0.$$

□

3.9 The main trace-variance bound

The following budget polynomial and row weights reduce the coupling to 561/200 in the same purity correction.

3.10 Budget and row weights

Use exactly the potential Eq. (5) and the trace coupling Eq. (4), now with

$$\iota := \frac{561}{200} = 2.805, \qquad \eta := 1 + 10^{-6}, \qquad \widehat{\beta}(y) := (1 - y^2)^{3/4} \widehat{f}(|y|), \tag{45}$$

where

$$\begin{aligned} \widehat{f}(z) := 1 &+ 0.19412141 z^2 + 0.08160473 z^4 + 0.00960041 z^6 + 0.08872177 z^8 \\ &- 0.11240176 z^{10} + 0.38843279 z^{12} - 0.90582086 z^{14} \\ &+ 1.32826096 z^{16} - 1.01931910 z^{18} + 0.34398483 z^{20}. \end{aligned} \tag{46}$$

Every decimal in this subsection denotes its exact terminating rational value. All scalar quantities in Eq. (38) through Eq. (43) are recomputed with $f := \widehat{f}$ and the new ι.

Let T_r be the Chebyshev polynomials specified by $T_0(X) := 1$, $T_1(X) := X$, and $T_{r+1}(X) := 2XT_r(X) - T_{r-1}(X)$. Define four exact polynomials

$$C_j(z) := 10^{-8} \sum_{r=0}^{5} N_{rj} T_r(2z - 1), \qquad 0 \le j \le 3, \tag{47}$$

using the following integer table:

r	N_{r0}	N_{r1}	N_{r2}	N_{r3}
0	-66012	3720900	-2297870	-1246900
1	2484721	-332791	1277971	324090
2	4050382	-7452932	5927687	2399040
3	2093785	195222	-3026842	709804
4	1703723	-117919	1683624	-174923
5	745971	-247570	1106145	83976

Start from the three displayed polynomials in Section 3.6 and replace them by

$$\widehat{j}_0 := j_0 + HC_0(z), \qquad \widehat{b}_0 := b_0 + HC_1(z), \qquad \widehat{b}_1 := b_1 + HC_2(z).$$

Use these new polynomials to compute F, B, m_0, n_0, f_*, u_* by the same formulas as before. Then replace the remaining row choice by

$$\widehat{a} := f_*/u_* + HC_3(z), \qquad \widehat{\gamma}_* := m_0 \widehat{a} + n_0 + \widehat{b}_1. \tag{48}$$

In all subsequent formulas use these hatted helpers. In particular, the change to a is made *before* computing γ_*, Z_*, D_* and the weights. The actual four row auxiliaries are

$$a := \widehat{a}, \qquad b := \Lambda s \widehat{b}_0, \qquad r := \Lambda s \widehat{\gamma}_*, \qquad d := \widehat{j}_0/\delta.$$

The identities $F = \delta Z_{22}$ and $D_* = \delta \det Z$ continue to hold: directly,

$$Z_{12} = \Lambda s (1 + \widehat{j}_0) \widehat{b}_1.$$

This cancellation uses only the displayed definition of $\widehat{\gamma}_*$, and does not require $a = f_*/u_*$. The next lemma verifies the scalar test for these row weights on the full compact domain.

Lemma 3.6 (Full-domain certificate for the improved parameters)**.** *For the preceding exact choices, the denominators $f, d_c, F, u_*, D_*, 1 + j_0$ are positive on the cube used in Section 3.7. The robust margin*

$$G := 1 - \mathcal{E}_0 - \Phi(R_U^*, d_U) - \Phi(R_V^*, d_V)$$

is nonnegative for every actual light state, including all arbitrary-rank local density data through Lemma 3.3.

Proof. Algorithm 5 substitutes the coefficients above into the rescaled formulas of Section 3.7. It uses the same integer-enclosed 2^{-40} interval arithmetic, positivity tests, full-cube subdivision, and treatment of $S_0 = 0$.

A further rigorous enclosure retains cancellations between the rank-one margin and the purity penalty. On the active branch of Φ, put $\theta := d/(2R) \in [0, 1]$. Direct differentiation gives

$$\partial_R \Phi = 1 - \theta^2, \qquad \partial_d \Phi = \theta - 1. \tag{49}$$

On the inactive branch both derivatives vanish, and these formulas agree at the branch boundary by taking $\theta = 1$. For interval inputs, intersect d with $[0, \infty)$; positivity of the weight matrix justifies this intersection. If $\overline{R} \le \underline{d}/2$, both derivatives are zero. If $\underline{R} > \overline{d}/2$, enclose θ by $[\underline{d}, \overline{d}]/(2[\underline{R}, \overline{R}])$. Otherwise the enclosure $[\underline{d}/(2\overline{R}), 1]$ contains all active values and also the inactive derivative pair. Intersect each such enclosure with $[0, 1]$. Every division here is outward enclosed.

The product and chain rules therefore enclose the gradient of the *complete* function G. Algorithm 5 also uses the lower endpoint of

$$G(x_0) + \sum_{a=1}^{3} (\partial_a G)(Q)(Q_a - x_{0,a})$$

in addition to the separate enclosures used earlier. Apply this mean-value bound first to interior states and then take limits; derivatives of the purity inputs are not used when the interval for S_0 contains zero. The acceptance condition is unchanged. The deterministic interval subdivision in Appendix B.3 verifies the inequality on the whole cube.

Algorithm 5 accepts 1,765,603 terminal boxes after visiting 3,531,198 nodes. Each octant has exactly covered volume 1/8, so the complete cube is covered. The maximum bisection depth from the unit cube is 37. The separate counts are

Octant	Nodes	Accepted boxes	Depth from octant
0	41,299	20,650	19
1	121,783	60,892	20
2	704,351	352,176	23
3	1,077,301	538,651	23
4	19,223	9,612	17
5	85,913	42,957	18
6	375,769	187,885	34
7	1,105,559	552,780	33

Separately, the Bernstein formula in Appendix B.2 checks rational coefficients on $[0, 1/2]$ and $[1/2, 1]$. It proves

$$\widehat{f} - 1 \ge 0, \qquad \widehat{f'} \ge 0, \qquad -q_c \ge 0, \qquad d_c - \tfrac{1}{2} \ge \frac{109577777}{200000000} > 0.$$

Since $\widehat{\beta}' = R^{-1/4}q_c$ and $-\widehat{\beta}'' = R^{-5/4}d_c$, these imply all budget conditions used in the matrix lemmas. In particular, $1 - y^2 \le \widehat{\beta}(y) \le 1$ and $-\widehat{\beta}'' > 1/2$. Moreover

$$0 \le -\frac{\widehat{\beta}'(z)}{\widehat{\beta}(z)} = \frac{3z}{2(1-z^2)} - \frac{\widehat{f}'(z)}{\widehat{f}(z)} \le \frac{3}{4(1-z)}.$$

Thus the response scales are positive throughout the light region. □

The improved certificate gives the following formal version of the trace-variance bound.

Theorem 3.7 (Formal version of Theorem 1.1)**.** *For positive semidefinite Hermitian matrices $A_1, \dots, A_m$ of arbitrary ranks, there are signs $\varepsilon_i \in \{-1, 1\}$ such that*

$$\|\sum_{i=1}^m \varepsilon_i A_i\| \le 2\sqrt{\frac{561}{200}}\|\sum_{i=1}^m \operatorname{tr}[A_i]A_i\|^{1/2} < 3.34963\,\sigma_{\mathrm{tr}}. \tag{50}$$

The strict comparison is for $\sigma_{\mathrm{tr}} > 0$; if $\sigma_{\mathrm{tr}} = 0$, all inputs are zero.

In particular, if $\sum_{i=1}^m A_i = I_n$ and $\operatorname{tr}[A_i] \le \delta$, there is a subset $S \subseteq [m]$ with

$$\|\sum_{i\in S} A_i - \tfrac{1}{2}I_n\| \le \sqrt{\frac{561}{200}}\sqrt{\delta} \le 1.674815\sqrt{\delta}.$$

Proof. Every stationary light state maps to the same compact cube as before. Lemma 3.6 and the exact purity identity Eq. (33) give $\mathcal{E}_i \le 1$ for every active coordinate, with no rank restriction. The general row-certificate lemma supplies a nonzero response covariance. The exact Hessian identity then yields

$$\tfrac{1}{2}\mathbb{E}[r^{\mathsf{T}}\nabla^2 P r] \le -(\eta - 1)\sum_{i\in I} k_i \mathbb{E}[Y_i^2] < 0.$$

The budget conditions just checked validate the regularity, lightness and endpoint arguments with the new potential parameters. As in the proof of Theorem 3.5, a global minimizer with the fewest active coordinates must therefore be a vertex. The initial potential now satisfies

$$P(0) \le 2\sqrt{\frac{561}{200}\sigma_{\mathrm{tr}}^2 + 2n\rho}.$$

Sending ρ to zero and choosing a recurrent vertex proves the non-strict estimate in Eq. (50). Finally,

$$(3.34963)^2 - 4\frac{561}{200} = \frac{211369}{10^{10}} > 0.$$

For the partition assertion, $T \preceq \delta I_n$. Taking $S := \{i : \varepsilon_i = 1\}$ gives

$$\sum_{i\in S} A_i - \tfrac{1}{2}I_n = \tfrac{1}{2}\sum_i \varepsilon_i A_i.$$

The signing bound and $\sqrt{561/200} < 1.674815$ prove the stated partition estimate. □

3.11 Shared machinery for the trace-variance algorithms

The numerical primitives for grouped signing, root evaluation, and spectral filtering are proved in Sections 5–7. We use the routines of Song–Yue [SY26a] with the trace-map interface below; the operation counts include their dense higher-rank matrix actions.

Regard Hermitian matrices as a real inner-product space with $\langle X, Y\rangle = \operatorname{tr}[XY]$; its real dimension is n^2, including for complex Hermitian inputs. For nonnegative weights γ_i, set

$$\mathcal{K}_\gamma(X) = \sum_i \gamma_i \operatorname{tr}[A_i X] A_i, \qquad \Phi e_i = \operatorname{vec}(A_i).$$

Here vec uses an orthonormal real Hermitian basis.

Lemma 3.8 (Trace-map interface)**.** *The map $\mathcal{K}_\gamma$ is self-adjoint and completely positive, and $\mathcal{K}_\gamma = \Phi \operatorname{diag}(\gamma)\Phi^*$. For $X \succ 0$ and Hermitian D,*

$$|\operatorname{tr}[A_i D]|^2 \le \operatorname{tr}[A_i X] \operatorname{tr}[A_i D X^{-1} D], \tag{51}$$

$$\mathcal{K}_\gamma(D)^2 \preceq \|\mathcal{K}_\gamma(X)\| \mathcal{K}_\gamma(D X^{-1} D). \tag{52}$$

If $|\xi_i| \le t\gamma_i$, the signed map $\mathcal{S}(X) = \sum_i \xi_i \operatorname{tr}[A_i X] A_i$ satisfies $-t\mathcal{K}_\gamma(X) \preceq \mathcal{S}(X) \preceq t\mathcal{K}_\gamma(X)$ for $X \succeq 0$; moreover,

$$\mathcal{S}(D)^2 \preceq t^2 \|\mathcal{K}_\gamma(X)\| \mathcal{K}_\gamma(D X^{-1} D). \tag{53}$$

All parameter derivatives of the trace coupling are signed maps of this form. One application of any such map or its adjoint to a dense matrix costs $O(kn^2)$ for k retained inputs.

Proof. Writing $A_i = \sum_a \lambda_{ia} u_{ia} u_{ia}^*$, the matrices $\sqrt{\gamma_i \lambda_{ia} \lambda_{ib}}\, u_{ia} u_{ib}^*$ form Kraus operators for $\mathcal{K}_\gamma$. Self-adjointness and the factorization follow directly by taking traces. Hilbert–Schmidt Cauchy–Schwarz applied to $A_i^{1/2} X^{1/2}$ and $A_i^{1/2} D X^{-1/2}$ proves Eq. (51). The block matrix

$$\begin{pmatrix} X & D \\ D & D X^{-1} D \end{pmatrix} \succeq 0$$

is positive. Complete positivity and its Schur complement give Eq. (52); one may first add a positive multiple of the identity to the upper left block and take a limit if its image is singular. For the signed assertion, apply each completely positive summand instead to

$$\begin{pmatrix} X & c_i D \\ c_i D & D X^{-1} D \end{pmatrix} \succeq 0, \qquad c_i = \xi_i/(t\gamma_i), \quad |c_i| \le 1.$$

Zero weights contribute zero and can be discarded; $t = 0$ is immediate. Summation and the same Schur-complement argument prove Eq. (53). The order bound on positive inputs follows term by term. Finally, differentiating $\gamma_i = \iota\beta(y_i)$ only changes its scalar coefficients:

$$D^j \mathcal{K}_y[h_1, \dots, h_j](X) = \iota \sum_i \beta^{(j)}(y_i) \prod_{a=1}^{j} (h_a)_i \operatorname{tr}[A_i X] A_i.$$

Each trace and each dense accumulation costs $O(n^2)$, proving the operation count without a decomposition of A_i into separate signs. □

On a retained interval, the ratios $|\beta^{(j)}|/\beta$ for $j \le 4$ are bounded by a fixed polynomial in $\log(4n)$. Thus the signed derivative bounds required for the root continuation and predictor estimates follow from the lemma. Pairing them with a positive dual block uses the exact identity $\langle Q, \mathcal{K}(X)\rangle = \langle \mathcal{K}(Q), X\rangle$. Consequently the inverse-Lyapunov primal energy, the logarithmic entropy dual energy, and their mixed forcing estimates in Sections 6.1–6.4 of the reproduced root analysis have the same form for this map. At off-root iterates we use the fixed-anchor inverse refinement from those sections; comparison of matrix eigenvalues alone would not justify replacing that inverse by a newly computed root metric. We count dense trace-map actions explicitly below.

There are two further interfaces to check separately. The grouped reduction of Lemma 5.9 acts on arbitrary vectors, so its dimension here is n^2. The Gibbs feature identity and ridge concentration of Lemma 6.5 are identities for the Gibbs covariance operator and an arbitrary real-linear forcing map F_1, so they also apply. For these features, Eq. (419) and Eq. (421) give their dense higher-rank costs. The finite-step estimates use the trace-map inequalities just proved, and are detailed below in the deterministic and curved constructions. These estimates specify the required finite numerical procedures.

3.12 Normalization and numerical primitives

Normalize $T = \sum_i \operatorname{tr}[A_i]A_i \preceq I$, reserving a fixed small allowance for an upper approximation to $\|T\|^{1/2}$. If $T = 0$, every input is zero. For scale computation, first divide by the total input trace. Then $1/(mn) \le \|T\| \le 1$, so trace powers and scalar bisection give the required relative approximation in $\widetilde{O}(mn^2 + n^3)$ operations independently of the original scale. For Hermitian D,

$$\sum_i \operatorname{tr}[A_i D]^2 \le \sum_i \operatorname{tr}[A_i]\operatorname{tr}[A_i D^2] = \operatorname{tr}[TD^2] \le \|D\|_F^2. \tag{54}$$

Hence the matrix with columns $\operatorname{vec}(A_i)$ has operator norm at most one, and each column has Frobenius norm at most one. Use the deterministic grouped reduction with dimensional slack in Lemma 5.9, in real Hermitian dimension n^2. It costs $\widetilde{O}(mn^2+n^{2\omega})$, preserves the signed sum up to a chosen fixed error, and leaves at most $2n^2$ fractional coordinates. Entrywise short approximations are needed only for this vector reduction; the maintained matrices can be the original inputs. Their accumulated vector-approximation error is included in the prescribed residual. Padding the matrices to an order $N < 4n$ gives $k \le N^2$. Below, n denotes this working order.

For the matrix stage replace the hard maximum by

$$P_s(y) = \min_{U,V\succ 0}\{\rho \operatorname{tr}[U+V] + \tau_s \log \operatorname{tr}[\exp(\frac{C(U,V,y)}{\tau_s})]\}, \tag{55}$$

where C is the block diagonal pair of left sides in Eq. (5), $\rho = c_\rho/n$, and $\tau_s = c_s/\log(4n)$. Here $c_\rho, c_s > 0$ are fixed sufficiently small constants. Then $P \le P_s \le P + \tau_s \log(2n)$. The coupling remains the trace map, with the improved budget. It is completely positive: if $A = \sum_a \lambda_a u_a u_a^*$, the operators $\sqrt{\lambda_a\lambda_b}u_a u_b^*$ are Kraus operators for $D \mapsto \operatorname{tr}[AD]A$. The Gibbs dual pair $Q = (W, Z)$ satisfies the same strict dual equations as Eq. (6). On bounded sublevels they imply

$$aI \preceq U, V \preceq C\sqrt{n}I, \qquad (a/n)I \preceq W, Z \preceq I. \tag{56}$$

Constants in this subsection are allowed to depend on the fixed certificate and fixed error allowances, but not on m, n or the ranks. Let P_n be a fixed sufficiently large polynomial in $\log(4n)$, enlarged a finite number of times as needed.

Use an endpoint cutoff $\delta_{\text{end}} = c_e/\log(4n)$ and defer coordinates nearer to an endpoint. On the retained interval all budget derivative ratios through order four are bounded by P_n. At a light

state, $u_i, v_i \le P_n$ and the response scales and their reciprocals are bounded by P_n. Arbitrary-rank normalization gives $\operatorname{tr}[A_i] \le \sqrt{n}$; lightness and Eq. (56) supply the stronger P_n bound where it is needed.

3.13 Uniform margin and curvature

Apply Algorithm 5 with the strengthened acceptance rule that the scalar margin lower bound on every accepted box is at least

$$\mu = 2^{-30}. \tag{57}$$

All eight octants pass this guard, with the same subdivision counts as before. Thus the balanced local cost is at most $1-\mu$ for *every* allowed local density quartet. This is stronger than a non-strict stationary certificate.

The rescaled weight matrix $\overline{D}_i = L_i^{-1} D_i L_i^{-1}$, with $L_i = \operatorname{diag}(1, \sqrt{\delta_i})$, and its inverse are uniformly bounded on the compact certificate domain. Indeed its entries are $W_e, W_f, \sqrt{\delta}\Lambda s W_c$, its determinant is $1/D_*$, and $(\sqrt{\delta}\Lambda s)^2 = 4Js$ is bounded. Positivity of the certified denominators and compactness give computable absolute bounds. In particular

$$D_i \preceq CI, \quad |h_*| + |j_*| \le C, \quad |q_*| \le C\sqrt{\delta}, \quad |p_*| \le C/\sqrt{\delta}. \tag{58}$$

These constants can be enclosed on the same finite interval cover. The next lemma uses their uniform bounds to extend the row certificate to states with a small gradient.

Lemma 3.9 (Stability away from stationarity)**.** *Define*

$$G = \operatorname{diag}(\varpi_i), \qquad \varpi_i = \iota(-\beta_i'')(v_i w_i + u_i z_i).$$

At a retained light state, either

$$\max_i \frac{|\partial_i P_s|}{\sqrt{\varpi_i}} \ge \frac{a}{P_n\sqrt{n}}, \tag{59}$$

or $G^{-1/2}\nabla^2 P_s G^{-1/2}$ has at least ak eigenvalues at most $-a/P_n$. The constants $a > 0$ may be decreased between inequalities.

Proof. By exchanging U, V and changing the coordinate sign, it suffices to treat $y_i \ge 0$, as in the scalar certificate. Suppress the coordinate index, put $x = cu$, $v = dv_i$, and set $\alpha_{\mathrm{loc}} = cw$, $\zeta = dz_i$. The repeated letter v in this paragraph is the rescaled trace. Let

$$D_{\mathrm{loc}} = x\alpha_{\mathrm{loc}} + v\zeta, \qquad \epsilon_i = \frac{|\alpha_{\mathrm{loc}} - \zeta|}{\alpha_{\mathrm{loc}} + \zeta}, \qquad r_b = \alpha_{\mathrm{loc}}/\zeta.$$

Now use the paired basis defined by the square roots of $x\alpha_{\mathrm{loc}}$ and $v\zeta$, rather than its balanced specialization. For $r_b \in [1/2, 2]$ the exact local response block is

$$T_{ee} = -\delta\frac{x\chi_U + v r_b \chi_V}{x r_b + v}, \qquad T_{ff} = -T_{ee},$$
$$T_{ef} = \delta\sqrt{xv}\frac{r_b^{3/2}\chi_V - r_b^{-1/2}\chi_U}{x r_b + v}, \qquad T_{fe} = \frac{\delta\sqrt{r_b}(x^2\chi_U - v^2\chi_V)}{\sqrt{xv}(x r_b + v)}.$$

The local energy block is the orthogonal change of basis of $\operatorname{diag}(e_U, e_V)$. Keep the density invariants fixed and compare these blocks with $r_b = 1$. Differentiation in $\log r_b$ bounds the changes in

T_{ee}, T_{ff}, T_{ef} by $C\delta\epsilon_i$, and the change in T_{fe} by $C\sqrt{\delta}(x+v)\epsilon_i$. The energy block changes by at most $C\epsilon_i$. Using Eq. (58) in the exact local cost gives

$$\mathcal{E}_i \le 1 - \mu + P_n\epsilon_i. \tag{60}$$

This comparison is legitimate even though the balanced density quartet need not come from a global optimizer: the scalar certificate covers every such quartet.

The scalar mixed-variation identity also holds without stationarity. For the normalized response, define $Y_i := e_i^{\mathsf{T}}(I - \mathsf{T})\xi$, $E_i := e_i^{\mathsf{T}}\xi$, $F_i := f_i^{\mathsf{T}}\xi$, and set

$$\upsilon_i = \frac{v\sqrt{\zeta} - x\sqrt{\alpha_{\mathrm{loc}}}}{\sqrt{xv}(\sqrt{\alpha_{\mathrm{loc}}} + \sqrt{\zeta})}, \qquad g_i = \beta_i'/\beta_i.$$

Directly from the coefficient response equations,

$$m_i(r) = \Lambda_i Y_i F_i + \partial_i P_s \frac{g_i}{\tau_i^2 D_{\mathrm{loc},i}} Y_i(E_i - Y_i + \upsilon_i F_i), \tag{61}$$

where $m_i(r) = \iota\beta_i' r_i(w_i\dot{\upsilon}_i + z_i\dot{u}_i)$. For example, the only scalar factorization required is

$$(x+v)\sqrt{\alpha_{\mathrm{loc}}\zeta} - (x\alpha_{\mathrm{loc}} + v\zeta) = (\alpha_{\mathrm{loc}} - \zeta)\upsilon_i\sqrt{xv}.$$

The row-completion proof applies in the actual paired basis, with $\mathcal{E}_i$ as above. It can also be restricted to any physical subspace of codimension q: if A is its weighted response covariance, $\mathrm{tr}[D^{-1}A] = k - q$. Therefore it gives

$$\mathbb{E}Q + \sum_i \Lambda_i \mathbb{E}[Y_i F_i] \le \sum_i k_i \mathbb{E}Y_i^2 - \sum_i (1 - \mathcal{E}_i) + q. \tag{62}$$

The energy block has trace at most $2k$, so coercivity and $A \preceq CI$ imply

$$\sum_i \mathbb{E}Y_i^2 \le Ck, \qquad \sum_i \Gamma_i \mathbb{E}[(E_i - Y_i)^2 + F_i^2] \le Ck.$$

Also $\delta_i(1 + \upsilon_i^2) \le C(x+v)^2$ and $|\partial_i P_s|/(\tau_i^2 D_{\mathrm{loc},i}) \le P_n\epsilon_i/(x+v)$. Cauchy–Schwarz in the last two displays bounds the sum of absolute expected remainder terms in Eq. (61) by $P_n(\max_i \epsilon_i)k$.

Moreover $Y_i^2 = \tau_i^2 D_{\mathrm{loc},i} r_i^2$ and

$$\frac{\varpi_i}{\tau_i^2 D_{\mathrm{loc},i}} = 2\eta k_i \frac{x\zeta + v\alpha_{\mathrm{loc}}}{x\alpha_{\mathrm{loc}} + v\zeta}, \qquad \left|\frac{x\zeta + v\alpha_{\mathrm{loc}}}{x\alpha_{\mathrm{loc}} + v\zeta} - 1\right| \le C\epsilon_i.$$

Choose $\max_i \epsilon_i \le a/P_n$, with a sufficiently small relative to μ and $\eta - 1$. Combining the exact Hessian comparison, Eq. (60), and Eq. (62) yields

$$\tfrac{1}{2}\,\mathrm{tr}[K_q \nabla^2 P_s] \le -\mu k/4 + q, \qquad \mathrm{tr}[GK_q] \le P_n k. \tag{63}$$

Here K_q is the physical covariance. The fixed-constraint comparison path defining the Hessian upper bound is available at a smoothed root because the same positive inverse J^{-1} exists. Its log-sum-exp term is constant, so it has exactly the earlier comparison variation.

It remains to justify this balance from a small gradient. Weighted trace Cauchy–Schwarz and the dual equations give $w_i \ge \rho u_i^2/t_i$, $z_i \ge \rho v_i^2/t_i$, and $u_i, v_i \ge a t_i$. Consequently

$$\frac{\sqrt{\varpi_i}}{\tau_i(\alpha_{\mathrm{loc}} + \zeta)} \le P_n \sqrt{\frac{u_i + v_i}{w_i + z_i}} \le P_n\sqrt{n}.$$

Failure of Eq. (59), with its constant and polylogarithm chosen accordingly, implies the required balance. If fewer than $\mu k/16$ generalized eigenvalues were below $-\mu/(16P_n)$, restrict to the inverse image of their orthogonal complement. Its codimension $q < \mu k/16$ would make the left side of Eq. (63) at least $-\mu k/32$, whereas its right side is at most $-3\mu k/16$, a contradiction. This proves the claimed multiplicity, also when the fraction is smaller than one. □

The trace-map estimates and the curvature alternative yield a deterministic signing procedure.

Theorem 3.10 (A common baseline and finishing construction)**.** *For m dense positive semidefinite Hermitian inputs of order n, a deterministic algorithm finds one sign per input with*

$$\|\sum_i \varepsilon_i A_i\| \le 3.34963 \|\sum_i \operatorname{tr}[A_i] A_i\|^{1/2}.$$

Its arithmetic-operation count is

$$\widetilde{O}(mn^2 + n^{\max\{2\omega, 11/2\}}), \tag{64}$$

where $\omega \in [2,3]$ is an admissible square-multiplication exponent. Thus ordinary matrix multiplication gives $\widetilde{O}(mn^2 + n^6)$; the current bound $\omega < 2.371177 < 11/4$ of Dupont et al. [DEK+26] gives $\widetilde{O}(mn^2 + n^{11/2})$. The hidden constants depend on the fixed scalar certificate and the fixed discrepancy allowance. The count includes comparisons and elementary scalar approximations.

Proof. Here are the higher-rank replacements needed to use the deterministic construction in Section 5.10. Small-input rounding uses $\|A_i\|_F \le t_i$ and the conditional expectation of squared Frobenius norm. At the start of a phase with K active coordinates, remove traces below $c/(L\sqrt{K})$, where $L = O(\log n)$, spending at most c/L in discrepancy. The retained coordinates obey

$$t_i \ge a/(L\sqrt{k}), \qquad \varpi_i \ge a/(P_n nk), \qquad K/2 < k \le K.$$

For deferred endpoint rounding use

$$\sum_{t_i > \tau} A_i \preceq T/\tau, \qquad \sum_i A_i^2 \preceq T \preceq I.$$

The first inequality handles heavy traces by nearest-sign rounding. For lighter traces, use the deterministic trace-exponential estimator of Lemma 5.3. Its matrix logarithmic moment term is now evaluated by functional calculus on A_i; it is not replaced by the rank-one scalar formula. The proof uses only the two inequalities just displayed and Lieb's trace-exponential concavity, in the form used in the proof of Lemma 5.3. It costs $\widetilde{O}(kn^3)$ with ordinary matrix operations.

All response-energy and root-forcing estimates transfer by Eq. (20). In particular, the higher-rank replacement for the mixed absolute majorant is

$$|\operatorname{tr}[A_i D]| \le \sqrt{\operatorname{tr}[A_i U] \operatorname{tr}[A_i D U^{-1} D]}. \tag{65}$$

Multiply the resulting scalar Young bound by the PSD output A_i and sum. The two sums are controlled respectively by $\mathcal{K}(U)$ and by the dual equation. Their operator norms and dual-weighted traces have precisely the bounds required for the linear-primal/dual path estimate Eq. (187). Complete positivity gives its root-preconditioning inequality as well:

$$\mathcal{K}(D)^2 \preceq \|\mathcal{K}(U)\| \mathcal{K}(DU^{-1}D).$$

Thus the entropy root inverse has polylogarithmic condition number in the Lyapunov energy metric. Sections 6.1–6.4 provide an explicit root evaluation with $\widetilde{O}(\sqrt{n})$ corrected stages from the scalar

zero-data root. These estimates use positivity, complete positivity, the strict dual equations, and the displayed trace Cauchy–Schwarz bound; none requires rank-one inputs.

Use the phase reservoir $\lambda_K \sum_i \beta(y_i)$ with $\lambda_K := a/(P_n n\sqrt{K})$, and define $F_s := P_s + \lambda_K \sum_i \beta(y_i)$ and $G_+ := G + \lambda_K \operatorname{diag}(-\beta_i'')$. Its total positive injection is an arbitrarily small fixed constant. The failed gradient test for F_s, measured using G, implies the hypothesis of Lemma 3.9 for P_s. Adding the reservoir strengthens Eq. (63) in Lemma 3.9 by $-\lambda_K \operatorname{tr}[K_q \operatorname{diag}(-\beta_i'')]/2$. The same subspace argument therefore supplies a fixed fraction of negative eigenvalues for the G_+-scaled Hessian, at inverse- polylogarithmic magnitude. Also $G_+ \succeq a/(P_n n\sqrt{k})I$.

The deterministic spectral-filter and conditional-expectation procedures in Section 5.10 now give a unit physical direction r with

$$r^{\mathsf{T}}\nabla^2 F_s r \le -\theta/P_n, \quad \theta = r^{\mathsf{T}} G_+ r, \quad \|r\|_\infty + \max_{\ell=1}^{4} \|\mathcal{T}_\ell r\| \le P_n d_k \sqrt{\theta}, \qquad d_k = \min\{\sqrt{n}, n/\sqrt{k}\}.$$

Here the four maps are the primal and dual optimizer responses, conjugated respectively by $U^{-1/2}, V^{-1/2}, W^{-1/2}, Z^{-1/2}$. Their energy-to-Frobenius bounds use Eq. (56); the matrix conditional estimator uses no assumption on the ranks of its Hermitian summands. In this transfer the negative-eigenvalue threshold may be inverse polylogarithmic; increasing P_n absorbs that change in the filter accuracy and step size. The linear-path entropy-residual argument and Eq. (65) justify a step $t = a/(P_n d_k \sqrt{\theta})$ after orienting r against the gradient. Enlarging P_n reserves all value-comparison errors and gives decrease at least $a/(P_n d_k^2)$. In each phase the number of curvature moves is therefore

$$I_K = \widetilde{O}(\min\{n, n^2/K\}). \tag{66}$$

For regular endpoint moves, use the following resolvent update. At a positive trigger put $d = 1 - y_i \ge \delta_{\text{end}}$ and $h = \iota\beta_i u_i + d$. Test the new face at

$$U, \qquad V_+ = (V^{-1} + hA_i)^{-1}.$$

Its minus constraint is unchanged and its plus constraint does not increase. The high-rank resolvent inequality gives

$$\operatorname{tr}[V - V_+] \ge \frac{h \operatorname{tr}[A_i V^2]}{1 + hv_i} \ge a\frac{hv_i}{1 + hv_i} \ge a\delta_{\text{end}}^2.$$

The middle inequality uses $V \succeq aI$. The last uses the trigger $v_i \ge d/(\iota\beta_i) \ge \delta_{\text{end}}/\iota$ and $h \ge d$. Conservative score tests lose only a fixed fraction of this decrease by tightening their tolerance. Thus each regular endpoint decreases P_s by at least $a\rho\delta_{\text{end}}^2$, and there are $\widetilde{O}(n)$ such endpoints. Recompute the new optimizer with the zero-data homotopy after each regular endpoint; charge $\widetilde{O}(\sqrt{n})$ stages each.

For gradient moves, define $\kappa_i := \iota(v_i w_i + u_i z_i) + \lambda_K$. Choose an inverse-polylogarithmic $h_n > 0$ small enough that $|\partial_i F_s| \le h_n\sqrt{\kappa_i/n}$ for every retained coordinate implies the small-gradient alternative of Lemma 3.9, and set $l_n := h_n/64$. The high test compares an approximation to $|\partial_i F_s|$ with $(3h_n/4)\sqrt{\kappa_i/n}$, with error at most $(h_n/8)\sqrt{\kappa_i/n}$. A sweep uses the frozen low threshold $2l_n\sqrt{\kappa_i^0/n}$ with error at most $(l_n/4)\sqrt{\kappa_i^0/n}$. These are the separated tests in Lemmas 6.23–6.24. Use the frozen-root sweeps of those lemmas and their trace-budget adaptation in Section 5.10. Their scalar surrogate is

$$\phi_i(t) = (w_i^0 - z_i^0)t + [\iota(v_i^0 w_i^0 + u_i^0 z_i^0) + \lambda_K]\beta(t).$$

The root-forcing and gradient-error estimates follow from Eq. (65), so a saturated sweep decreases the potential by $a/(P_n n)$. An unsaturated sweep exits at a separated low normalized-gradient

threshold. Marked coordinates are removed in one positive-slot attenuation batch. For a removed positive submap $\mathcal{K}_J$, set $\mathcal{K}_s = \mathcal{K}_{\mathrm{rest}} + e^{-s}\mathcal{K}_J$ and $v(s) = \langle Q, e^{-s}\mathcal{K}_J(U,V)\rangle$. The envelope and energy bounds give

$$P_s'(s) = -v(s), \quad e(P'(s)) + \mathcal{E}(Q'(s)) \le P_n\sqrt{v(s)}, \quad |v'(s)| \le P_n v(s).$$

Here e and $\mathcal{E}$ are the primal Lyapunov and dual entropy energy norms. Hence a batch of drop Δ has energy arc at most $P_n\sqrt{\Delta}$. The separated gradient thresholds imply that a batch can reactivate a gradient test only if its drop is at least $a/(P_n n)$, or if a curvature move, regular endpoint, or phase change intervenes. This is the charging argument of Section 5.10; it uses only positivity of the removed submap. There are $\widetilde{O}(n)$ sweeps and batches and, in total, $\widetilde{O}(n)$ corrected stages for batches and sweeps. All scalar sweep updates together cost $\widetilde{O}(n^4)$.

Algorithm 1 organizes these moves into halving phases. All tests use tolerances smaller than fixed fractions of the margins established above.

We next bound the number of arithmetic operations. One dense trace-coupling action, a signed sum, or a complete score scan costs $O(kn^2)$. Root evaluation uses order-n spectral preparations and matrix products, giving corrected-stage cost

$$F_{\mathrm{dense}}(n,k) = \widetilde{O}(n^3 + kn^2). \tag{67}$$

All matrix functions and inverses are approximated to the inverse-polynomial tolerances fixed above, using series, spectral preparations, or polynomially preconditioned iterations. At a fixed root, define $\Phi_i := \mathrm{vec}(A_i)$. The coupling is exactly $\Phi\,\mathrm{diag}(\iota\beta_i)\Phi^*$. Compute all k response columns together as in Lemma 5.10. Applying Φ and Φ^* to these blocks costs $\widetilde{O}(n^2k^{\omega-1})$, and the k matrix congruences cost $\widetilde{O}(kn^\omega)$. Preparation, Hessian filtering, and conditional selection therefore have total per-curvature cost

$$B(n,k) = \widetilde{O}(n^3 + kn^\omega + n^2k^{\omega-1} + k^\omega). \tag{68}$$

This block formula is valid for arbitrary-rank columns; the trace coupling is used explicitly in place of the sandwich identity.

Multiply Eq. (68) by Eq. (66) in each halving phase. For $K \le n$ the result is at most $\widetilde{O}(n^4 + n^{2+\omega})$. For $n < K \le n^2$ it is at most

$$\widetilde{O}(n^5/K + n^{2+\omega} + n^4K^{\omega-2} + n^2K^{\omega-1}).$$

Summing phases costs $\widetilde{O}(n^{2\omega})$. The cold roots after curvature moves cost, per phase,

$$I_K\sqrt{n}\,(n^3 + Kn^2) = \widetilde{O}(n^{9/2}).$$

The $\widetilde{O}(n)$ regular endpoints each cost at most $\widetilde{O}(\sqrt{n}(n^3+n^4))$ for a fresh root, totaling $\widetilde{O}(n^{11/2})$. Gradient sweeps and attenuation batches cost $\widetilde{O}(nF_{\mathrm{dense}}(n,n^2)) = \widetilde{O}(n^5)$. Small-input rounding and final deferred rounding cost at most $\widetilde{O}(n^5)$. Adding the grouped preprocessing proves Eq. (64).

It remains to verify the discrepancy bound. Before adding the phase reservoirs, the normalized initial value is at most

$$2\sqrt{561/200 + 2n\rho} + \tau_s\log(2n).$$

Choose their fixed constants, partial-signing error, small-input rounding, deferred rounding, normalization error and accepted comparison tolerances so that their total effect is at most 10^{-6}. The total number of events is polynomial, so inverse-polynomial per-comparison tolerances suffice. The preceding bounds retain the same powers of n when these fixed constants decrease. Exactly,

$$(3.34963 - 10^{-6})^2 - 4(561/200) = \frac{14437641}{10^{12}} > 0.$$

Consequently the final coefficient is below 3.34963. This proves the discrepancy bound and the operation count. □

Algorithm 1 records the procedure proved in Theorem 3.10. Its numerical tests use the accuracy and separated thresholds of the frozen-root sweeps in Lemmas 6.23–6.24, with the normalization and error allowances specified above.

Algorithm 1 Deterministic trace-variance signing

procedure TraceVarianceSigning($A_1, \dots, A_m$)
  Form $T \leftarrow \sum_i \operatorname{tr}[A_i] A_i$.
  if $T = 0$ **then**
    $\varepsilon \leftarrow (1, \dots, 1)$.
    return ε.
  end if
  Normalize and apply grouped partial signing. ▷ Lemma 5.9
  Initialize the modeled sum, improved budget, and rounding ledger. ▷ Eq. (45)
  while active coordinates remain **do**
    if a new halving phase begins **then**
      Round the small-trace inputs and set the phase reservoir.
      Compute a certified root of the smoothed potential. ▷ Eq. (55)
    end if
    Defer near-endpoint coordinates in an attenuation batch.
    if a regular endpoint move is available **then**
      Perform the endpoint move and reinitialize the root.
    else
      Perform frozen-root gradient sweeps. ▷ Lemmas 6.23–6.24
      Attenuate the marked slots.
      if the state is retained and light and the gradient tests fail **then**
        Assemble the scaled Hessian and response columns.
        Select a spread direction by filtering and conditional expectation. ▷ Lemma 3.9
        Take the certified decreasing step and reinitialize the root.
      end if
    end if
  end while
  Jointly round the deferred coordinates by conditional expectation. ▷ Lemma 5.3
  Recover all recorded signs in ε and restore the input scale.
  return ε.
end procedure

3.14 Fixed-background estimates

At a corrected state, classify the coordinates in the retained interval $[-1+\delta_{\mathrm{end}}, 1-\delta_{\mathrm{end}}]$. Coordinates outside this interval may be removed in a deferred attenuation batch or kept fixed as pending slots. The non-light count below concerns only retained coordinates. Define the two endpoint scores

$$s_i^+ = \iota\beta_i v_i - (1 - y_i), \qquad s_i^- = \iota\beta_i u_i - (1 + y_i).$$

Use enclosures of width at most $\delta_{\mathrm{end}}/64$ and put in B_{nl} every coordinate whose upper score enclosure is at least $-\delta_{\mathrm{end}}/8$. Thus for retained $i \notin B_{\mathrm{nl}}$ both true scores are less than $-\delta_{\mathrm{end}}/8$. A marked

coordinate has $u_i \geq \delta_{\mathrm{end}}/(2\iota)$ or $v_i \geq \delta_{\mathrm{end}}/(2\iota)$, after allowing the stated numerical slack. The contraction Eq. (54) and the strict dual equations imply

$$\sum_i (u_i^2 + v_i^2) \leq \operatorname{tr}[U^2 + V^2] \leq 1/\rho.$$

Consequently, at every classification,

$$|B_{\mathrm{nl}}| \leq H_n := \lceil \frac{16\iota^2}{\rho\delta_{\mathrm{end}}^2} \rceil = \widetilde{O}(n). \tag{69}$$

Fix

$$K_c = \lceil (1 + 512/\mu) H_n \rceil, \qquad \mu = 2^{-30}.$$

Let B contain B_{nl} and any additional fixed pending slots, and define $J := I \setminus B$, $p := |J|$, and $q := |B|$. The calculation below applies whenever J is retained and light and $q \leq \mu p/512$. If there are no pending slots, taking $B = B_{\mathrm{nl}}$ and $k > K_c$ guarantees this condition. Both large-count constructions use a larger switching threshold and explicitly bound the pending count to guarantee the same condition. The coordinates in B retain their signed coefficients, coupling slots, and reservoir terms. They are neither rounded nor deleted. Only the coordinates in J move.

The response constraints below preserve curvature in the presence of fixed coordinates. At a fixed root decompose the trace coupling as $\mathcal{K} = \mathcal{K}_J + \mathcal{K}_B$. For a pair (D_U, D_V), let

$$\mathcal{L}_J(D_U, D_V) = (U^{-1} D_U U^{-1} - \mathcal{K}_J(D_V), V^{-1} D_V V^{-1} - \mathcal{K}_J(D_U)).$$

Its inverse is positive: the strict dual inequalities for the full coupling remain strict for its positive submap $\mathcal{K}_J$. For a physical direction $h \in \mathbb{R}^J$ let $D_J(h) = \mathcal{L}_J^{-1} F_1 h$ be its fixed-constraint response. Impose the $2q$ real linear equations

$$\operatorname{tr}[A_b D_J(h)_U] = \operatorname{tr}[A_b D_J(h)_V] = 0 \quad (b \in B). \tag{70}$$

Each original input contributes only two equations, regardless of its matrix rank. On this subspace $\mathcal{K}_B(D_J(h)_U) = \mathcal{K}_B(D_J(h)_V) = 0$. Hence $D_J(h)$ is exactly the full-coupling response. The omitted coupling also contributes zero to its quadratic inversion energy. This is the trace-map version of the fixed-background construction in Lemma 6.28. Pair the second-response fixed-constraint equation with the full dual optimizer. The identity $\mathcal{L}Q = \rho(I, I)$ then gives the original full Hessian energy and mixed terms.

Repeat Lemma 3.9 using only J and add Eq. (70) to any tested physical subspace of codimension r. The local density inequalities still hold: the unused background adds positive terms to the dual coercivity bounds. On the annihilator the full and reduced inversion energies agree. Thus the restricted row calculation gives

$$\tfrac{1}{2} \operatorname{tr}[K_r \nabla_{JJ}^2 P_s] \leq -\mu p/4 + r + 2q, \qquad \operatorname{tr}[G_J K_r] \leq P_n p. \tag{71}$$

Here K_r is supported on the tested subspace and the annihilator; it is used in the proof, not constructed by the algorithm. The same small-gradient balance on J suffices; no gradient assumption is made on B. If fewer than $\mu p/32$ scaled Hessian eigenvalues were below $-\mu/(32P_n)$, their complementary subspace would make the left side at least $-\mu p/64$. The right side would be at most $-\mu p/4 + \mu p/32 + \mu p/256$, a contradiction. Adding the concave reservoir preserves this argument as before.

3.15 Fast evaluation of deferred rounding

The common deferred conditional-expectation rounding can also be evaluated in $\widetilde{O}(kn^{\omega})$, also for high-rank inputs. For each coordinate compute its two exact logarithmic moment matrices as functions of A_i, to the prescribed tolerance. On the light-trace part their exponential arguments have norm $O(\log n)$ and the positive logarithm arguments have known polynomial spectral bounds. The polynomial/resolvent matrix-function algorithms of Lemmas 6.34–6.35 thus cost $\widetilde{O}(n^{\omega})$ per coordinate; the conditional trace-exponential comparisons have the same cost. Heavy traces retain nearest-sign rounding. Matrix sums are updated in $O(n^2)$ per coordinate. Lieb's estimator inequality and $\sum A_i^2 \preceq T$ are unchanged. In particular this step uses full matrix functional calculus rather than the rank-one scalar moment formula.

3.16 Small-count finishing

Once the active set has size $\widetilde{O}(n)$, the baseline construction completes the signing within the remaining discrepancy allowance.

Lemma 3.11 (Finishing a bounded-sublevel retained state)**.** *Suppose the current retained state lies on the bounded sublevel used above, has $k \le K_c = \widetilde{O}(n)$ active coordinates, and carries the existing phase reservoir and rounding ledger. The baseline endpoint, sweep, and curvature procedure finishes the signing in*

$$\widetilde{O}(n^{9/2} + n^{\omega+2})$$

arithmetic operations, including final rounding of at most $O(n^2)$ deferred coordinates. It uses the existing discrepancy allowance.

Proof. The proof of Theorem 3.10 applies to every retained face on the bounded sublevel, not only to the initial zero signing. Keep the current reservoir phase and the existing rounding ledger. The procedure resumes from the current corrected state; it does not restart the input normalization or spend a fresh error allowance.

There are $\widetilde{O}(n)$ regular endpoints. A corrected stage now costs

$$F_{\mathrm{dense}}(n,k) = \widetilde{O}(n^3 + kn^2) = \widetilde{O}(n^3).$$

Thus cold endpoint initializations cost $\widetilde{O}(n^{9/2})$. The curvature count in Eq. (66) in Theorem 3.10 and the block cost in Eq. (68) in Theorem 3.10, summed over the remaining halving phases, give $\widetilde{O}(n^4 + n^{\omega+2})$ block work and $\widetilde{O}(n^{9/2})$ cold-root work. The sweep and attenuation accounting gives $\widetilde{O}(n)$ corrected stages and hence $\widetilde{O}(n^4)$ further work.

Use Section 3.15 for the final ledger. Its $O(n^2)$ deferred coordinates cost $\widetilde{O}(n^{\omega+2})$. All accepted moves and rounding steps have the same discrepancy accounting as in the baseline construction. These bounds prove the assertion, for either preceding large-count algorithm. □

3.17 Coordinate acceleration and negative eigenvalues

The coordinate acceleration below applies to either large-count algorithm.

Work first on a light movable set J, with the full background kept in the root. Define $g_i := \partial_i F_s$, $g_i^0 := \partial_i P_s$, $R_i := \lambda_K(-\beta_i'')$, and

$$G_+ = \mathrm{diag}(\varpi_i + R_i), \qquad \varpi_i = \iota(-\beta_i'')(v_i w_i + u_i z_i).$$

Uniform light-state bounds, without a gradient assumption, give a known $C_0 \le P_n$ such that

$$\nabla^2_{JJ} P_s \succeq -C_0 G_J, \qquad (\nabla^2 P_s)_{ii} \le C_0 \varpi_i. \tag{72}$$

The Schur formula is $\nabla^2 P_s = -G + F_1^* \mathcal{Q} F_1 - (\mathcal{L}\mathcal{Q}F_1 - F_2)^* \mathcal{S}^{-1}(\mathcal{L}\mathcal{Q}F_1 - F_2)$. The weighted F_2 estimate and completion of its square give the lower bound. For the upper diagonal, use $\mathrm{tr}[A_i^2 W] \le t_i w_i$, $\mathrm{tr}[A_i^2 Z] \le t_i z_i$, $u_i, v_i \ge a t_i$, and the light bounds on p_i, q_i. They give $(F_1^* \mathcal{Q} F_1)_{ii} \le P_n t_i (w_i + z_i) \le P_n \varpi_i$. All operators in this identity use the full coupling.

Let $\epsilon_i = |c_i w_i - d_i z_i|/(c_i w_i + d_i z_i)$ as in Lemma 3.9. Choose a fixed inverse-polylogarithmic ϵ_0 so that $\epsilon_i \le 2\epsilon_0$ suffices for that lemma's restricted row estimate. Call such a coordinate balanced; otherwise put it in L, the set to be accelerated. Numerical classification uses an overlap: every unaccelerated coordinate has $\epsilon_i \le 2\epsilon_0$ and every accelerated one has $\epsilon_i \ge \epsilon_0/2$. This is obtainable from sufficiently accurate scalar enclosures without an exact equality test.

On L, $|g_i^0| \ge a\epsilon_0(w_i + z_i)$, with a polylogarithmic change of constants. The small-input cutoff and dual equations give $w_i + z_i \ge a t_i/n$ and $t_i \ge a/(P_n\sqrt{K})$. The constant in $\lambda_K = a/(P_n n\sqrt{K})$ can therefore be chosen so that $|\lambda_K \beta_i'| \le |g_i^0|/4$ on L. In particular g_i has the same sign as g_i^0 there. Define the following new acceleration coefficients (distinct from d_i in Eq. (10)):

$$a_i = \frac{(C_0+2)\varpi_i}{g_i} \quad (i \in L), \qquad a_i = 0 \quad (i \notin L), \qquad H^\sharp = (\nabla^2 F_s)_{JJ} - 2\,\mathrm{diag}(a_i g_i). \tag{73}$$

These divisions have inverse-polynomial lower bounds. Crucially,

$$|a_i| \le P_n(u_i + v_i) \le P_n, \qquad |a_i|(w_i + z_i) \le P_n \varpi_i, \qquad |a_i g_i| \le P_n \varpi_i. \tag{74}$$

The numerator in Eq. (73) contains only ϖ_i; the reservoir already contributes a negative Hessian.

We next bound the number of negative eigenvalues of the accelerated Hessian. Put $p = |J|$ and let q count all fixed resident coordinates. Require $q \le \mu p/4096$, which will be enforced below. The matrix $A^\sharp = G_+^{-1/2} H^\sharp G_+^{-1/2}$ has lower bound $-P_n I$, diagonal entries bounded in absolute value by P_n, and hence upper norm at most $P_n p$. If $|L| \ge \mu p/4096$, its principal submatrix on L has diagonal entries at most -1: by Eq. (72),

$$(H^\sharp)_{ii} \le C_0 \varpi_i - R_i - 2(C_0+2)\varpi_i \le -(\varpi_i + R_i).$$

A trace argument with its lower bound $-P_n I$ gives at least $|L|/(2P_n)$ eigenvalues below $-1/2$, after enlarging P_n. Interlacing transfers them to $A^\sharp$.

If $|L| < \mu p/4096$, temporarily treat L as additional background *only in the covariance proof*. On $J \setminus L$ the acceleration is zero and the local balance condition holds. Eq. (71), with the additional $2|L|$ annihilator equations, and the same complementary-subspace argument give a fixed inverse-polylogarithmic fraction of negative eigenvalues. The original background plus L has size at most $\mu p/2048$, well within that argument's slack. Thus in both cases there are known $\pi_n, \delta_n > 0$, with $\pi_n^{-1}, \delta_n^{-1} \le P_n$, such that

$$\#\{\lambda(A^\sharp) \le -4\delta_n\} \ge \pi_n p. \tag{75}$$

The two cases give the same spectral guarantee for the algorithm. The scalar certificate is applied only to balanced coordinates.

3.18 A direct comparison for accelerated paths

The linear primal–dual comparison in Eq. (189) extends to accelerated coordinate paths. The additional forcing requires both an operator-norm bound and a dual-weighted trace bound, which we verify explicitly.

Lemma 3.12 (Accelerated finite-step comparison)**.** *Define $F := F_s$, $X := (U, V)$, and $Q := (W, Z)$, with pairs identified with block diagonal matrices, so $\operatorname{tr}[Q] = 1$. Let $C(X, y)$ be the block constraint matrix in Eq. (55), and put $\tau = \tau_s$. Let r be supported on the retained light set J, put $\theta = r^{\mathsf{T}} G_+ r$, and let $A(r)$ and $B(r)$ be the true first primal and dual optimizer responses at the base root. Suppose*

$$e(A(r))^2 + \mathcal{E}(B(r))^2 \le P_n \theta.$$

Let R be $\|r\|_\infty$ plus the maximum of the four Hermitian relative response norms, obtained by conjugating the two primal responses by $U^{-1/2}, V^{-1/2}$ and the two dual responses by $W^{-1/2}, Z^{-1/2}$. Freeze the coefficients in Eq. (73), satisfying Eq. (74), and take

$$y_i(t) = y_i + t r_i - t^2 a_i r_i^2 \quad (i \in J), \qquad y_i(t) = y_i \quad (i \notin J).$$

If $|t| R \le a / P_n$, then, after enlarging P_n,

$$F(y(t)) \le F(y) + t\, g^T r + \frac{1}{2} t^2 r^T H^\sharp r + P_n R \theta |t|^3 + P_n R^2 \theta t^4. \tag{76}$$

The bound uses linear comparison paths for the primal and dual matrices.

Proof. Use the linear matrix paths $X_t = X + tA(r)$, $Q_t = Q + tB(r)$. They remain positive and comparable to their base values and $\operatorname{tr}[Q_t] = 1$. Also $y(t)$ stays in the enlarged retained interval. Set

$$S(t) = \rho \operatorname{tr}[X_t] + \langle Q_t, C(X_t, y(t)) \rangle - \tau \operatorname{tr}[Q_t \log Q_t] + \lambda_K \sum_i \beta(y_i(t)).$$

Stationarity and its first derivative give

$$S(0) = F(y), \qquad S'(0) = g^T r, \qquad S''(0) = r^T \nabla^2 F r - 2 \sum_i g_i a_i r_i^2 = r^T H^\sharp r.$$

The acceleration term must be included in this last equality.

We check all new terms relative to the straight-path comparison. Put $h_i(t) = y_i(t)$. Then

$$h_i' = r_i - 2 t a_i r_i^2, \quad h_i'' = -2 a_i r_i^2, \quad h_i''' = 0, \quad |h_i'| \le P_n |r_i|.$$

For either primal block, differentiation of the coupling gives exactly

$$\frac{d^2}{dt^2}[\mathcal{K}_{h(t)}(X_t)] = \mathcal{K}''_{h(t)}(X_t) + 2\mathcal{K}'_{h(t)}(A),$$
$$\frac{d^3}{dt^3}[\mathcal{K}_{h(t)}(X_t)] = \mathcal{K}'''_{h(t)}(X_t) + 3\mathcal{K}''_{h(t)}(A),$$

where the scalar coefficients in these three operator derivatives are

$$\beta'(h_i) h_i', \quad \beta''(h_i)(h_i')^2 + \beta'(h_i) h_i'', \quad \beta'''(h_i)(h_i')^3 + 3\beta''(h_i) h_i' h_i''.$$

In addition $M'' = -2 \sum_i a_i r_i^2 A_i$ and $M''' = 0$. These formulas list every new acceleration term.

For the new signed-matrix forcing use its PSD absolute majorant $J_M = 2 \sum_i |a_i| r_i^2 A_i$. By Eq. (74), and the lower bound $\beta_i \ge P_n^{-1}$ at movable base coordinates,

$$\|J_M\| \le P_n \|r\|_\infty^2 \Big\| \sum_{i \in J} (u_i + v_i) A_i \Big\| \le P_n R^2,$$

$$\operatorname{tr}[(W+Z)J_M] \le P_n \sum_{i\in J} \varpi_i r_i^2 \le P_n\theta.$$

The same bounds hold against Q_t by comparability. The new $\beta' h_i''$ coupling term has majorant proportional to $\sum_i |\beta'(h_i)a_i| r_i^2 \operatorname{tr}[A_i X_t] A_i$. Its norm is at most $P_n R^2$ using $|a_i| \le P_n$ and the base coupling bound. Its paired dual trace is at most $P_n\theta$ using budget derivative ratios and X_t, Q_t comparability. The remaining second derivative terms are the old inverse quadratic, the $\beta''(h_i)(h_i')^2$ term, and the mixed term. The last is bounded by the trace Cauchy–Schwarz inequality

$$|\operatorname{tr}[A_i A_X]|^2 \le \operatorname{tr}[A_i X_t] \operatorname{tr}[A_i A_X X_t^{-1} A_X]$$

and scalar Young's inequality. Pairing its positive second summand with the dual uses the full strict dual equations. This gives respectively norm $P_n R^2$ and weighted trace $P_n\theta$, just as for the straight path. Only $h_i' = O(P_n r_i)$ has changed.

Let c_0, c_1 be the scalar Gibbs multipliers at the root and its first derivative. The Hermitian Gibbs residual

$$E(t) = C(X_t, y(t)) - \tau \log Q_t - (c_0 + c_1 t) I$$

has $E(0) = E'(0) = 0$. The logarithm term has its usual PSD second derivative majorant; its norm is bounded by $P_n R^2$ and its weighted trace by $P_n \mathcal{E}(B)^2$. Integrating the preceding majorants twice gives a PSD matrix $J(t)$ with

$$-J(t) \preceq E(t) \preceq J(t), \qquad \|J(t)\| \le P_n R^2 t^2, \qquad \operatorname{tr}[Q_t J(t)] \le P_n \theta t^2. \tag{77}$$

Base and path dual matrices are comparable, so the same final Q_t can be used when integrating majorants from intermediate parameters.

The third derivative of S is

$$S'''(t) = \langle Q_t, C''' \rangle + 3\langle B, C'' \rangle - \tau \frac{d^3}{dt^3} \operatorname{tr}[Q_t \log Q_t] + \lambda_K \sum_i [\beta'''(h_i)(h_i')^3 + 3\beta''(h_i) h_i' h_i''].$$

The new $3\beta'' h_i' h_i''$ term has paired trace at most $P_n \|r\|_\infty \theta$. For $3\mathcal{K}''(A)$, use

$$|\operatorname{tr}[A_i A_X]| \le P_n R \operatorname{tr}[A_i X_t].$$

The new acceleration terms again cost at most $P_n R\theta$. The M'' term paired with B has this bound since $-P_n R Q_t \preceq B \preceq P_n R Q_t$. Inverse cubic and entropy cubic terms are bounded by $P_n R$ times their response energies. The reservoir terms are bounded by $P_n \|r\|_\infty \lambda_K \sum_i (-\beta_i'') r_i^2$. Thus $|S'''(t)| \le P_n R\theta$.

Finally, for $Q_t \succ 0$ of trace one and $\|E\| \le \tau/2$, Golden–Thompson and $e^x \le 1 + x + x^2$ for $|x| \le 1/2$ imply

$$\tau \log \operatorname{tr}[\exp(\log Q_t + E/\tau)] - \operatorname{tr}[Q_t E] \le \tau^{-1} \operatorname{tr}[Q_t E^2].$$

If $-J \preceq E \preceq J$, then $E^2 \preceq \|J\| J$ by factoring $E = J^{1/2} T J^{1/2}$ with $\|T\| \le 1$. Consequently Eq. (77) bounds this gap by $P_n R^2 \theta t^4$. Evaluating the minimized primal objective at X_t, then applying the Taylor bound for $S(t)$, proves Eq. (76). The restriction $|t| R \le a/P_n$ also makes $\|E(t)\| \le \tau/2$ after enlarging P_n. □

3.19 Resident coordinates and bulk deletion

Choose $\zeta = \mu/2^{20}$, decreasing it by a fixed factor if needed, and increase the switching threshold to $K_c = \widetilde{O}(n)$ so $H_n \le \zeta k$ whenever $k > K_c$. Near-endpoint coordinates are kept resident and fixed, including their signed contributions, positive coupling slots, and reservoir terms. Delete them together by positive-submap attenuation only when their count reaches $\lceil \zeta k \rceil$. Before this trigger, combine them with the non-light set from Eq. (69). Its size is at most $2\zeta k$, so $q \le \mu p/4096$. Only the remaining light coordinates enter the accelerated path. Derivatives of the pending coordinates are zero; no derivative bound at their near-endpoint values is used.

A triggered deletion removes a fixed fraction of resident coordinates. Including initial cleanup, phase small-input processing, and final cleanup gives $\widetilde{O}(1)$ attenuation calls. Their drops sum to a constant, and their charged stage bound is $\widetilde{O}(1+\sqrt{n\Delta_j})$ each. Consequently all bulk continuations use $\widetilde{O}(\sqrt{n})$ corrected stages. Classification itself changes neither the root nor the potential, and costs one score scan per signing invocation. This rule introduces no gradient sweeps. Each large-count construction below charges its signing moves to its own proved decrease; the shared deletion rule adds only the corrected stages counted above. Once $k \le K_c$, clear the pending set and use the common small-count finishing procedure in Lemma 3.11. No fresh rounding or reservoir allowance is allocated at this switch.

4 Potential identities and supporting scalar certificates

Sections 4.1 and 4.2 introduce the reference potential and its variation formulas. The projection and covariance estimates in Sections 4.3–4.5 give strict descent. Section 4.6 verifies the scalar inequalities, and Sections 4.7 and 4.8 incorporate the trace correction and certify the improved budget.

The reference potential, variation formulas, covariance estimates, and scalar certificates below support the numerical tools in Sections 5 and 6. The transfer to the higher-rank trace map is established in Lemma 3.8; the higher-rank scalar analysis is in Section 3.9.

4.1 Normalization, budget, and potential

Discard zero inputs. Define $L := \sum_i |\operatorname{tr}[A_i]|$ and $\bar{A}_i := \operatorname{sign}(\operatorname{tr}[A_i])A_i/L$. Then $\bar{A}_i \succeq 0$ has rank one, $\sum_i \operatorname{tr}[\bar{A}_i] = 1$, and, writing $\zeta := \|\sum_i \bar{A}_i^2\|$,

$$(mn)^{-1} \le \zeta \le 1, \qquad \|\bar{A}_i\| \le \sqrt{\zeta}.$$

A signing for the $\bar{A}_i$ transfers to one for the A_i. For clarity we first analyze nonzero PSD inputs with a variance upper bound $\bar{\zeta}$. The identities also hold on a face with a fixed signed matrix contribution. We use the variance upper bound $\bar{\zeta} := \zeta$ for the existence proof.

Until Section 4.7, use the following budget, denoted also by β_{fast} when both algorithms are discussed. Let $r_{\mathrm{aux}}(z) := 1 - z^2$ and $f := f_{\mathrm{fast}}$, where f_{fast} is the explicit polynomial in Eq. (122). Define

$$\beta(y) := (1-y^2)^p f(|y|), \qquad -1 \le y \le 1.$$

Define

$$p := \frac{97}{200}, \qquad \iota := \frac{14779}{2500} = 5.9116, \qquad \tau := \frac{4902302619}{5000000000}, \qquad \vartheta_* := \frac{101}{200}.$$

The following properties are certified in Section 4.6:

$$\beta(0) = 1, \quad \beta(\pm 1) = 0, \quad \sqrt{1-y^2} \le \beta(y) \le 1, \tag{78}$$

$$f'(z) \geq 0, \qquad \beta'(z) \leq 0, \qquad -\beta''(z) > 1/2, \qquad 0 \leq z < 1, \tag{79}$$

$$\frac{(-\beta'(z))(1+z)}{\beta(z)(-\beta''(z))} \leq 1. \tag{80}$$

The even extension β is C^2 and has locally Lipschitz second derivative on $(-1, 1)$. Its third derivative may jump at zero; no third derivative at zero is needed.

For a fractional signing y, define $M(y) := \sum_i y_i \bar{A}_i$ and $\mathcal{K}_y(D) := \iota \sum_i \beta(y_i) \bar{A}_i D \bar{A}_i$. For $\rho > 0$ define

$$\begin{aligned} \mathcal{P}(y) := \min_{b,\ U,V \succ 0} \ & b + \rho \operatorname{tr}[U + V] \\ \text{subject to } & U^{-1} + M(y) + \mathcal{K}_y(V) \preceq bI, \\ & V^{-1} - M(y) + \mathcal{K}_y(U) \preceq bI. \end{aligned} \tag{81}$$

Our total potential is $\mathcal{F} := \mathcal{P} + \gamma \mathcal{B}$, where $\mathcal{B}(y) := \sum_i \beta(y_i)$ and $\gamma > 0$. For every feasible point the constraints imply $b > \|M(y)\|$. At $y = 0$, the scalar choice $U := \ell I$, $V := \ell I$, and $b := \ell^{-1} + \iota \bar{\zeta} \ell$ is feasible. Minimizing its objective over $\ell > 0$ gives

$$\|M(y)\| \leq \mathcal{P}(y), \qquad \mathcal{F}(0) \leq 2\sqrt{\iota \bar{\zeta} + 2n\rho} + \gamma m. \tag{82}$$

The trace penalty makes the optimizer unique. The PSD dual multipliers W, Z are positive definite and satisfy

$$\operatorname{tr}[W + Z] = 1, \qquad W = U\mathcal{K}_y(Z)U + \rho U^2, \qquad Z = V\mathcal{K}_y(W)V + \rho V^2, \tag{83}$$

$$U^{-1} + M + \mathcal{K}_y(V) = bI, \qquad V^{-1} - M + \mathcal{K}_y(U) = bI. \tag{84}$$

The next lemma justifies differentiating the optimized potential twice within each open face of the cube.

Lemma 4.1. *On each open face of the cube, the program in Eq. (81) has a unique optimizer. Its primal and dual variables, and $\mathcal{P}$, are C^2 with locally Lipschitz second derivatives. The optimizer satisfies Eqs. (83) and (84). The value $\mathcal{P}$ is continuous on the closed cube.*

Proof. The map $\mathcal{K}_y$ is positive and self-adjoint. Matrix inversion is operator convex, and $U := I$ and $V := I$ with large b is strictly feasible. On any bounded objective sublevel, the constraints bound U^{-1}, V^{-1} above, and the trace penalty bounds U, V above. The sublevel is therefore compact inside the positive definite cone, so a minimizer exists. Convex duality gives $W, Z \succeq 0$. Differentiating the Lagrangian in b, U, V gives Eq. (83). In particular $W \succeq \rho U^2 \succ 0$ and $Z \succeq \rho V^2 \succ 0$, so complementary slackness gives Eq. (84).

For a nonzero Hermitian A,

$$\frac{d^2}{dt^2} \operatorname{tr}[W(U + t\mathsf{A})^{-1}]|_{t=0} = 2 \operatorname{tr}[WU^{-1}\mathsf{A}U^{-1}\mathsf{A}U^{-1}] > 0.$$

Thus the Lagrangian is strictly convex in (U, V); it has only one minimizer, and the tight equations determine b.

For completeness, define the self-adjoint linearization

$$\mathscr{J}(\mathsf{A}, \mathsf{B}) := (U^{-1}\mathsf{A}U^{-1} - \mathcal{K}_y(\mathsf{B}), V^{-1}\mathsf{B}V^{-1} - \mathcal{K}_y(\mathsf{A})).$$

The positive map $\mathcal{N}(\mathsf{A}, \mathsf{B}) := (U\mathcal{K}_y(\mathsf{B})U, V\mathcal{K}_y(\mathsf{A})V)$ satisfies $\mathcal{N}(W, Z) = (W, Z) - \rho(U^2, V^2) \preceq r_0(W, Z)$ for some $r_0 < 1$. Every Hermitian pair is bounded in absolute semidefinite order by a

multiple of (W, Z), so the Neumann series for $I - \mathcal{N}$ converges. Hence $\mathscr{J}$ is invertible and $\mathscr{J}^{-1}$ preserves positivity. This also proves uniqueness of the dual pair.

Let $\mathscr{E}$ be the positive definite Hessian of the Lagrangian in (U, V). A vector $(d_{\mathrm{b}}, \mathsf{A}, \mathsf{B}, W', Z')$ in the kernel of the full KKT Jacobian satisfies

$$\mathscr{J}(\mathsf{A}, \mathsf{B}) = -d_{\mathrm{b}}(I, I), \qquad \mathscr{J}(W', Z') = \mathscr{E}(\mathsf{A}, \mathsf{B}), \qquad \operatorname{tr}[W' + Z'] = 0.$$

Pairing the middle equation with (A, B) shows $\langle (\mathsf{A}, \mathsf{B}), \mathscr{E}(\mathsf{A}, \mathsf{B}) \rangle = 0$. Thus $\mathsf{A} = \mathsf{B} = 0$, then $d_{\mathrm{b}} = 0$ and $W' = Z' = 0$. The implicit function theorem proves the asserted regularity; the locally Lipschitz second derivatives follow from those of β. Finally the sublevel bounds can be chosen uniformly in y for fixed $\rho > 0$. A convergent sequence of optimizers gives lower semicontinuity at every boundary point. Conversely, keeping U, V fixed and increasing b by a vanishing amount restores feasibility after a small change in y, proving upper semicontinuity. □

4.2 Endpoint moves and variation formulas

Fix a state and let its k_{act} active coordinates be indexed by $[k_{\mathrm{act}}]$. Write $\bar{A}_i = a_i a_i^*$ and

$$u_i := a_i^* U a_i, \quad v_i := a_i^* V a_i, \quad w_i := a_i^* W a_i, \quad z_i := a_i^* Z a_i.$$

If $\iota\beta_i v_i \geq 1 - y_i$, rounding y_i to 1 preserves feasibility with the same (b, U, V) and decreases $\mathcal{F}$ by at least $\gamma\beta_i$. Likewise $\iota\beta_i u_i \geq 1 + y_i$ permits rounding to -1. Hence the remaining case is

$$\iota\beta_i u_i < 1 + y_i, \qquad \iota\beta_i v_i < 1 - y_i. \tag{85}$$

Define

$$p_i := 1 + \iota\beta_i' v_i, \quad q_i := 1 - \iota\beta_i' u_i, \quad c_i := \sqrt{\beta_i p_i / q_i}, \quad d_i := \sqrt{\beta_i q_i / p_i}, \quad \tau_i := \sqrt{p_i q_i / \beta_i}.$$

For $y_i \geq 0$, $f' \geq 0$ gives

$$-\beta_i'/\beta_i \leq \frac{y_i}{1 - y_i^2}, \qquad p_i \geq 1/(1 + y_i) \geq 1/2, \quad q_i \geq 1, \quad p_i q_i \leq 1 - (1 + y_i)\beta_i'/\beta_i.$$

For negative y_i, exchange p_i, q_i and u_i, v_i. In particular all scales above are positive. Write

$$\widetilde{u}_i := c_i u_i, \quad \widetilde{v}_i := d_i v_i, \qquad \widetilde{w}_i := c_i w_i, \quad \widetilde{z}_i := d_i z_i.$$

We will use $c_i d_i = \beta_i$, $\tau_i c_i = p_i$, $\tau_i d_i = q_i$, and

$$\widetilde{u}_i \widetilde{v}_i = \beta_i u_i v_i < \frac{1 - y_i^2}{\iota^2 \beta_i} \leq \frac{1}{\iota^2}. \tag{86}$$

For a real direction r, hold b fixed and differentiate the two tight constraints. The resulting feasible local path has derivatives satisfying

$$\begin{aligned} U^{-1}\dot{U}U^{-1} &= \sum_i p_i r_i \bar{A}_i + \mathcal{K}_y(\dot{V}), \\ V^{-1}\dot{V}V^{-1} &= -\sum_i q_i r_i \bar{A}_i + \mathcal{K}_y(\dot{U}). \end{aligned} \tag{87}$$

Define the nonnegative quadratic form

$$\mathcal{Q}(r) := \operatorname{tr}[W U^{-1}\dot{U}U^{-1}\dot{U}U^{-1}] + \operatorname{tr}[Z V^{-1}\dot{V}V^{-1}\dot{V}V^{-1}].$$

Differentiating the fixed-b feasible path gives

$$\partial_i \mathcal{P} = p_i w_i - q_i z_i = \tau_i(\widetilde{w}_i - \widetilde{z}_i). \tag{88}$$

More explicitly, its second-variation calculation gives

$$r^\top \nabla^2 \mathcal{P} r \le 2\mathcal{Q}(r) + 2\sqrt{\mathcal{Q}(r)\mathcal{M}(r)} - \iota \sum_i (-\beta_i'')(v_i w_i + u_i z_i) r_i^2,$$

where $\mathcal{M}(r) := \iota \sum_i (\beta_i'^2/\beta_i)(v_i w_i + u_i z_i) r_i^2$. Here is a direct derivation of the second-variation bound. Let $\varphi_r(t) := b + \rho \operatorname{tr}[U(t) + V(t)]$ along the fixed-b feasible path. Then $\varphi_r(t) \ge \mathcal{P}(y + tr)$ with equality at zero. Also $\mathscr{J}(W, Z) = \rho(I, I)$, so differentiating the constraints once gives $\varphi_r'(0) = \sum_i (p_i w_i - q_i z_i) r_i$. Differentiating twice gives

$$\varphi_r''(0) = 2\mathcal{Q}(r) + 2\iota \sum_i \beta_i' r_i (w_i a_i^* \dot{V} a_i + z_i a_i^* \dot{U} a_i) + \iota \sum_i \beta_i'' (v_i w_i + u_i z_i) r_i^2.$$

By the dual equations,

$$\mathcal{Q}(r) \ge \iota \sum_i \beta_i (w_i a_i^* \dot{V} V^{-1} \dot{V} a_i + z_i a_i^* \dot{U} U^{-1} \dot{U} a_i).$$

The inequalities $|a_i^* \dot{U} a_i|^2 \le u_i a_i^* \dot{U} U^{-1} \dot{U} a_i$ and its V analogue, followed by scalar Cauchy–Schwarz, bound the absolute mixed term by $2\sqrt{\mathcal{Q}(r)\mathcal{M}(r)}$. This proves the stated bound because $r^\top \nabla^2 \mathcal{P} r \le \varphi_r''(0)$.

Apply $2\sqrt{EB} \le \tau E + B/\tau$. For any centered random r, with $\pi_i := \tau_i^2 \mathbb{E}[r_i^2]$, this gives

$$\frac{1}{2}\mathbb{E}[r^\top \nabla^2 \mathcal{P} r] \le (1 + \tau/2)\mathbb{E}[\mathcal{Q}(r)] - \frac{\iota}{2} \sum_i (-\beta_i'' - \frac{\beta_i'^2}{\tau \beta_i})(v_i w_i + u_i z_i)\mathbb{E}[r_i^2]. \tag{89}$$

4.3 A weighted projection inequality

The covariance must retain a different coupling coefficient at each coordinate. The next lemma permits these coefficients without requiring that the response projection commute with their weights.

Lemma 4.2. *Let $e_1, \dots, e_{k_{\mathrm{act}}}, f_1, \dots, f_{k_{\mathrm{act}}}$ be an orthonormal basis of $\mathbb{R}^{2k_{\mathrm{act}}}$. Let Π_0 project onto the span of the e_i, and let Π be a rank-k_{act} orthogonal projection. Suppose*

$$(I - \Pi_0)(I - T)\Pi = 0, \qquad \langle e_i, T e_i\rangle + \langle f_i, T f_i\rangle = 0, \qquad \langle e_i, T e_i\rangle \le 0,$$

$$\langle e_i, D_{\mathrm{aux}} e_i\rangle = \langle f_i, D_{\mathrm{aux}} f_i\rangle = 1, \qquad D_{\mathrm{aux}} \succeq T^\top \Gamma T,$$

where $\Gamma e_i = \Gamma_i e_i$, $\Gamma f_i = \Gamma_i f_i$, and $\Gamma_i > 2$. Define $0 < \vartheta_i < 1$ and k_i by

$$\vartheta_i := \frac{2}{\sqrt{1 + 4\Gamma_i} - 1}, \qquad k_i := \frac{1 + \vartheta_i}{1 - \vartheta_i^3}.$$

There is a PSD matrix A, positive definite on $\operatorname{range}\Pi$ and zero on its orthogonal complement, such that

$$\operatorname{tr}[D_{\mathrm{aux}}\mathsf{A}] \le \sum_i k_i \langle e_i, (I - T)\mathsf{A}(I - T)^\top e_i\rangle. \tag{90}$$

Moreover $\operatorname{tr}[(I - T)\mathsf{A}(I - T)^\top] > 0$.

Proof. For $0 < \vartheta < 1$ define

$$A := 1 + \vartheta^2 + \vartheta^3, \quad w_e := \frac{1}{(1+\vartheta)A}, \quad w_f := \frac{\vartheta}{(1+\vartheta)(2+\vartheta)}, \quad h_{\text{aux}} := \frac{2}{A}, \quad j := \frac{2}{2+\vartheta}.$$

With $\Gamma := (1+\vartheta)/\vartheta^2$ and $k := (1+\vartheta)/(1-\vartheta^3)$, direct rational identities give

$$\frac{h_{\text{aux}}}{w_e} - \frac{{h_{\text{aux}}}^2}{4w_e^2}(1/k - 1/\Gamma) = 1/w_e, \tag{91}$$

$$\frac{j}{w_f} + \frac{j^2}{4\Gamma w_f^2} = 1/w_f, \tag{92}$$

$$\frac{{h_{\text{aux}}}^2}{4\Gamma w_e} + \frac{j^2}{4\Gamma w_f} = 1 - w_e - w_f. \tag{93}$$

Also

$$h_{\text{aux}} - j = \frac{2(1-\vartheta)(1+\vartheta)^2}{(2+\vartheta)A} > 0.$$

Use these values at ϑ_i for each coordinate and define

$$\mathsf{D} := \sum_i (w_i^e e_i e_i^\top + w_i^f f_i f_i^\top), \qquad \mathsf{A} := \mathsf{E}(\mathsf{E}^\top \mathsf{D}^{-1} \mathsf{E})^{-1} \mathsf{E}^\top, \qquad \mathsf{B} := \mathsf{D} - \mathsf{A},$$

where E is an orthonormal basis matrix for range Π. Since $\mathsf{D}^{-1/2}\mathsf{A}\mathsf{D}^{-1/2}$ is an orthogonal projection of rank k_{act},

$$\mathsf{A}, \mathsf{B} \succeq 0, \quad \mathsf{A}\mathsf{D}^{-1}\mathsf{A} = \mathsf{A}, \quad \mathsf{A}\mathsf{D}^{-1}\mathsf{B} = 0, \quad \operatorname{tr}[\mathsf{D}^{-1}\mathsf{A}] = k_{\text{act}}.$$

The energy outside A satisfies

$$\operatorname{tr}[D_{\text{aux}}\mathsf{A}] \le \operatorname{tr}[D_{\text{aux}}\mathsf{D}] - \operatorname{tr}[\Gamma T \mathsf{B} T^\top], \qquad \operatorname{tr}[D_{\text{aux}}\mathsf{D}] = \sum_i (w_i^e + w_i^f).$$

We compare this gap and the movement by completing squares in each row.

For the e_i row define $t_i := e_i^\top T$ and $C_i := \Gamma_i \mathsf{B} + k_i \mathsf{A}$. This matrix is positive definite, with

$$C_i^{-1} = \Gamma_i^{-1}\mathsf{D}^{-1} + (k_i^{-1} - \Gamma_i^{-1})\mathsf{D}^{-1}\mathsf{A}\mathsf{D}^{-1}.$$

Completing a square gives

$$\begin{aligned} &\Gamma_i t_i \mathsf{B} t_i^\top + k_i (e_i^\top - t_i)\mathsf{A}(e_i - t_i^\top) + h_{\text{aux}i} t_i e_i \\ &\ge k_i \langle e_i, \mathsf{A} e_i \rangle - (k_i \mathsf{A} e_i - h_{\text{aux}i} e_i/2)^\top C_i^{-1} (k_i \mathsf{A} e_i - h_{\text{aux}i} e_i/2) \\ &= \frac{\langle e_i, \mathsf{A} e_i \rangle}{w_i^e} - \frac{{h_{\text{aux}i}}^2}{4\Gamma_i w_i^e}. \end{aligned}$$

The last equality uses Eq. (91). For the f_i row define $s_i := f_i^\top T$. The response condition gives $s_i \mathsf{A} = f_i^\top \mathsf{A}$, so

$$s_i f_i = \frac{\langle f_i, \mathsf{A} f_i \rangle + s_i \mathsf{B} f_i}{w_i^f}.$$

Completing a square with the PSD matrix B, which need not be invertible, and using Eq. (92), we obtain

$$\Gamma_i s_i \mathsf{B} s_i^\top + j_i s_i f_i \ge \frac{j_i}{w_i^f}\langle f_i, \mathsf{A} f_i \rangle - \frac{j_i^2}{4\Gamma_i (w_i^f)^2}\langle f_i, \mathsf{B} f_i \rangle = \frac{\langle f_i, \mathsf{A} f_i \rangle}{w_i^f} - \frac{j_i^2}{4\Gamma_i w_i^f}.$$

Let

$$\mathfrak{q} := \operatorname{tr}[\Gamma T \mathsf{B} T^\top] + \sum_i k_i \langle e_i, (I-T)\mathsf{A}(I-T)^\top e_i\rangle.$$

Summing the row inequalities and Eq. (93) yields

$$\mathfrak{q} + \sum_i (h_{\mathrm{aux}\,i}\langle e_i, Te_i\rangle + j_i\langle f_i, Tf_i\rangle) \geq \operatorname{tr}[\mathsf{D}^{-1}\mathsf{A}] - \sum_i (1 - w_i^e - w_i^f) = \operatorname{tr}[D_{\mathrm{aux}}\mathsf{D}].$$

The added sum equals $\sum_i (h_{\mathrm{aux}\,i} - j_i)\langle e_i, Te_i\rangle \leq 0$. Thus $\mathfrak{q} \geq \operatorname{tr}[D_{\mathrm{aux}}\mathsf{D}]$, proving Eq. (90).

Finally define $\Gamma_{\min} := \min_i \Gamma_i > 2$ and $w_{\min} := \lambda_{\min}(\mathsf{D}) > 0$. Taking traces in the hypothesis gives $\|T\|_F^2 \leq 2k_{\mathrm{act}}/\Gamma_{\min} < k_{\mathrm{act}}$, and $\mathsf{A} \succeq w_{\min}\Pi$. Consequently

$$\operatorname{tr}[(I-T)\mathsf{A}(I-T)^\top] \geq w_{\min} k_{\mathrm{act}} (1 - \sqrt{2/\Gamma_{\min}})^2 > 0. \tag{94}$$

No dimension-independent lower bound on $w_{\min}$ is needed. □

4.4 A covariance with coordinate-dependent coefficients

We apply the weighted projection inequality to the response system to obtain a covariance for coordinate moves.

Lemma 4.3. *Under Eq. (85), define*

$$\Gamma_i := \frac{1}{\iota \widetilde{u}_i \widetilde{v}_i} = \frac{1}{\iota \beta_i u_i v_i}, \qquad \vartheta_i := \frac{2}{\sqrt{1 + 4\Gamma_i} - 1}, \qquad k_i := \frac{1 + \vartheta_i}{1 - \vartheta_i^3}.$$

There is a centered, finitely supported direction r with nonzero covariance such that, with $\pi_i := \tau_i^2 \mathbb{E}[r_i^2]$,

$$\mathbb{E}[\mathcal{Q}(r)] \leq \sum_i k_i (\widetilde{u}_i \widetilde{w}_i + \widetilde{v}_i \widetilde{z}_i)\pi_i. \tag{95}$$

Proof. Define symmetric PSD, entrywise nonnegative matrices

$$\mathscr{U}_{ij} := \iota c_i c_j |a_i^* U a_j|^2, \qquad \mathscr{V}_{ij} := \iota d_i d_j |a_i^* V a_j|^2.$$

For $P, Q \in \mathbb{R}^{k_{\mathrm{act}}}$ let $M_P := \sum_i c_i P_i \bar{A}_i$ and $M_Q := \sum_i d_i Q_i \bar{A}_i$. The responses $\dot{U} := U M_P U$, $\dot{V} := V M_Q V$ satisfy Eq. (87) precisely when

$$J := P - \mathscr{V} Q = \mathscr{U} P - Q, \qquad r_i := J_i/\tau_i. \tag{96}$$

The linearized constraint map is invertible, so these are the actual responses. Define

$$E_P := \operatorname{diag}(\widetilde{u}_i \widetilde{w}_i), \quad E_Q := \operatorname{diag}(\widetilde{v}_i \widetilde{z}_i), \quad E := \operatorname{diag}(E_P, E_Q).$$

The energy $\operatorname{tr}[W M_P U M_P] + \operatorname{tr}[Z M_Q V M_Q]$ is a PSD quadratic form with matrix M_{raw} and diagonal E. Hence $\widehat{D_{\mathrm{aux}}} := E^{-1/2} M_{\mathrm{raw}} E^{-1/2}$ has unit diagonal and zero cross-block entries. Define

$$\widehat{T} := \begin{pmatrix} 0 & E_P^{1/2} \mathscr{V} E_Q^{-1/2} \\ E_Q^{1/2} \mathscr{U} E_P^{-1/2} & 0 \end{pmatrix}, \qquad \Gamma := \operatorname{diag}(\Gamma_1, \dots, \Gamma_{k_{\mathrm{act}}}, \Gamma_1, \dots, \Gamma_{k_{\mathrm{act}}}).$$

The full coefficient at each coordinate gives

$$\widehat{D_{\mathrm{aux}}} \succeq \widehat{T}^\top \Gamma \widehat{T}, \qquad \Gamma_i > \frac{\iota \beta_i}{1 - y_i^2} \geq \iota > 2. \tag{97}$$

Indeed the dual equation and Cauchy–Schwarz imply, for every real P,

$$\operatorname{tr}[WM_PUM_P] \geq \iota \sum_i \beta_i z_i a_i^* UM_PUM_PUa_i \geq \sum_i \frac{\widetilde{z}_i}{\iota \widetilde{u}_i} (\mathscr{U}P)_i^2 = \sum_i \Gamma_i \widetilde{v}_i \widetilde{z}_i (\mathscr{U}P)_i^2.$$

Here $|a_i^* UM_PUa_i|^2 \leq u_i a_i^* UM_PUM_PUa_i$ and $(\mathscr{U}P)_i = \iota c_i a_i^* UM_PUa_i$. The analogous inequality for Q, followed by diagonal normalization, proves the matrix inequality for all (P, Q). The light bounds give $\iota^2 \beta_i^2 u_i v_i < 1 - y_i^2$; also $\beta_i \geq 1 - y_i^2$. These prove the remaining assertions. The argument applies to complex Hermitian input matrices.

Let

$$\mathcal{V}_{\mathrm{resp}} := \{(P, Q) : (I - \mathscr{U})P + (I - \mathscr{V})Q = 0\}.$$

This space has dimension k_{act}. In fact the strict dual equations give $\mathscr{U}\widetilde{z} < \widetilde{w}$ and $\mathscr{V}\widetilde{w} < \widetilde{z}$ coordinatewise, so $\rho(\mathscr{U}\mathscr{V}) < 1$. A nonzero left kernel vector of $[I - \mathscr{U} \;\; I - \mathscr{V}]$ would be a common eigenvector with eigenvalue one, a contradiction. Let Π project onto $E^{1/2}\mathcal{V}_{\mathrm{resp}}$. Writing $d_{P,i} := (E_P)_{ii}$ and $d_{Q,i} := (E_Q)_{ii}$, define

$$e_i := \frac{\sqrt{d_{P,i}} e_{P,i} - \sqrt{d_{Q,i}} e_{Q,i}}{\sqrt{d_{P,i} + d_{Q,i}}}, \qquad f_i := \frac{\sqrt{d_{Q,i}} e_{P,i} + \sqrt{d_{P,i}} e_{Q,i}}{\sqrt{d_{P,i} + d_{Q,i}}}.$$

Let Π_0 project onto the span of the e_i. Eq. (96) gives $(I - \Pi_0)(I - \widehat{T})\Pi = 0$. The local two-coordinate blocks of $\widehat{D_{\mathrm{aux}}}$ are identity matrices, so its quadratic forms at e_i, f_i equal one. Directly,

$$\langle e_i, \widehat{T} e_i \rangle = -\frac{d_{P,i}\mathscr{V}_{ii} + d_{Q,i}\mathscr{U}_{ii}}{d_{P,i} + d_{Q,i}} \leq 0, \qquad \langle f_i, \widehat{T} f_i \rangle = -\langle e_i, \widehat{T} e_i \rangle.$$

The matrix Γ is scalar on each coordinate pair, so this change of basis preserves its form. Apply Lemma 4.2. For independent uniform signs $\epsilon \in \{\pm 1\}^{2k_{\mathrm{act}}}$ define $(P, Q) := E^{-1/2}\mathsf{A}^{1/2}\epsilon$ and define J, r by Eq. (96). This pair belongs to $\mathcal{V}_{\mathrm{resp}}$, and

$$\mathbb{E}[\mathcal{Q}(r)] = \operatorname{tr}[\widehat{D_{\mathrm{aux}}}\mathsf{A}], \qquad \langle e_i, (I - \widehat{T})\mathsf{A}(I - \widehat{T})^\top e_i \rangle = (d_{P,i} + d_{Q,i})\mathbb{E}[J_i^2].$$

Thus Eq. (90) proves the desired comparison, and Eq. (94) implies $\operatorname{tr}[\mathbb{E}[rr^\top]] > 0$. The signs here supply only a finite-support existence certificate. □

4.5 Joint curvature and strict descent

The coupling coefficient and endpoint position must be bounded together. This replaces the use of a single lower bound for the scalar curvature ratio.

Lemma 4.4. *At every light state,*

$$\iota(-\beta_i'' - \frac{\beta_i'^2}{\tau\beta_i}) \geq (1 + \kappa_{\mathrm{fast}})(2 + \tau)k_i p_i q_i, \qquad \kappa_{\mathrm{fast}} := 10^{-8}. \tag{98}$$

Proof. By symmetry suppose $y_i \geq 0$ and define $z := y_i$. Write

$$q_{\log} := -\beta'/\beta \geq 0, \qquad u := \iota\beta_i u_i, \qquad v := \iota\beta_i v_i.$$

The light inequalities give $0 < u < 1 + z$, $0 < v < 1 - z$, and

$$pq = (1 - q_{\log} v)(1 + q_{\log} u), \qquad \Gamma = \frac{\iota\beta}{uv}.$$

Here $1 - q_{\log} v > 0$. The function $k(\Gamma)$ decreases in Γ, since $(1+\vartheta)/\vartheta^2$ decreases in ϑ whereas $(1+\vartheta)/(1-\vartheta^3)$ increases. For fixed v, the product $k(\Gamma)(1 - q_{\log} v)(1 + q_{\log} u)$ therefore increases with u. Define $u := 1 + z$ for an upper bound, define $v_{\mathrm{loc}} := v/(1-z)$, and define ϑ by

$$\vartheta := \frac{2}{\sqrt{1 + 4\iota\beta/((1-z^2)v_{\mathrm{loc}})} - 1}.$$

The case $v_{\mathrm{loc}} = 0$ is its continuous limit $\vartheta = 0$. The constraint $0 \le v_{\mathrm{loc}} \le 1$ becomes

$$\iota f(z)\vartheta^2 \le (1+\vartheta)(1-z^2)^{1-p}. \tag{99}$$

All such parameters satisfy $0 \le \vartheta < \vartheta_*$, because $f \ge 1$, $(1-z^2)^{1-p} \le 1$, and $\iota\vartheta_*^2 > 1 + \vartheta_*$. A cancellation retains the negative endpoint term:

$$k_i p_i q_i \le (1 + (1+z)q_{\log}) \frac{1 + \vartheta - \iota(-\beta')\vartheta^2/(1+z)}{1 - \vartheta^3}. \tag{100}$$

Indeed substitute $v_{\mathrm{loc}} = \iota\beta\vartheta^2/((1-z^2)(1+\vartheta))$ into $(1 - q_{\log} v)k$.

Define

$$r_{\mathrm{aux}} := 1 - z^2, \quad q := r_{\mathrm{aux}} f' - 2pzf, \quad v_0 := r_{\mathrm{aux}} q' + 2(1-p)zq,$$
$$d_0 := r_{\mathrm{aux}} f - (1+z)q, \quad w_0 := -\tau f v_0 - q^2.$$

Then $\beta' = {r_{\mathrm{aux}}}^{p-1} q$ and $\beta'' = {r_{\mathrm{aux}}}^{p-2} v_0$. After multiplying Eq. (100) by the positive denominators, Eq. (98) follows from $\mathcal{A} \ge {r_{\mathrm{aux}}}^{1-p} G_{\mathrm{cert}}$, where

$$\mathcal{A} := \iota[(1+z)w_0(1-\vartheta^3) - (1+\kappa_{\mathrm{fast}})(2+\tau)\tau d_0 q\vartheta^2],$$
$$G_{\mathrm{cert}} := (1+\kappa_{\mathrm{fast}})(2+\tau)\tau d_0(1+z)(1+\vartheta). \tag{101}$$

The scalar certificate below proves $w_0, d_0 > 0$ and $q \le 0$, so $\mathcal{A}, G_{\mathrm{cert}} > 0$ on $[0,1] \times [0, \vartheta_*]$. It also proves the full-domain implication

$$H_{\mathrm{cert}} := \mathcal{A} - {r_{\mathrm{aux}}}^{1-p} G_{\mathrm{cert}} \ge 0 \quad \text{whenever} \quad \mathcal{G} := \iota f\vartheta^2 - {r_{\mathrm{aux}}}^{1-p}(1+\vartheta) \le 0. \tag{102}$$

The constraint is exactly Eq. (99). All multipliers used above are positive for $z < 1$, so this proves Eq. (98). □

Define $C_{2,i} := (1 + \tau/2)k_i$ and define

$$h_i := -\frac{C_{2,i}}{\tau_i}(\widetilde{u}_i - \widetilde{v}_i)\pi_i, \qquad K := \mathbb{E}[rr^\top].$$

Define $\mathscr{L}g := \nabla g \cdot h + \frac{1}{2}\operatorname{tr}[\nabla^2 g K]$. The cancellation is coordinatewise:

$$(\widetilde{u}_i\widetilde{w}_i + \widetilde{v}_i\widetilde{z}_i) - (\widetilde{u}_i\widetilde{z}_i + \widetilde{v}_i\widetilde{w}_i) = (\widetilde{u}_i - \widetilde{v}_i)(\widetilde{w}_i - \widetilde{z}_i).$$

Also

$$(\widetilde{u}_i\widetilde{z}_i + \widetilde{v}_i\widetilde{w}_i)\pi_i = p_i q_i (u_i z_i + v_i w_i)\mathbb{E}[r_i^2].$$

Combining Eq. (88), Eq. (89) and the two preceding lemmas gives

$$\mathscr{L}\mathcal{P} \le \sum_i [C_{2,i} p_i q_i - \frac{\iota}{2}(-\beta_i'' - \frac{\beta_i'^2}{\tau\beta_i})](u_i z_i + v_i w_i)\mathbb{E}[r_i^2] \le 0. \tag{103}$$

No derivative of k_i or h_i is taken: the drift and covariance are pointwise witnesses for this generator inequality.

To bound the budget drift suppose $y_i \geq 0$. Since

$$\tau_i(\widetilde{u}_i - \widetilde{v}_i) = u_i - v_i + 2\iota\beta_i' u_i v_i \leq u_i,$$

Eqs. (85) and (80) give

$$\beta_i' h_i \leq \frac{C_{2,i}}{\iota} \frac{(-\beta_i')(1+y_i)}{\beta_i} \mathbb{E}[r_i^2] \leq \frac{C_{2,i}}{\iota}(-\beta_i'')\mathbb{E}[r_i^2].$$

Negative coordinates follow by symmetry. The exact uniform bound is

$$\frac{C_{2,i}}{\iota} \leq \frac{(1+\tau/2)(1+\vartheta_*)}{\iota(1-\vartheta_*^3)} = \frac{4485593088319}{10300518152100} < \frac{9}{20}. \tag{104}$$

Together with $-\beta'' > 1/2$, this proves

$$\mathscr{L}\mathcal{B} \leq -\frac{1}{20}\sum_i (-\beta_i'')\mathbb{E}[r_i^2] \leq -\frac{1}{40}\operatorname{tr}[K] < 0. \tag{105}$$

The covariance is nonzero by Eq. (94). Thus the total potential has strict local descent whenever an endpoint move is unavailable.

The same cancellation retains the strict surplus:

$$\mathscr{L}\mathcal{P} \leq -\kappa_{\mathrm{fast}} \sum_i C_{2,i} p_i q_i (v_i w_i + u_i z_i) K_{ii}. \tag{106}$$

This additional margin will be used in the faster algorithm. The earlier budget estimate also gives $\mathscr{L}\mathcal{B} \leq -\frac{1}{20}\sum_i (-\beta_i'')K_{ii}$. Minimizing $\mathcal{F}$ on the cube now proves the 4.8628 discrepancy bound: an endpoint move strictly decreases $\mathcal{F}$, and a light nonvertex minimum contradicts its negative generator. Letting ρ, γ tend to zero in Eq. (82) gives $2\sqrt{\iota} < 4.8628$.

4.6 Exact scalar certificate

The scalar inequalities follow from finite rational coefficient tests. With $r_{\mathrm{aux}}, q, v_0, d_0, w_0$ defined above, these tests give

$$f - 1 \geq 0, \quad f' \geq 0, \quad -q \geq 0, \quad -v_0 - 1/2 > 0, \quad d_0 > 0, \quad w_0 > 0,$$
$$(-f v_0)^2 - (1+z)^2 q^2 r_{\mathrm{aux}} > 0$$

on $[0,1]$. Each inequality has a rational Bernstein certificate on $[0,1/2]$ and $[1/2,1]$; the maximum degree is 84. The derivative identities give Eq. (79). Since $p \leq 1/2$, we have ${r_{\mathrm{aux}}}^{1-p} \leq \sqrt{r_{\mathrm{aux}}}$. Taking nonnegative square roots in the last inequality therefore gives

$$\frac{(-\beta')(1+z)}{\beta(-\beta'')} = {r_{\mathrm{aux}}}^{1-p} \frac{-(1+z)q}{-f v_0} \leq 1.$$

Also $f \geq 1$, monotonicity of β, and $\beta(0) = 1$ imply $\sqrt{r_{\mathrm{aux}}} \leq {r_{\mathrm{aux}}}^p \leq \beta \leq 1$. The missing linear term in f gives a C^2 even extension with locally Lipschitz second derivative.

We describe the certificate for Eq. (102) without replacing a fractional power by a high-degree polynomial. Define $p_{\mathrm{cert}} := 1 - p = 103/200$. On each dyadic rectangle in $[0,1] \times [0, \vartheta_*]$, let z_0 be

the midpoint of its z interval and define $r_0 := 1 - z_0^2 > 0$. Define $u_0 := k/2^{80}$, where k is the least positive integer satisfying $u_0^{200} \geq r_0^{103}$. This is an exact integer comparison. Concavity of $t^{p_{\mathrm{cert}}}$ on $[0,1]$ gives the polynomial upper bound

$$r_{\mathrm{aux}}{}^{p_{\mathrm{cert}}} \leq L(z) := u_0(1 - p_{\mathrm{cert}} + p_{\mathrm{cert}} \frac{r_{\mathrm{aux}}(z)}{r_0}).$$

Indeed the tangent with $r_0^{p_{\mathrm{cert}}}$ in place of u_0 is an upper bound, and its multiplier $1 - p_{\mathrm{cert}} + p_{\mathrm{cert}} r_{\mathrm{aux}}/r_0$ is positive. Define two polynomials on this rectangle by

$$F_L := \mathcal{A} - L G_{\mathrm{cert}}, \qquad G_L := \iota f \vartheta^2 - L(1 + \vartheta).$$

Their common Bernstein bidegree is $(43, 3)$. Each rectangle passes one of three tests: all coefficients of F_L are nonnegative; all coefficients of G_L are strictly positive; or all coefficients of $F_L + \lambda G_L$ are nonnegative for some rational $\lambda \geq 0$. These tests suffice. Since $G_{\mathrm{cert}} > 0$, the first gives $H_{\mathrm{cert}} \geq F_L \geq 0$. The second excludes every feasible point, because $\mathcal{G} \geq G_L$. For the third, feasibility gives $G_L \leq \mathcal{G} \leq 0$, and hence $H_{\mathrm{cert}} \geq F_L \geq -\lambda G_L \geq 0$.

We give the coefficient rules used here. For a polynomial P on a box $\prod_{j=1}^{d} [a_j, b_j]$, substitute $x_j = a_j + (b_j - a_j)u_j$ and expand $P(x) = \sum_\nu A_\nu u^\nu$. If n_j bounds its degree in u_j, its tensor-product Bernstein coefficients are

$$B_\mu := \sum_{\nu \leq \mu} A_\nu \prod_{j=1}^{d} \frac{\binom{\mu_j}{\nu_j}}{\binom{n_j}{\nu_j}}, \qquad 0 \leq \mu_j \leq n_j. \tag{107}$$

The corresponding basis functions are nonnegative and sum to one. Thus P lies between the smallest and largest coefficients on the box. For midpoint subdivision in one coordinate, apply

$$c_i^{(0)} := B_i, \qquad c_i^{(r+1)} := \frac{c_i^{(r)} + c_{i+1}^{(r)}}{2} \quad (0 \leq i < n - r). \tag{108}$$

The left and right coefficient lists are $(c_0^{(r)})_{r=0}^{n}$ and $(c_i^{(n-i)})_{i=0}^{n}$, respectively. Apply this operation to every coefficient line parallel to the selected coordinate. All operations are rational.

For the joint inequality, use degree $(43, 3)$ and scale ϑ by ϑ_*. Initialize coefficient arrays for the six polynomials

$$\mathcal{A}, \quad G_{\mathrm{cert}}, \quad r_{\mathrm{aux}} G_{\mathrm{cert}}, \quad \iota f \vartheta^2, \quad 1 + \vartheta, \quad r_{\mathrm{aux}}(1 + \vartheta).$$

Keep a common positive scale through subdivision. This preserves the relative coefficients when forming F_L and G_L. If their coefficient pairs are (a_μ, b_μ), define

$$\lambda_- := \max(\{0\} \cup \{-a_\mu/b_\mu : b_\mu > 0\}), \qquad \lambda_+ := \min\{-a_\mu/b_\mu : b_\mu < 0\}, \tag{109}$$

where the minimum of the empty set is $+\infty$. The third test succeeds exactly when $a_\mu \geq 0$ for every $b_\mu = 0$ and $\lambda_- \leq \lambda_+$. In that case define $\lambda := \lambda_-$. This specifies every multiplier from the polynomial coefficients.

Start with $[0,1] \times [0, \vartheta_*]$ and apply the three tests in the displayed order. If none succeeds, split in z when $d_z < d_\vartheta + 3$ and in ϑ otherwise, where (d_z, d_ϑ) are the coordinate depths. Recompute L whenever the z interval changes. Exact rational evaluation gives 1421 nodes and 711 leaves: 636 pass the first test, 7 the second, and 68 the third. The maximum depths are 14 in z and 11 in ϑ. Each split partitions its parent, and the normalized leaf areas sum to $\sum 2^{-d_z - d_\vartheta} = 1$. This proves Eq. (102) on the entire feasible domain. The drift margin follows from Eq. (104); the final error allowances are treated in Section 5.

4.7 A covariance correction from the local trace

We now retain a diagonal and both off-diagonal local traces and control the mixed variation directly. The potential, its regularity, and endpoint moves are unchanged. In this section and the next, β, ι denote the polynomially weighted 3/4-power budget and coupling specified below. This budget satisfies $\beta(y) \geq 1 - y^2$ and has positive response scales, as required by the response and coercivity identities above.

Define $D_i := \widetilde{u}_i\widetilde{w}_i + \widetilde{v}_i\widetilde{z}_i$ and $\rho_{\mathrm{rat}\,i} := (\widetilde{u}_i\widetilde{z}_i + \widetilde{v}_i\widetilde{w}_i)/D_i$. The paired basis gives the exact identities

$$\langle e_i, \widehat{T} e_i\rangle = -\rho_{\mathrm{rat}\,i}/\Gamma_i, \qquad \langle f_i, \widehat{T} f_i\rangle = \rho_{\mathrm{rat}\,i}/\Gamma_i. \tag{110}$$

Indeed, $\mathscr{U}_{ii} = \iota\widetilde{u}_i^2$ and $\mathscr{V}_{ii} = \iota\widetilde{v}_i^2$. Substitution into the paired basis proves Eq. (110). Also

$$D_i(1 - \rho_{\mathrm{rat}\,i}) = (\widetilde{u}_i - \widetilde{v}_i)(\widetilde{w}_i - \widetilde{z}_i), \qquad \partial_i\mathcal{P} = \tau_i(\widetilde{w}_i - \widetilde{z}_i).$$

Define

$$\upsilon_i := \frac{\widetilde{v}_i\sqrt{\widetilde{z}_i} - \widetilde{u}_i\sqrt{\widetilde{w}_i}}{\sqrt{\widetilde{u}_i\widetilde{v}_i}(\sqrt{\widetilde{w}_i} + \sqrt{\widetilde{z}_i})}, \qquad L_i := \iota\sqrt{\widetilde{u}_i\widetilde{v}_i}(\widetilde{u}_i - \widetilde{v}_i), \qquad \delta_i := \Gamma_i^{-1}.$$

The additional off-diagonal identity is

$$\langle f_i, \widehat{T} e_i\rangle = L_i + \frac{\delta_i(\widetilde{u}_i - \widetilde{v}_i)\upsilon_i}{D_i}(\widetilde{w}_i - \widetilde{z}_i). \tag{111}$$

To verify it, direct substitution gives

$$\langle f_i, \widehat{T} e_i\rangle = L_i\frac{(\widetilde{u}_i + \widetilde{v}_i)\sqrt{\widetilde{w}_i\widetilde{z}_i}}{D_i},$$

and the factorization

$$(\widetilde{u}_i + \widetilde{v}_i)\sqrt{\widetilde{w}_i\widetilde{z}_i} - D_i = (\widetilde{w}_i - \widetilde{z}_i)\upsilon_i\sqrt{\widetilde{u}_i\widetilde{v}_i}$$

proves Eq. (111). Both trace deviations are explicit gradient terms, including when that gradient coordinate is zero. The other off-diagonal entry also factors through the gradient:

$$\langle e_i, \widehat{T} f_i\rangle = \Omega_i\partial_i\mathcal{P}, \qquad \Omega_i := \frac{\iota(\widetilde{u}_i\widetilde{v}_i)^{3/2}(\widetilde{w}_i + \widetilde{z}_i)}{D_i\tau_i\sqrt{\widetilde{w}_i\widetilde{z}_i}}. \tag{112}$$

Indeed, substitution into the paired basis gives $\iota(\widetilde{u}_i\widetilde{v}_i)^{3/2}(\widetilde{w}_i^2 - \widetilde{z}_i^2)/(D_i\sqrt{\widetilde{w}_i\widetilde{z}_i})$. Factoring the difference of squares proves the identity without any division by a gradient coordinate.

Let $\eta := 1 + \kappa_{\mathrm{tr}}$, with κ_{tr} specified in Section 4.8, and define the mixed coefficient and target

$$g_i := \beta_i'/\beta_i, \qquad \Lambda_i := \frac{g_i}{\tau_i\sqrt{\widetilde{u}_i\widetilde{v}_i}}, \qquad k_i := \frac{\iota(-\beta_i'')}{2\eta p_i q_i}.$$

We give the local scalar construction before stating the covariance comparison. Suppress the index, define $\delta := 1/\Gamma$ and $s_{\mathrm{inv}} := 1/k$, and define $H_i := \iota^2\beta_i^2 u_i v_i/(1-y_i^2) \in (0,1)$. Define $j_0 := j_0(|y_i|, H_i)$, $b_0 := b_0(|y_i|, H_i)$, and $b_1 := b_1(|y_i|, H_i)$ from Section 4.8. Define

$$j := \frac{\Lambda^2\delta s_{\mathrm{inv}}}{4}, \qquad \ell := \Lambda L s_{\mathrm{inv}},$$
$$F := {j_0}^2 + 2j_0 - j - 4js_{\mathrm{inv}}b_0 - 4js_{\mathrm{inv}}(s_{\mathrm{inv}} - \delta)b_0^2,$$
$$B := 2j_0 - j - \delta(1 + 2j_0) - 4js_{\mathrm{inv}}b_0 - 4j{s_{\mathrm{inv}}}^2b_0^2,$$

$$
\begin{aligned}
m_0 &:= \frac{1/2 + b_0(s_{\mathrm{inv}} - \delta)}{1 + j_0}, & n_0 &:= -\frac{b_0}{1 + j_0},\\
\mathfrak{f} &:= F + 4j s_{\mathrm{inv}} n_0^2 (F - B) + 2F\ell n_0,\\
\mathfrak{u} &:= B + \delta F + 4j s_{\mathrm{inv}} m_0 n_0 (B - F) - F\ell m_0,\\
\mathcal{P}_* &:= m_0\mathfrak{f} + (n_0 + b_1)\mathfrak{u}, & O_* &:= (1 + j_0)b_1,\\
Z_* &:= 2\mathfrak{f}\mathfrak{u} - (s_{\mathrm{inv}} - \delta)\mathfrak{f}^2 + 4j s_{\mathrm{inv}}\mathcal{P}_*^2,\\
N_* &:= \mathfrak{u}^2 + \delta\mathfrak{f}^2 + 4j s_{\mathrm{inv}}\mathcal{P}_*^2 - 2\delta\mathfrak{f}\mathfrak{u} + 2\ell\mathcal{P}_*\mathfrak{u},\\
M_* &:= \delta(\mathfrak{f} - \mathfrak{u})b_0 + \mathcal{P}_*(j_0 + \delta),\\
\mathfrak{z} &:= FZ_* - 4j s_{\mathrm{inv}} O_*^2\mathfrak{u}^2,\\
\mathfrak{p} &:= BZ_* - FN_* - 4j s_{\mathrm{inv}} O_*^2\mathfrak{u}^2 + 8j s_{\mathrm{inv}} O_* M_*\mathfrak{u} + 2\ell j_0 O_*\mathfrak{u}^2.
\end{aligned}
$$

The scalar certificate below proves

$$
F > 0, \quad B > 0, \quad \mathfrak{f} > 0, \quad \mathfrak{u} > 0, \quad \mathfrak{z} > 0, \quad \mathfrak{p} > 0. \tag{113}
$$

Define the resulting weights

$$
\begin{aligned}
a_* &:= \mathfrak{f}/\mathfrak{u}, & \gamma_* &:= \mathcal{P}_*/\mathfrak{u}, & d_* &:= \Gamma j_0,\\
b_* &:= \Lambda s_{\mathrm{inv}} b_0, & r_* &:= \Lambda s_{\mathrm{inv}}\gamma_*, & o_* &:= \Lambda s_{\mathrm{inv}} O_*,\\
z_* &:= Z_*/\mathfrak{u}^2, & \Delta_* &:= Fz_* - \delta o_*^2 = \mathfrak{z}/\mathfrak{u}^2,\\
w_e &:= F/\Delta_*, & w_f &:= \delta z_*/\Delta_*, & w_{ef} &:= -\delta o_*/\Delta_*.
\end{aligned}
$$

The paired weight block is positive definite, since

$$
\begin{pmatrix} w_e & w_{ef} \\ w_{ef} & w_f \end{pmatrix}^{-1} = \begin{pmatrix} z_* & o_* \\ o_* & F/\delta \end{pmatrix}, \qquad F > 0, \quad \Delta_* > 0. \tag{114}
$$

Define its row multipliers by

$$
\begin{aligned}
h_* &:= 2(a_* w_e + b_* w_{ef}), & q_* &:= 2(a_* w_{ef} + b_* w_f),\\
p_* &:= 2(r_* w_e + d_* w_{ef}), & j_* &:= 2(r_* w_{ef} + d_* w_f).
\end{aligned}
$$

These helper scalars are distinct from the state variables. The extra trace information allows both row auxiliaries to use both paired directions. The weight block also retains an off-diagonal entry; the corresponding trace corrections are canceled by the drift.

Lemma 4.5. *There is a centered normalized response ξ with covariance A, positive definite on the response subspace and zero on its orthogonal complement, such that, with $\mathsf{F}_i := f_i^\top \xi$, $\mathsf{Y}_i := e_i^\top (I - \widehat{T})\xi$, $K := \mathbb{E}[rr^\top]$, and $\pi_i := \tau_i^2 K_{ii}$,*

$$
\mathbb{E}[\mathcal{Q}(r)] + \sum_i \Lambda_i \mathbb{E}[\mathsf{Y}_i \mathsf{F}_i] \le \sum_i k_i D_i \pi_i + \nu \cdot \nabla\mathcal{P}, \tag{115}
$$

where

$$
\nu_i := \frac{\delta_i[h_{*,i} - j_{*,i} + p_{*,i}\upsilon_i](\widetilde{u}_i - \widetilde{v}_i)}{D_i\tau_i} + q_{*,i}\Omega_i.
$$

For $R_{\mathsf{A}} := (I - \widehat{T})\mathsf{A}(I - \widehat{T})^\top$, there are absolute $a, C > 0$ such that

$$
a k_{\mathrm{act}} \le \mathrm{tr}[R_{\mathsf{A}}] \le C k_{\mathrm{act}}. \tag{116}
$$

Proof. Use the weights above at each coordinate and define

$$\mathsf{D} := \sum_i [w_i^e e_i e_i^\top + w_i^f f_i f_i^\top + w_i^{ef}(e_i f_i^\top + f_i e_i^\top)],$$

$$\mathsf{A} := \mathsf{E}(\mathsf{E}^\top \mathsf{D}^{-1}\mathsf{E})^{-1}\mathsf{E}^\top, \qquad \mathsf{B} := \mathsf{D} - \mathsf{A},$$

where E is an orthonormal basis matrix for the response subspace. As before, $\mathsf{A}, \mathsf{B} \succeq 0$, $\mathsf{A}\mathsf{D}^{-1}\mathsf{A} = \mathsf{A}$, $\mathsf{A}\mathsf{D}^{-1}\mathsf{B} = 0$, and $\mathrm{tr}[\mathsf{D}^{-1}\mathsf{A}] = k_{\mathrm{act}}$. No commutation is required.

Fix one pair and suppress its index. Define $v_* := a_* e + b_* f$, $w_* := r_* e + d_* f$, $\widehat{e} := e - \Lambda f/(2k)$, $t_e := e^\top \widehat{T}$, and $t_f := f^\top \widehat{T}$. For $C_* := \Gamma\mathsf{B} + k\mathsf{A} \succ 0$, the projection identities give

$$C_*^{-1} = \delta \mathsf{D}^{-1} + (1/k - \delta)\mathsf{D}^{-1}\mathsf{A}\mathsf{D}^{-1}.$$

Completing a square in t_e yields

$$\begin{aligned}\Gamma t_e \mathsf{B} t_e^\top + k(e^\top - t_e)\mathsf{A}(e - t_e^\top) - \Lambda(e^\top - t_e)\mathsf{A}f + t_e(2\mathsf{D}v_*) \\ \geq 2\widehat{e}^\top \mathsf{A} v_* - (1/k - \delta) v_*^\top \mathsf{A} v_* - \frac{\Lambda^2}{4k} f^\top \mathsf{A} f - \delta v_*^\top \mathsf{D} v_*.\end{aligned}$$

The response condition gives $t_f \mathsf{A} = f^\top \mathsf{A}$. A square with the PSD matrix B, which need not be invertible, gives

$$\Gamma t_f \mathsf{B} t_f^\top + t_f(2\mathsf{D}w_*) \geq 2 f^\top \mathsf{A} w_* + \delta w_*^\top \mathsf{A} w_* - \delta w_*^\top \mathsf{D} w_*.$$

The two diagonal coefficients of A in these bounds are z_* and $-\Lambda s_{\mathrm{inv}} b_* - (s_{\mathrm{inv}} - \delta) b_*^2 - \Lambda^2 s_{\mathrm{inv}}/4 + 2d_* + \delta d_*^2 = F/\delta$. The off-diagonal coefficient is

$$b_* - \frac{\Lambda s_{\mathrm{inv}} a_*}{2} - (s_{\mathrm{inv}} - \delta) a_* b_* + r_* + \delta r_* d_* = \Lambda s_{\mathrm{inv}}(1 + j_0) b_1 = o_*.$$

Indeed, the definitions give $b_0 - a_*/2 - (s_{\mathrm{inv}} - \delta) a_* b_0 + (1 + j_0)\gamma_* = (1 + j_0) b_1$. Multiplication by Λs_{inv} proves the identity, including when $\Lambda = 0$. By Eq. (114), the summed coefficient is exactly D^{-1}.

For the scalar constant term, define

$$\mathcal{E} := w_e + w_f + \delta(v_*^\top \mathsf{D} v_* + w_*^\top \mathsf{D} w_*) + (-h_* + j_*)\delta + p_* L.$$

The exact identity

$$1 - \mathcal{E} = \frac{\mathfrak{p}}{\mathfrak{z}} > 0 \tag{117}$$

follows by substituting the three block weights. More explicitly, $\delta \Lambda^2 {s_{\mathrm{inv}}}^2 = 4 j s_{\mathrm{inv}}$ gives

$$\begin{aligned}\Delta_*(1 - \mathcal{E}) = B z_* - F[1 + \delta a_*^2 + 4 j s_{\mathrm{inv}} \gamma_*^2 - 2\delta a_* + 2\ell\gamma_*] \\ - 4 j s_{\mathrm{inv}} O_*^2 + 8 j s_{\mathrm{inv}} O_*[\delta(a_* - 1) b_0 + \gamma_*(j_0 + \delta)] + 2\ell j_0 O_*.\end{aligned}$$

Multiplication by $\mathfrak{u}^2$ gives $\mathfrak{p}$, while $\mathfrak{u}^2 \Delta_* = \mathfrak{z}$, proving Eq. (117).

Sum the two row inequalities over all coordinates. Coercivity gives

$$\mathrm{tr}[\widehat{D_{\mathrm{aux}}}\mathsf{A}] \leq \mathrm{tr}[\widehat{D_{\mathrm{aux}}}\mathsf{D}] - \mathrm{tr}[\Gamma \widehat{T} \mathsf{B} \widehat{T}^\top], \qquad \mathrm{tr}[\widehat{D_{\mathrm{aux}}}\mathsf{D}] = \sum_i (w_i^e + w_i^f).$$

The second equality uses the identity local blocks of $\widehat{D_{\mathrm{aux}}}$, including their zero off-diagonal entries. Using the opposite diagonal traces, the remaining correction to the weighted movement is at most

$$\sum_i [(h_{*,i} - j_{*,i})(\langle e_i, \widehat{T} e_i\rangle + \delta_i) + q_{*,i}\langle e_i, \widehat{T} f_i\rangle + p_{*,i}(\langle f_i, \widehat{T} e_i\rangle - L_i)].$$

Here the terms $\mathcal{E}_i - 1$ were dropped using Eq. (117). Eqs. (110), (111), and (112) identify this correction as $\nu \cdot \nabla \mathcal{P}$. Since $\mathbb{E}[\mathcal{Q}(r)] = \mathrm{tr}[\widehat{D_{\mathrm{aux}}}\mathsf{A}]$ and $\langle e_i, R_{\mathsf{A}} e_i\rangle = D_i \pi_i$, this proves Eq. (115).

For movement, the compact scalar certificate in Lemma 4.6 bounds $F, \mathfrak{f}, \mathfrak{u}, \mathfrak{z}$ above and away from zero. The fixed polynomial b_1 is bounded, and $\delta o_*^2 = 4 j s_{\mathrm{inv}} O_*^2 \le C$. The block formula therefore gives $w_e \le C$, $w_f \le C\delta$, and $|w_{ef}| \le C\sqrt{\delta}$. Its inverse has $z_* \le C$, $F/\delta \le C\Gamma$, and $|o_*| \le C\sqrt{\Gamma}$. Bounding the cross term by its two diagonal terms gives

$$\mathsf{D} \preceq CI, \qquad \mathsf{D}^{-1} \preceq C\Pi_0 + C\Gamma\Pi_1, \qquad \Pi_1 := I - \Pi_0.$$

With $C_* := \mathsf{E}^\top \mathsf{D}^{-1} \mathsf{E}$, the response condition $\Pi_1 \mathsf{E} = \Pi_1 \widehat{T} \mathsf{E}$ and coercivity imply

$$\mathrm{tr}[C_*] \le C k_{\mathrm{act}} + C\|\Gamma^{1/2}\Pi_1 \widehat{T}\mathsf{E}\|_F^2 \le C k_{\mathrm{act}} + C\,\mathrm{tr}[\widehat{T}^\top \Gamma \widehat{T}] \le C k_{\mathrm{act}}.$$

Weighted Frobenius Cauchy–Schwarz gives

$$(k_{\mathrm{act}} - \mathrm{tr}[\widehat{T}\Pi])^2 \le \|(I - \widehat{T})\mathsf{E} C_*^{-1/2}\|_F^2 \|\mathsf{E} C_*^{1/2}\|_F^2 = \mathrm{tr}[R_{\mathsf{A}}]\,\mathrm{tr}[C_*].$$

Since $\|\widehat{T}\|_F^2 \le 2k_{\mathrm{act}}/\iota$ and $\iota > 2$, $|\,\mathrm{tr}[\widehat{T}\Pi]| \le k_{\mathrm{act}}\sqrt{2/\iota} < k_{\mathrm{act}}$. This proves the lower bound independently of the smallest weight. The upper bound follows from $\mathsf{A} \preceq CI$ and $\|I - \widehat{T}\|_F^2 \le C k_{\mathrm{act}}$. □

The mixed variation and its correction. Retain the exact mixed term in the derivation of Eq. (89), before applying Cauchy–Schwarz:

$$\frac{1}{2} r^\top \nabla^2 \mathcal{P} r \le \mathcal{Q}(r) + \sum_i m_i(r) - \frac{\iota}{2}\sum_i (-\beta_i'')(v_i w_i + u_i z_i) r_i^2, \qquad m_i(r) := \iota \beta_i' r_i (w_i \dot{v}_i + z_i \dot{u}_i). \tag{118}$$

For the normalized response above, write $\mathsf{Y}_i = \sqrt{D_i} J_i$ and $\mathsf{E}_i := e_i^\top \xi$. The exact identity

$$m_i(r) = \Lambda_i \mathsf{Y}_i \mathsf{F}_i + \partial_i \mathcal{P} \frac{g_i}{\tau_i^2 D_i} \mathsf{Y}_i (\mathsf{E}_i - \mathsf{Y}_i + \upsilon_i \mathsf{F}_i) \tag{119}$$

holds even when the gradient coordinate is zero. Indeed, the response equations give

$$m_i(r) = \frac{g_i J_i}{\tau_i}[\widetilde{w}_i P_i + \widetilde{z}_i Q_i + (\widetilde{z}_i - \widetilde{w}_i) J_i],$$

and the paired basis expresses the first two terms as

$$\widetilde{w}_i P_i + \widetilde{z}_i Q_i = \frac{\widetilde{w}_i - \widetilde{z}_i}{\sqrt{D_i}} \mathsf{E}_i + \frac{(\widetilde{u}_i + \widetilde{v}_i)\sqrt{\widetilde{w}_i \widetilde{z}_i}}{\sqrt{\widetilde{u}_i \widetilde{v}_i}\sqrt{D_i}} \mathsf{F}_i.$$

The factorization used in Eq. (111) proves Eq. (119) without dividing by $\widetilde{w}_i - \widetilde{z}_i$. Define

$$\theta_i := \frac{g_i}{\tau_i^2 D_i} \mathbb{E}[\mathsf{Y}_i (\mathsf{E}_i - \mathsf{Y}_i + \upsilon_i \mathsf{F}_i)].$$

Combining the mixed identity with Eq. (115) gives

$$\frac{1}{2}\mathbb{E}[r^\top \nabla^2 \mathcal{P} r] \le \sum_i k_i D_i \pi_i + (\nu + \theta) \cdot \nabla \mathcal{P} - \sum_i \frac{\iota(-\beta_i'')}{2 p_i q_i} \rho_{\mathrm{rat}_i} D_i \pi_i. \tag{120}$$

Thus the mixed variation is included directly in the covariance bound.

4.8 The improved constant and its exact certificate

We use a 3/4 endpoint power and adjust the curvature by a fixed polynomial. Define

$$\iota_{\mathrm{tr}} := \frac{699}{250}, \qquad p := \frac{3}{4}, \qquad \kappa_{\mathrm{tr}} := \frac{1}{10^6}, \qquad \eta := 1 + \kappa_{\mathrm{tr}}, \qquad \beta_{\mathrm{tr}}(y) := (1-y^2)^p f_{\mathrm{tr}}(|y|),$$

where the polynomial factor is

$$\begin{aligned} f_{\mathrm{tr}}(z) := 1 &+ 0.1915505z^2 + 0.07716335z^4 + 0.03917979z^6 \\ &+ 0.04234609z^8 - 0.08891557z^{10} + 0.40792039z^{12} \\ &- 0.91331458z^{14} + 1.30793628z^{16} - 1.02145746z^{18} \\ &+ 0.35713938z^{20}. \end{aligned} \tag{121}$$

For comparison, the polynomial factor of β_{fast} is

$$\begin{aligned} f_{\mathrm{fast}}(z) := 1 &+ 0.0498735123z^2 - 0.005706553z^3 + 0.3937540726z^4 \\ &- 11.0832914979z^5 + 203.0443576388z^6 - 2445.8186989094z^7 \\ &+ 20520.6541832702z^8 - 124448.8676898058z^9 + 559322.9751101618z^{10} \\ &- 1893873.5450890157z^{11} + 4878891.695382722z^{12} - 9600597.735909414z^{13} \\ &+ 14402676.373907391z^{14} - 16328233.13094071z^{15} + 13740497.154438846z^{16} \\ &- 8313646.91225735z^{17} + 3417391.239657022z^{18} - 854236.6692651988z^{19} \\ &+ 97990.3280047086z^{20}. \end{aligned} \tag{122}$$

All decimal coefficients in these two expressions are exact. The notation f_{fast} distinguishes the polynomial f used in Section 4.1 from f_{tr}.

The covariance helpers are the fixed polynomials

$$\begin{aligned} j_0(z,H) &:= \frac{84047268 - 52938705z + 133367380z^2 + 49643028z^3}{10^8} \\ &\quad + H\frac{22863212 + 66017181z - 147568585z^2}{10^8}, \\ b_0(z,H) &:= \frac{-41550398 + 11091757z - 14435171z^2}{10^8} \\ &\quad + H\frac{10275247 - 10715575z + 21317210z^2}{10^8}, \\ b_1(z,H) &:= \frac{3323951 + 40587513z - 130152459z^2 + 191253811z^3}{10^8} \\ &\quad + H\frac{-2355589 + 49791446z - 117532304z^2}{10^8}. \end{aligned}$$

Here $H_i := \iota^2\beta_i^2 u_i v_i/(1-y_i^2)$ is the normalized local product. The light inequalities give $0 < H_i < 1$. For the remainder of this subsection, define $\iota := \iota_{\mathrm{tr}}$, $\beta := \beta_{\mathrm{tr}}$, and $f := f_{\mathrm{tr}}$. For $z \in [0,1)$, define $r_{\mathrm{aux}} := 1 - z^2$ and define the polynomials

$$\mathfrak{a}(z) := r_{\mathrm{aux}} f'(z) - 2pzf(z), \qquad \mathfrak{d}(z) := 2p[1 + (1-2p)z^2]f(z) + 4pzr_{\mathrm{aux}} f'(z) - {r_{\mathrm{aux}}}^2 f''(z).$$

Direct differentiation gives

$$\beta'(z) = {r_{\mathrm{aux}}}^{p-1}\mathfrak{a}(z), \qquad -\beta''(z) = {r_{\mathrm{aux}}}^{p-2}\mathfrak{d}(z), \qquad -\frac{\beta'(z)}{\beta(z)} = -\frac{\mathfrak{a}(z)}{r_{\mathrm{aux}} f(z)}.$$

The exact polynomial certificate gives

$$\frac{f'(z)}{z} > 0, \quad -\frac{\mathfrak{a}(z)}{z} > 0, \quad \mathfrak{d}(z) > \frac{1}{2}, \qquad \frac{1}{4} < j_0(z, H) < 3. \tag{123}$$

The bounds hold for $z, H \in [0, 1]$, with the polynomial continuations of the quotients at zero. Since $f(0) = 1$, the first inequality implies $f \geq 1$. Thus

$$\beta(0) = 1, \quad \beta(\pm 1) = 0, \quad 1 - y^2 \leq \beta(y) \leq 1, \quad \beta(y) \geq (1 - y^2)^{3/4}. \tag{124}$$

The upper bound follows from $\beta' \leq 0$ and evenness. Also

$$0 \leq -\frac{\beta'(z)}{\beta(z)} = \frac{2pz}{1 - z^2} - \frac{f'(z)}{f(z)} \leq \frac{3}{4(1 - z)}.$$

At a light state with $y_i \geq 0$, this gives $p_i \geq 1/4$ and $q_i \geq 1$, with the roles exchanged for negative coordinates. The potential, endpoint, response, and coercivity arguments therefore apply. The budget is smooth on $(-1, 1)$, with $\beta'(0) = 0$. On the truncated interval its fixed polynomial factor gives $|\beta^{(j)}| \leq C\delta_{\text{end}}^{3/4-j}$ and $|\beta^{(j)}|/\beta \leq C\delta_{\text{end}}^{-j}$ for $1 \leq j \leq 3$.

The helpers $j_{0i} := j_0(|y_i|, H_i)$ and $b_{r_{\text{aux}},i} := b_{r_{\text{aux}}}(|y_i|, H_i)$, $r_{\text{aux}} = 0, 1$, are used pointwise in the covariance construction. No derivatives of them are needed. The remaining certificate verifies that construction on the entire light-state domain.

Lemma 4.6. *At every light state, the choice*

$$k_i := \frac{\iota(-\beta_i'')}{2\eta p_i q_i} \tag{125}$$

satisfies all six conditions in Eq. (113). The quantities $F, B, \mathfrak{f}, \mathfrak{u}, \mathfrak{z}$ are uniformly bounded above and away from zero.

Proof. By symmetry take $z = y_i \geq 0$ and define $q_0 := -\beta'(z)/\beta(z)$, $a_{\text{loc}} := \iota\beta u_i$, $b_{\text{loc}} := \iota\beta v_i$, and $P := a_{\text{loc}} b_{\text{loc}}$. The light rectangle is $0 < a_{\text{loc}} < 1 + z$, $0 < b_{\text{loc}} < 1 - z$. Writing $\mathcal{R} := p_i q_i$, direct substitution gives

$$\mathcal{R} = 1 + q_0(a_{\text{loc}} - b_{\text{loc}}) - q_0^2 P, \qquad \delta = \frac{P}{\iota\beta}, \qquad \Lambda^2\delta = \frac{\iota(\beta')^2}{\beta\mathcal{R}}, \qquad \Lambda L = \frac{1 + q_0^2 P - \mathcal{R}}{\mathcal{R}}.$$

At fixed $P \in [0, r_{\text{aux}}]$, the function $a_{\text{loc}} - P/a_{\text{loc}}$ is increasing. Its two boundary values, at $b_{\text{loc}} = 1 - z$ and $a_{\text{loc}} = 1 + z$, give the complete parametrization

$$P = r_{\text{aux}} H, \qquad a_{\text{loc}} - b_{\text{loc}} = z(1 + H) + (1 - H)(2V - 1), \qquad H, V \in [0, 1].$$

The boundary values are affine in P and agree at $P = r_{\text{aux}}$. Define $Q_y := z(1+H)+(1-H)(2V-1)$ and $\varrho := {r_{\text{aux}}}^{1/4}$. With the budget polynomials evaluated at z and the helpers at (z, H), the scalar data become

$$\delta = \frac{\varrho H}{\iota f}, \qquad j = \frac{\eta\mathfrak{a}^2}{2f\mathfrak{d}},$$
$$s_{\text{inv}} = \frac{2\eta\varrho[r_{\text{aux}} f^2 - \mathfrak{a} f Q_y - H\mathfrak{a}^2]}{\iota\mathfrak{d} f^2}, \qquad \sigma := \frac{2\eta\varrho r_{\text{aux}}}{\iota\mathfrak{d}}, \qquad \ell = \sigma + 4j\delta - s_{\text{inv}}.$$

These formulas keep the mixed coefficient and off-diagonal reference correlated. They remain finite at both endpoints.

For an exact certificate without singular fractional powers, define

$$z := 1 - t^4, \qquad r_{\mathrm{aux}} = t^4(2 - t^4), \qquad \varrho = t(2 - t^4)^{1/4}, \qquad 0 \le t \le 1.$$

The six expressions in Eq. (113) are now continuous functions on the closed cube $(t, H, V) \in [0, 1]^3$. Their only nonpolynomial operation is the fourth root of $2 - t^4 \in [1, 2]$.

The first three polynomial inequalities in Eq. (123) have strictly positive Bernstein coefficients on $[0, 1]$; their degrees are $18, 20, 22$, respectively. Since j_0 is affine in H, its bounds reduce to $H = 0, 1$. Their cubic Bernstein coefficients are positive on $[0, 1]$. For the full-state inequalities, use intervals with endpoints in $2^{-100}\mathbb{Z}$. Define $S := 2^{100}$ and define outward rounding by

$$\lfloor x \rfloor_S := \frac{\lfloor Sx \rfloor}{S}, \qquad \lceil x \rceil_S := \frac{\lceil Sx \rceil}{S}. \tag{126}$$

Enclose each rational constant by these two values. Addition and subtraction use the endpoint formulas; multiplication takes the minimum and maximum of the four endpoint products and rounds outwards. For $0 \notin [a, b]$, inversion gives $[\lfloor 1/b \rfloor_S, \lceil 1/a \rceil_S]$. Integer powers use their exact range on the interval, including zero when an even power crosses zero, followed by outward rounding. For $0 \le a \le b$ with integer endpoints, the fourth root of $[a/S, b/S]$ is enclosed by

$$\frac{1}{S}[\lfloor (aS^3)^{1/4} \rfloor, \lceil (bS^3)^{1/4} \rceil]. \tag{127}$$

Both integer roots are determined by comparisons of fourth powers. Define $\widehat{f}$ by $f(z) = \widehat{f}(z^2)$. Evaluate $\widehat{f}, \widehat{f}', \widehat{f}''$ by Horner's rule and use

$$f'(z) = 2z\widehat{f}'(z^2), \qquad f''(z) = 2\widehat{f}'(z^2) + 4z^2\widehat{f}''(z^2).$$

Evaluate the helper polynomials by Horner's rule in z, retaining their affine dependence on H. Substitute into the displayed scalar formulas, preserving their factorizations. In these formulas, evaluate squares of computed quantities as products of two copies of the same interval. Differentiate by the product, quotient, and chain rules, using the same outward interval operations to enclose every first derivative.

Here is the complete subdivision rule. Let $G_1, \ldots, G_6$ be the expressions in Eq. (113), in that order. Start with $[0, 1]^3$. On a box Q with midpoint x_0 and half-widths h_1, h_2, h_3, first compute interval enclosures $I_i(Q)$ of $G_i(Q)$. Accept Q if every lower endpoint is positive. Otherwise compute $I_i(x_0)$ and derivative enclosures $D_{ij}(Q)$, and define

$$c_i := \inf I_i(x_0), \qquad e_{ij} := \lceil h_j \sup\{|u| : u \in D_{ij}(Q)\} \rceil_S, \tag{128}$$

$$L_i := \max\{\inf I_i(Q), c_i - \sum_{j=1}^{3} e_{ij}\}. \tag{129}$$

The mean-value theorem gives $G_i(x) \ge L_i$ throughout Q. Accept Q when all six L_i are positive. If some remain nonpositive, bisect the coordinate maximizing

$$\max_{i : L_i \le 0} \frac{e_{ij}}{\max\{S^{-1}, c_i\}}, \tag{130}$$

breaking ties in the order t, H, V. If an interval denominator contains zero, bisect in t before continuing.

Exact evaluation of these rules gives 636477 nodes and 318239 accepted leaves, with maximum coordinate depths $(15, 19, 6)$. Every accepted lower bound exceeds $3{\cdot}10^{-9}$. Each bisection partitions its parent, so the leaves cover the cube and have disjoint interiors; their exact volumes sum to one. This verifies the inequalities on the full closed cube, including all light-state limits. Compactness gives the stated uniform bounds and proves the lemma. □

Define the drift

$$h_i := -\frac{k_i}{\tau_i}(\widetilde{u}_i - \widetilde{v}_i)\pi_i - \nu_i - \theta_i.$$

The last two terms cancel the gradient corrections in Eq. (120). The first changes $k_i D_i \pi_i$ into $k_i \rho_{\mathrm{rat}\,i} D_i \pi_i$. Since $\rho_{\mathrm{rat}\,i} D_i \pi_i = p_i q_i (v_i w_i + u_i z_i) K_{ii}$, Eq. (125) gives

$$\mathscr{L}\mathcal{P} \le -\kappa_{\mathrm{tr}} \sum_i k_i \rho_{\mathrm{rat}\,i} D_i \pi_i < 0. \tag{131}$$

The weights are positive and Eq. (116) ensures nonzero movement. No derivative of the covariance or drift is used. This proves the local assertion in [SY26a, Theorem 3.1]. Its polynomial implementation is given in Section 5.10.

5 Shared deterministic numerical tools

This section develops the reduction, rounding, and numerical response tools used by the deterministic algorithm in Section 9 and the Las Vegas algorithm in Section 10. The reference estimates retain their rank-one hypotheses; the higher-rank normalization, curvature guarantees, and total operation bounds are supplied in those sections.

Section 5.1 fixes the matrix multiplication exponents used in the operation bounds. Section 5.2 reduces the number of fractional coordinates while controlling the residual signed sum. Section 5.3 specifies the maintained matrix model and bounds the error from deferred rounding near the endpoints.

Section 5.4 establishes conditioning and stability estimates for the optimality equations. Section 5.5 shows how to find progress through either a large gradient coordinate or a direction of negative curvature. Section 5.6 controls directional derivatives and optimizer responses at the budget scale, giving the estimates needed to choose the step length.

Section 5.7 implements the required operator actions through rectangular matrix products. Section 5.8 controls feasibility and the number of regular endpoint moves. Section 5.9 bounds the arithmetic cost of the numerical solves and curvature search in each stage. Finally, Section 5.10 develops the trace-corrected local estimates and deterministic response constructions used in Section 9.

5.1 Multiplication parameters

For the operation bounds below, define

$$\omega_0 := 2.371177, \qquad \omega_{\mathrm{rect}0} := 2.042777, \qquad \chi := 3.250036, \qquad p_0 := 2(\omega_0 - \omega_{\mathrm{rect}0}) = 0.6568. \tag{132}$$

Dupont et al. [DEK+26] prove $\omega < 2.371177$ for the square matrix-multiplication exponent. Thus ω_0 is a fixed upper bound used in the operation counts, not the asserted value of ω. The rectangular bounds $\omega(1, 1/2, 1) \le 2.042776 < \omega_{\mathrm{rect}0}$ and $\omega(1, 2, 1) \le 3.250035 < \chi$ follow from Alman et al. [ADW+25, Table 1]. The strict slack permits fixed recursive algorithms at the stated exponents.

The shared fast-budget estimates use β_{fast}, $\iota = 14779/2500$, and its parameter τ from Section 4. Section 5.10 states the specialization to the trace-corrected budget. Constants $C, a > 0$ are absolute, respectively sufficiently large and small. The covariance and drift prove that a descent direction exists; the constructions compute directions from the gradient and Hessian.

5.2 Vector reduction and inherited coefficients

We reduce the number of fractional coordinates while allowing a controlled residual, so the dependence calculations use short precision.

Lemma 5.1 (Short-precision partial signing). *Let $b_1, \dots, b_m \in \mathbb{R}^{d_{\mathrm{vec}}}$ satisfy $\|b_i\|_2 \le 2$, with rational entries of $O(\log(2md_{\mathrm{vec}}/\eta))$ bits, where $0 < \eta < 1$. Define $\chi := 211/400$. There is a deterministic algorithm returning $y \in [-1,1]^m$ with at most d_{vec} fractional coordinates and*

$$\|\sum_i y_i b_i\|_2 \le \eta.$$

Its arithmetic and bit costs are

$$\widetilde{O}(m{d_{\mathrm{vec}}}^{1+\chi} + {d_{\mathrm{vec}}}^{5/2}),$$

where the suppressed logarithmic factors also include $\log(1/\eta)$. In particular, for $d_{\mathrm{vec}} = n^2$ and logarithmic input word length, both costs are $\widetilde{O}(mn^{3.055} + n^5)$.

Proof. All unprocessed coefficients start at zero. Retain the fractional columns in a matrix A, insert the next input with coefficient zero, and, whenever there are $d_{\mathrm{vec}} + 1$ retained columns, perform the following approximate nullspace move. Choose a positive dyadic δ with

$$\frac{1}{2}(\frac{\eta}{100md_{\mathrm{vec}}})^2 \le \delta \le (\frac{\eta}{100md_{\mathrm{vec}}})^2, \qquad G := \delta I + AA^\top, \qquad \Pi_{\mathrm{rid}} := I - A^\top G^{-1} A.$$

The matrix Π_{rid} is PSD and has trace at least one. Choose a largest diagonal entry, say $\Pi_{\mathrm{rid}\,jj}$, and define $r := \Pi_{\mathrm{rid}} e_j$. Then

$$r_j \ge \frac{1}{d_{\mathrm{vec}} + 1}, \qquad \|Ar\|_2 \le \frac{\sqrt{\delta}}{2}.$$

The latter inequality follows by applying singular value decomposition to $A\Pi_{\mathrm{rid}} = \delta(\delta I + AA^\top)^{-1}A$: its singular values are $\delta s/(\delta + s^2) \le \sqrt{\delta}/2$. Singular value decomposition is used only in this proof, not by the algorithm.

Define $\tau := \eta/(10^4 m{d_{\mathrm{vec}}}^2)$. Compute the diagonal entries and the chosen column to entrywise error at most τ. Choosing the largest computed diagonal still gives a computed direction $\widetilde{r}$ with $\|\widetilde{r}\|_\infty \ge 1/(2(d_{\mathrm{vec}} + 1))$. Move in its positive direction until the first coefficient reaches a boundary. The step length is at most $4(d_{\mathrm{vec}} + 1)$. The boundary ratios are formed and compared exactly from the short dyadic coefficients and direction; this requires only a constant multiple of their word length. Set the hitting coefficient exactly to its boundary value, round the other coefficients inward to a sufficiently fine common dyadic grid, and remove every coefficient at a boundary. Each move removes at least one retained column. Since $\|A\| \le 2\sqrt{d_{\mathrm{vec}} + 1}$, the change in the signed sum has norm at most

$$4(d_{\mathrm{vec}} + 1)(\frac{\sqrt{\delta}}{2} + 2(d_{\mathrm{vec}} + 1)\tau) + 2(d_{\mathrm{vec}} + 1)u,$$

where u is the coefficient mesh size. With $u \le \tau$, the sum of these bounds over at most m moves is less than η. Insertions do not change the signed sum. This proves the output guarantee once the matrix evaluations are justified.

We use two levels of batching. An epoch ends after d_{vec} column insertions or deletions. At its start let C contain all currently retained columns and the next at most d_{vec} input columns. There are at most $d_{\mathrm{vec}}+1$ retained columns, so C has at most $2d_{\mathrm{vec}}+1$ columns. Keep this same C throughout the epoch, including columns subsequently deleted from the active set. Form the current G and $\mathsf{G} := C^{\top}G^{-1}C$ from scratch. Square multiplication with exponent $5/2$ and short-precision inversion cost $\widetilde{O}({d_{\mathrm{vec}}}^{5/2})$. The valid exponent $5/2$ is already more than sufficient here. During the epoch, inserting or deleting column C_i changes G by $\varepsilon_{\mathrm{upd}}C_iC_i^{\top}$, with $\varepsilon_{\mathrm{upd}} \in \{1,-1\}$, and hence changes G by

$$\mathsf{G}_{\mathrm{new}} = \mathsf{G} - \frac{\varepsilon_{\mathrm{upd}}}{1+\varepsilon_{\mathrm{upd}}\mathsf{G}_{ii}}\mathsf{G}_{:,i}\mathsf{G}_{i,:}. \tag{133}$$

Define $b_{\mathrm{batch}} := \lceil {d_{\mathrm{vec}}}^{\chi} \rceil$. Keep these updates as a list of at most b_{batch} outer products. Recovering a requested column costs $O(d_{\mathrm{vec}}b_{\mathrm{batch}})$; updating all diagonal entries costs $O(d_{\mathrm{vec}})$. The restriction of $I-\mathsf{G}$ to the retained columns supplies Π_{rid}. Before the next update would exceed this limit, collect the pending terms as $\mathsf{A}\,\mathrm{diag}(\alpha_{\mathrm{aux}})\mathsf{A}^{\top}$, add this product to the explicit base matrix G, and clear the list. Here A has $O(d_{\mathrm{vec}})$ rows and at most b_{batch} columns. The rectangular multiplication bound $\omega(1,\chi,1) < 2.055$ of Alman et al. [ADW+25, Table 1] gives cost $O({d_{\mathrm{vec}}}^{1+2\chi})$ for this operation. This writes the sum of pending updates explicitly; it does not recompute an inverse. At an epoch boundary discard the list and recompute from the actual retained columns. No product is needed merely to clear the final incomplete list. If several coordinates hit boundaries, an epoch may end between their deletions; the new epoch includes every column still retained.

There are at most $2m$ insertions and deletions, at most m nullspace moves, and $O(m/b_{\mathrm{batch}})$ products that write pending updates explicitly. The total cost is therefore

$$\begin{aligned}&\widetilde{O}((m/d_{\mathrm{vec}}+1){d_{\mathrm{vec}}}^{5/2} + (m/b_{\mathrm{batch}}){d_{\mathrm{vec}}}^{1+2\chi} + md_{\mathrm{vec}}b_{\mathrm{batch}})\\ &\quad = \widetilde{O}(m{d_{\mathrm{vec}}}^{1+\chi} + {d_{\mathrm{vec}}}^{5/2}),\end{aligned}$$

because $\chi > 1/2$.

Section 7.2 proves that the same batching attains the prescribed entry errors and residual allowance in logarithmic precision, and establishes the bit bound.

□

Grouping before applying a small reduction also appears in the Carathéodory construction of Maalouf, Jubran, and Feldman [MJF19]. Here a balanced tree of groups combines cube constraints, inherited coefficients, and controlled numerical error.

Lemma 5.2 (Partial signing by grouped coordinates)**.** *Let $b_1,\dots,b_m \in \mathbb{R}^{d_{\mathrm{vec}}}$ have norm at most two and entries with a common rational denominator. The numerators and denominator have $O(\log(2md_{\mathrm{vec}}/\eta))$ bits, where $0<\eta<1$. A deterministic algorithm returns $y \in [-1,1]^m$ with at most d_{vec} fractional coordinates and*

$$\|\sum_i y_i b_i\|_2 \le \eta$$

in $\widetilde{O}(md_{\mathrm{vec}} + {d_{\mathrm{vec}}}^{2.5275})$ arithmetic and bit operations. The suppressed factors include $\log(1/\eta)$.

Proof. Write $D := d_{\mathrm{vec}}$. We first give a small rounding procedure. Let A have $r_{\mathrm{aux}} \le 2D$ columns of norm at most two and let $y_{\mathrm{aux}} \in [-1,1]^{r_{\mathrm{aux}}}$ be any short dyadic coefficient vector. For $0<\varepsilon<1$, we can leave at most D fractional coordinates while changing Ay_{aux} by norm at most ε, at cost $\widetilde{O}(D^{2.5275})$ in both models. Coordinates already at a boundary are fixed and removed before this procedure.

For each remaining active matrix define $G := \delta I + AA^\top$ and $\Pi_{\mathrm{rid}} := I - A^\top G^{-1} A$, with a positive dyadic δ satisfying

$$\tfrac{1}{2}(\varepsilon/(100D^2))^2 \le \delta \le (\varepsilon/(100D^2))^2.$$

When $r_{\mathrm{aux}} > D$, $\mathrm{tr}[\Pi_{\mathrm{rid}}] \ge r_{\mathrm{aux}} - D \ge 1$, so some diagonal entry is at least $1/(2D)$. As in Lemma 5.1, the corresponding column r satisfies $\|Ar\|_2 \le \sqrt{\delta}/2$. Approximate the diagonal and this column to entrywise error $\tau \le \varepsilon/(10^4 D^3)$. Choosing the largest computed diagonal gives a computed direction of maximum absolute entry at least $1/(4D)$. From any y_{aux} in the cube, the positive boundary step therefore has length at most $8D$. Form and compare the boundary ratios exactly, set a hitting coordinate exactly to its endpoint, and round the other coordinates inward on a common dyadic grid of mesh $u \le \tau$. Remove every boundary coordinate. At most D moves are needed. Their total change in Ay_{aux} has norm at most

$$D(8D(\sqrt{\delta}/2 + 4D\tau) + 4Du) < \varepsilon.$$

Here the rounding error is at most the sum of the column norms times the coefficient errors. The argument does not require $y_{\mathrm{aux}} = 0$.

We implement the small procedure by batching its deletion updates. Keep its original matrix C, which has at most $2D$ columns, including columns subsequently removed from the active set. Define $a := 211/400$ and $b_{\mathrm{batch}} := \lceil D^a \rceil$. Initially form $\mathsf{G} := C^\top G^{-1} C$ in $\widetilde{O}(D^{5/2})$ operations, using square multiplication and short-precision Newton–Schulz inversion. Represent subsequent changes to G by an explicit base matrix and a list of at most b_{batch} rank-one terms. Keep its diagonal separately. A requested column is the corresponding base column plus the pending terms, and costs $O(Db_{\mathrm{batch}})$ to recover.

For a deletion at position j, recover the full column $z = \mathsf{G}_{:,j}$. Eq. (133) adds $zz^\top/(1-\mathsf{G}_{jj})$ to G. Append this term to the list and update every diagonal entry in $O(D)$ operations. Consequently choosing a direction, taking a boundary step, rounding coefficients, and processing a deletion cost $O(Db_{\mathrm{batch}})$ in total per deleted coordinate. The inherited vector y_{aux} changes only the boundary ratios; it does not enter G or its update formula. Process simultaneous boundary hits as consecutive deletions before the next rounding move. There are at most $2D$ deletions, and no inserted columns in this small procedure.

After every b_{batch} deletions, add the pending sum $\mathsf{U}\,\mathrm{diag}(\alpha_{\mathrm{aux}})\mathsf{U}^\top$ to the explicit base and clear the list. Here U has $O(D)$ rows and at most b_{batch} columns. The bound $\omega(1,a,1) \le 2.054999 < 1 + 2a$ from Alman et al. [ADW+25, Table 1] gives cost $O(D^{1+2a})$ per such product. A final incomplete list need not be written to the base after the stopping condition is met. Therefore the total arithmetic cost of the small procedure is

$$\widetilde{O}(D^{5/2} + D^2 b_{\mathrm{batch}} + (D/b_{\mathrm{batch}})D^{1+2a} + D^2) = \widetilde{O}(D^{2+a}), \qquad 2 + a = 2.5275.$$

The stability argument in Section 7.2 gives the same bit bound and preserves the residual, cube, and stopping guarantees.

We now organize the full input into a balanced binary tree whose leaves are the indices $1, \dots, m$. Each internal node splits its leaves into two sets with sizes differing by at most one. Precompute the vector $W_{\mathcal{J}} := \sum_{i \in \mathcal{J}} b_i$ at every node $\mathcal{J}$ by additions. There are fewer than $2m$ nodes, so this costs $O(mD)$ operations. The common denominator implies that every sum has only $O(\log(2mD/\eta))$ bits per entry. Initially the root is the only active group and its coefficient is zero. A group coefficient is shared by all leaves of that group. A group at coefficient 1 or -1 is frozen permanently.

Let $L_{\mathrm{tree}} := \max\{1, \lceil \log_2 m \rceil\}$ and choose a power of two M_{grp} with $m \le M_{\mathrm{grp}} < 2m$. At each of at most L_{tree} rounds, replace each active nonsingleton group by its two children, with both children inheriting the parent coefficient; retain active singletons. Since

$$y_{\mathrm{aux}_{\mathcal{J}}} W_{\mathcal{J}} = y_{\mathrm{aux}_{\mathcal{J}}} W_{\mathcal{J}_1} + y_{\mathrm{aux}_{\mathcal{J}}} W_{\mathcal{J}_2},$$

splitting makes no change to the signed sum. There are at most $2D$ groups after splitting, provided there were at most D before it. Apply the small procedure to the columns $W_{\mathcal{J}}/M_{\mathrm{grp}}$, whose norms are at most two, with $\varepsilon := \eta/(M_{\mathrm{grp}} L_{\mathrm{tree}})$ and the inherited coefficients. Freeze every group that reaches a boundary. At most D active groups remain, and the change in the original signed sum is at most $M_{\mathrm{grp}}\varepsilon = \eta/L_{\mathrm{tree}}$. If there are already at most D groups, this rounding step is empty.

The root satisfies the required group count. After L_{tree} rounds every remaining active group is a singleton, because its size is at most $\lceil m/2^{L_{\mathrm{tree}}} \rceil = 1$. Hence at most D original coordinates are fractional. Starting from zero and summing the residual allowances over the rounds gives the asserted bound η. Store each coefficient only at its current tree node; a final tree traversal assigns coefficients to all leaves in $O(m)$ operations. No update scans the leaves of an active group. The total cost is

$$\widetilde{O}(mD + L_{\mathrm{tree}} D^{2.5275}) = \widetilde{O}(mD + D^{2.5275}).$$

All column sums, normalized columns, and coefficients have $O(\log(2mD/\eta))$ bits with a sufficiently large absolute multiplier. This proves both complexity bounds. □

For the matrix inputs, use an orthonormal real basis of the Hermitian matrices and approximate the coordinate vector of each C_i to Euclidean error at most $\eta_{\mathrm{ps}}/(4m)$ on a common dyadic grid. Since $\|C_i\|_F \le 1$, these approximations have norm at most two and $O(\log(2mn))$-bit entries. Rank one and the variance computation above give total cost $\widetilde{O}(mn^2 + n^3)$ for scaling and vectorization. Apply Lemma 5.2 with $\eta := \eta_{\mathrm{ps}}/2$ and $d_{\mathrm{vec}} := n^2$. Because every output coefficient has absolute value at most one, the resulting coefficients for the true matrices satisfy

$$e_0 := \sum_i y_i C_i, \qquad \|e_0\|_F \le \eta_{\mathrm{ps}}.$$

The mathematical model below uses the true retained matrices. Only this single initial residual is charged to the approximate procedure; its error does not recur in the nonlinear phase. For exact partial signing, define $e_0 := 0$.

5.3 The maintained model and deferred rounding

For the reference lemmas, assume normalized PSD rank-one inputs C_i with $\sum_i C_i^2 \preceq I$ and at most $k \le n^2$ active coordinates. Set $\varepsilon_{\mathrm{scale}} := 10^{-8}$, $\rho := \varepsilon_{\mathrm{scale}}/n$, and $\eta_{\mathrm{ps}} := 10^{-7}$. The input conversion and centering for the higher-rank algorithms are specified in their main proofs. Define

$$\gamma := 10^{-6}/k, \qquad \varepsilon_{\mathrm{round}} := 10^{-6}.$$

Use the budget $\beta(y) = (1-y^2)^{97/200} f(|y|)$ from Section 4; $f \ge 1$, $0 \le \beta \le 1$, and $\iota = 14779/2500$. At endpoint distance at least $\delta_{\mathrm{end}}/2$,

$$|\beta^{(j)}(y)| \le C\delta_{\mathrm{end}}^{97/200-j} \quad (1 \le j \le 3), \qquad |\beta'(y)| \le C\delta_{\mathrm{end}}^{-1}\beta(y). \tag{134}$$

The relative derivative bound follows directly from the logarithmic derivative of $(1-y^2)^{97/200}$ and boundedness of f'/f. Third derivatives at zero are interpreted one-sided; β is C^2 and its second derivative is Lipschitz on every truncated interval.

There are two deliberately separated objects: the exact mathematical model, and rounded numerical data used to evaluate it. All discrepancy and descent inequalities refer to the exact model.

Let y^0 be the fractional coefficients from the chosen partial signing. Round them to a common fine dyadic grid, obtaining $\widetilde{y}^0 \in [-1,1]^k$ with each error at most δ_0, where $\delta_0\sqrt{k} \le \varepsilon_{\mathrm{round}}/2$. Initialize the model with

$$F := -\sum_{i=1}^{k} \widetilde{y}_i^0 C_i, \qquad M(y) := F + \sum_{i\in I} y_i C_i.$$

Thus $M(\widetilde{y}^0) = 0$. The actual previously frozen sum is $e_0 - \sum_i y_i^0 C_i$, so this initialization creates a fixed final discrepancy error of norm at most $\eta_{\mathrm{ps}} + \varepsilon_{\mathrm{round}}/2$. For exact partial signing $e_0 = 0$ and the η_{ps} term is omitted. Indeed

$$\|\sum_i r_i C_i\| \le \|\sum_i r_i C_i\|_F \le \|r\|_2. \tag{135}$$

For rank-one inputs, the adjoint inequality is $\sum_i \mathrm{tr}[C_i D_{\mathrm{aux}}]^2 \le \mathrm{tr}[{D_{\mathrm{aux}}}^2 \sum_i C_i^2] \le \|D_{\mathrm{aux}}\|_F^2$.

Define $L_{\log} := \lceil \log_2(4n) \rceil$ and choose a dyadic cutoff

$$\frac{{\varepsilon_{\mathrm{round}}}^2}{8192 L_{\log}} < \delta_{\mathrm{end}} \le \frac{{\varepsilon_{\mathrm{round}}}^2}{4096 L_{\log}}. \tag{136}$$

When a coordinate t_i has endpoint distance at most δ_{end}, update $F \leftarrow F + t_i C_i$, remove its coupling slot, and record t_i. Its output sign is chosen jointly with the other recorded coordinates at the end. This removal preserves the modeled sum and cannot increase the potential. Process these removals before testing regular endpoint moves. The continuation paths below remove the coupling while keeping the modeled sum fixed. A regular endpoint move fixes its sign immediately. Lemma 5.3 pays for the final joint rounding.

For the fast-budget estimates, the potential is $\mathcal{F} := \mathcal{P} + \gamma\mathcal{B}$, with $\mathcal{B}(y) := \sum_{i\in I} \beta(y_i)$ and

$$\mathcal{P}(y) := \min\{b + \rho\,\mathrm{tr}[U+V] : U^{-1} + M(y) + \mathcal{K}_y(V) \preceq bI, \quad V^{-1} - M(y) + \mathcal{K}_y(U) \preceq bI\},$$

$$\mathcal{K}_y(D) := \iota \sum_{i\in I} \beta(y_i) C_i D C_i.$$

We work in the original coordinates y. Write $A \lesssim B$ when $A \le P_n B$ for a fixed positive polynomial P_n in $\log(2n)$, enlarged when needed. Since $\delta_{\mathrm{end}}^{-1} = O(\log(2n))$, the explicit budget satisfies, throughout endpoint distance at least $\delta_{\mathrm{end}}/2$,

$$-\beta'' > 1/2, \qquad |\beta^{(j)}| \lesssim -\beta'', \qquad \frac{|\beta^{(j)}|^2}{\beta} \lesssim -\beta'' \quad (1 \le j \le 4). \tag{137}$$

At zero, the derivatives of orders three and four are interpreted one-sided. Derivatives through order two are continuous.

Coordinates close to an endpoint can be rounded together within the reserved discrepancy allowance, completing the partial signing.

Lemma 5.3 (Deferred rounding)**.** *Suppose $C_i \succeq 0$ have rank at most one, satisfy $\sum_i C_i^2 \preceq I$, and $|t_i| \ge 1 - \delta_{\mathrm{end}}$, where Eq. (136) holds. All the coefficients can be rounded deterministically to signs with*

$$\|\sum_i (\sigma_i - t_i) C_i\| < \varepsilon_{\mathrm{round}}/2.$$

For at most k inputs already available in a short PSD cache, the cost is $\widetilde{O}(kn^3)$ arithmetic and bit operations. The conclusion includes cache error and comparison error.

Proof. First work with short rational PSD rank-one matrices $\widehat{C}_i$ satisfying

$$\|\widehat{C}_i - C_i\|_{\mathrm{F}} \le u, \quad \sum_i \widehat{C}_i^2 \preceq 2I, \quad 2ku \le \varepsilon_{\mathrm{round}}/64.$$

The construction in Section 7.3 supplies this after making its fixed accuracy exponent sufficiently large. For example, $\sum_i \widehat{C}_i^2 \preceq (1 + 2ku + ku^2)I$. Define

$$\theta := 8L_{\log}/\varepsilon_{\mathrm{round}}, \qquad \tau := \frac{1}{4\theta} = \frac{\varepsilon_{\mathrm{round}}}{32L_{\log}}.$$

For every matrix with $\widehat{\nu}_i := \mathrm{tr}[\widehat{C}_i] > \tau$, choose the nearest sign. Rank one gives $\widehat{C}_i^2 = \widehat{\nu}_i \widehat{C}_i$, so the norm of the total error from these matrices is at most

$$\delta_{\mathrm{end}} \|\sum_{\widehat{\nu}_i > \tau} \widehat{C}_i\| \le 2\delta_{\mathrm{end}}/\tau \le \varepsilon_{\mathrm{round}}/64. \tag{138}$$

For $\widehat{\nu}_i \le \tau$, temporarily let $\epsilon_i \in \{-1, 1\}$ have mean t_i, and write $\mathsf{E}_i := (\epsilon_i - t_i)\widehat{C}_i$. Then

$$\mathbb{E}\mathsf{E}_i = 0, \quad \mathbb{E}\mathsf{E}_i^2 = (1 - t_i^2)\widehat{C}_i^2 \preceq 2\delta_{\mathrm{end}}\widehat{C}_i^2, \quad \|\theta\mathsf{E}_i\| \le 1/2.$$

For $|y| \le 1$, $e^y \le 1 + y + y^2$. Spectral calculus and scalar Jensen therefore give, for either sign,

$$0 \preceq G_i^{\pm} := \log \mathbb{E} e^{\pm\theta\mathsf{E}_i} \preceq 2\delta_{\mathrm{end}}\theta^2\widehat{C}_i^2, \qquad \|\sum_i G_i^{\pm}\| \le 4\delta_{\mathrm{end}}\theta^2 \le L_{\log}/16. \tag{139}$$

For a chosen prefix with error sum E, define the estimator

$$\mathcal{E}_{\mathrm{rnd}}(E) := \mathrm{tr}[\exp(\theta E + \sum_{\text{remaining}} G_i^{+})] + \mathrm{tr}[\exp(-\theta E + \sum_{\text{remaining}} G_i^{-})]. \tag{140}$$

Lieb's concavity theorem, in the form $\mathbb{E}\,\mathrm{tr}[e^{H+U}] \le \mathrm{tr}[e^{H+\log \mathbb{E} e^U}]$, implies that the weighted average of the two next estimators is at most the current one [Tro12, Corollary 3.3]. Thus one of the two signs does not increase it. Initially $\mathcal{E}_{\mathrm{rnd}0} \le 2ne^{L_{\log}/16}$. Choose the smaller estimator with additive evaluation error at most $1/(8\max\{1, k\})$. The total increase is at most $1/4$, hence the final estimator is at most $2\mathcal{E}_{\mathrm{rnd}0}$. At the end, the remaining sums vanish and

$$e^{\theta\|E\|} \le \mathrm{tr}[e^{\theta E}] + \mathrm{tr}[e^{-\theta E}] \le 4ne^{L_{\log}/16}.$$

Consequently

$$\|E\| \le \frac{\log(4n) + L_{\log}/16}{\theta} \le \frac{17\varepsilon_{\mathrm{round}}}{128}. \tag{141}$$

Combining (138), (141), and cache error gives

$$\|\sum_i (\sigma_i - t_i)C_i\| \le \frac{\varepsilon_{\mathrm{round}}}{64} + \frac{17\varepsilon_{\mathrm{round}}}{128} + \frac{\varepsilon_{\mathrm{round}}}{64} = \frac{21\varepsilon_{\mathrm{round}}}{128} < \varepsilon_{\mathrm{round}}/2.$$

It remains to justify the evaluation cost. Rank one gives the explicit expression

$$G_i^{\pm} = \frac{\log m_i^{\pm}}{\widehat{\nu}_i}\widehat{C}_i, \quad m_i^{\pm} := \frac{1 + t_i}{2}e^{\pm\theta(1-t_i)\widehat{\nu}_i} + \frac{1 - t_i}{2}e^{\mp\theta(1+t_i)\widehat{\nu}_i},$$

with zero matrices omitted. Store and subtract the two remaining sums rather than reconstructing them at every decision. At every accepted prefix, (139) and the estimator bound imply $\|\theta E\| = O(\log(2n))$. At either candidate this changes by at most $1/2$; both exponential arguments therefore have norm $O(\log(2n))$. Matrix Taylor series with scaling and squaring evaluate their traces to the required additive accuracy in a polylogarithmic number of matrix multiplications. Scalar exponentials, logarithms, and roots have the same accuracy overhead. Cache entries have $O(\log(2kn))$ bits; nonzero traces have inverse-polynomial magnitude, or the equivalent divided series can be used. Recomputing the estimator from the recorded prefix and controlled remaining sums prevents unaccounted accumulation. Ordinary multiplication gives the claimed $\widetilde{O}(kn^3)$ cost. □

5.4 KKT conditioning and weighted norms

Matrix spaces carry Frobenius norms and pairs carry product norms. Order inequalities for pairs are componentwise. Write $\mathsf{P} := (U, V)$, $\mathsf{Q} := (W, Z)$, and $E := (I, I)$ for the primal and dual pairs at the unique optimizer. The dual and tight-constraint identities give

$$\operatorname{tr}[W + Z] = 1, \quad W \succeq \rho U^2, \quad Z \succeq \rho V^2, \quad U, V \succeq \kappa_0 I, \quad \|U^{-1}\|, \|V^{-1}\| \le C,$$

$$\rho \operatorname{tr}[U^2 + V^2] \le 1.$$

Let

$$\mathscr{J}(\mathsf{A}_1, \mathsf{A}_2) := (U^{-1}\mathsf{A}_1 U^{-1} - \mathcal{K}_y(\mathsf{A}_2), V^{-1}\mathsf{A}_2 V^{-1} - \mathcal{K}_y(\mathsf{A}_1)).$$

It is self-adjoint and $\mathscr{J}^{-1}$ preserves the PSD order by Lemma 4.1. Also $\mathscr{J}\mathsf{Q} = \rho E$ and $\|\mathscr{J}^{-1}\| \le \rho^{-1}$: for $\|B\|_F \le 1$, positivity gives $-\mathsf{Q}/\rho \preceq \mathscr{J}^{-1}B \preceq \mathsf{Q}/\rho$, whose Frobenius norm is at most $\operatorname{tr}[W + Z]/\rho = \rho^{-1}$. Let $\mathscr{E}$ be the block diagonal Hessian of the Lagrangian in P; its U block is

$$\mathscr{E}_U(\mathsf{A}) := U^{-1}\mathsf{A}U^{-1}WU^{-1} + U^{-1}WU^{-1}\mathsf{A}U^{-1}.$$

In particular $a\rho^2 I \preceq \mathscr{E} \preceq CI$ and $\|\mathscr{J}\| \le C$. For the lower bound, $W \succeq \rho U^2$ implies

$$\operatorname{tr}[WU^{-1}\mathsf{A}U^{-1}\mathsf{A}U^{-1}] \ge \rho \operatorname{tr}[\mathsf{A}U^{-1}\mathsf{A}] \ge \rho^{3/2}\|\mathsf{A}\|_F^2.$$

The V block is identical, and $\rho < 1$ gives the stated weaker bound. The full KKT Jacobian satisfies

$$\|\mathsf{R}^{-1}\| \le C\rho^{-2} = Cn^2. \tag{142}$$

For completeness, consider the full KKT inverse, including the scalar normalization. Write $E := (I, I)$. Its linear system has the form

$$\mathscr{J}\mathsf{A} + d_{\mathrm{b}}E = F_1, \qquad \mathscr{J}\mathsf{B} - \mathscr{E}\mathsf{A} = F_2, \qquad \langle E, \mathsf{B}\rangle = f_3.$$

Define $\mathsf{f} := \mathscr{J}^{-1}E$, $\mathsf{g}_0 := \mathscr{J}^{-1}F_1$, and $D_0 := \langle \mathsf{f}, \mathscr{E}\mathsf{f}\rangle$. Elimination gives

$$d_{\mathrm{b}} = \frac{\langle \mathsf{f}, \mathscr{E}\mathsf{g}_0\rangle + \langle \mathsf{f}, F_2\rangle - f_3}{D_0}, \qquad \mathsf{A} = \mathsf{g}_0 - \mathsf{f}d_{\mathrm{b}}, \qquad \mathsf{B} = \mathscr{J}^{-1}(\mathscr{E}\mathsf{A} + F_2).$$

Use the following projection bounds before estimating these expressions:

$$\|I - \mathsf{f}\mathsf{f}^*\mathscr{E}/D_0\| \le \sqrt{\|\mathscr{E}\|/\lambda_{\min}(\mathscr{E})} = O(\rho^{-1}),$$

$$0 \preceq \mathscr{E} - \mathscr{E}\mathsf{f}\mathsf{f}^*\mathscr{E}/D_0 \preceq \mathscr{E},$$
$$\|\mathsf{f}\|^2/D_0 = O(\rho^{-2}), \qquad \|\mathscr{E}\mathsf{f}\|\|\mathsf{f}\|/D_0 = O(\rho^{-1}), \qquad \|\mathsf{f}\| \geq \|E\|/\|\mathscr{J}\| \geq a\sqrt{n}.$$

Applying these separately to the F_1, F_2, f_3 terms gives

$$\|\mathsf{A}\| + \|\mathsf{B}\| + |d_{\mathrm{b}}| \leq C\rho^{-2}(\|F_1\| + \|F_2\| + |f_3|).$$

For a Hermitian pair $D := (D_U, D_V)$, write

$$\widehat{D}_U := U^{-1/2} D_U U^{-1/2}, \qquad T_U := U^{-1/2} W U^{-1/2},$$

and use the analogous notation for the V block. Define

$$e(D)^2 := \langle D, \mathscr{E} D\rangle = 2\,\mathrm{tr}[T_U \widehat{D}_U^2] + 2\,\mathrm{tr}[T_V \widehat{D}_V^2].$$

Sums over the two blocks below are denoted by $\sum_{A=U,V}$. We have $\sum_A \mathrm{tr}[T_A] \leq C$, $T_U \succeq \kappa_0 \rho I$, and $\mathscr{E} \succeq 2\rho^{3/2} I$. Consequently

$$\|D\|_F \leq C n^{3/4} e(D), \qquad \max_A \|\widehat{D}_A\| \leq C\sqrt{n} e(D), \qquad e(D) \leq C\|D\|_F. \tag{143}$$

The second inequality follows by bounding $\mathrm{tr}[\widehat{D}_A^2]$ using $T_A \succeq \kappa_0 \rho I$. We also have $\|\widehat{D}_A\| \leq C\|D_A\|_F$.

For a Hermitian pair B, define the weighted order bound

$$q(B) := \inf\{\langle \mathsf{Q}, A\rangle : A \succeq 0,\ -A \preceq B \preceq A\}.$$

It is subadditive and homogeneous. Positivity and self-adjointness of $\mathscr{J}^{-1}$ give

$$\|\mathscr{J}^{-1} B\|_F \leq \rho^{-1} q(B), \qquad e(\mathscr{J}^{-1} B) \leq C n q(B). \tag{144}$$

Indeed, for any admissible A, define $C_0 := \mathscr{J}^{-1} A \succeq 0$. Then $-C_0 \preceq \mathscr{J}^{-1} B \preceq C_0$, and the sum of the nuclear norms of the two blocks is at most $\langle E, C_0\rangle = \rho^{-1}\langle \mathsf{Q}, A\rangle$. For this nuclear-norm inequality, pair the two spectral projections of $\mathscr{J}^{-1} B$ with its upper and lower bounds. Taking the infimum proves the first assertion; $\mathscr{E} \preceq CI$ proves the second.

Define $\mathcal{A}(D_U, D_V) := (\mathcal{K}_y(D_V), \mathcal{K}_y(D_U))$. For the weighted maps below, $\mathcal{A}_t^{(j)}$ denotes a derivative along any path whose coupling coefficients have bounded relative derivatives. Along a reset path this bound is absolute; along a local y segment it is at most P_n. In the following display, replace C by P_n in the latter case.

Two useful consequences are

$$q(\mathcal{A}_t^{(j)} D) \leq C e(D), \qquad q(\mathscr{E} D) \leq C e(D). \tag{145}$$

For the first, define $D_U^{\#} := U^{1/2} |\widehat{D}_U| U^{1/2}$, and similarly for V. Then $\pm \mathcal{A}_t^{(j)} D \preceq C \mathcal{A}_t(D^{\#})$. The dual equation implies $\mathcal{A}_t \mathsf{Q} \preceq \mathsf{S}$, where $\mathsf{S} := (U^{-1} W U^{-1}, V^{-1} Z V^{-1})$. Hence its weighted cost is at most

$$C\langle \mathsf{S}, D^{\#}\rangle = C \sum_A \mathrm{tr}[T_A |\widehat{D}_A|] \leq C \sqrt{\sum_A \mathrm{tr}[T_A]} \sqrt{\sum_A \mathrm{tr}[T_A \widehat{D}_A^2]} \leq C e(D).$$

For the second, the normalized U block of $\mathscr{E} D$ is $\widehat{D}_U T_U + T_U \widehat{D}_U$. For any $a > 0$, its positive and negative are bounded by $a\widehat{D}_U^2 + a^{-1} T_U^2$. The total weighted cost is at most $a e(D)^2/2 + a^{-1} \sum_A \mathrm{tr}[T_A^3]$. Optimizing a and using $\sum_A \mathrm{tr}[T_A^3] \leq C$ proves the claim. This argument does not require the matrices to commute.

There is also an exact description of the order norm:

$$q(B) = \sum_{A=U,V} \|\mathsf{Q}_A^{1/2} B_A \mathsf{Q}_A^{1/2}\|_*. \tag{146}$$

Congruence reduces its defining minimization to the trace of a positive majorant of a Hermitian matrix and its negative, whose minimum is the nuclear norm. For an inverse-derivative block $U^{-1/2} D U^{-1/2}$ this equals $\|T_U^{1/2} D T_U^{1/2}\|_*$, by polar decomposition. This also justifies the noncommuting products below. Define $\mathsf{f} := \mathscr{J}^{-1} E = \mathsf{Q}/\rho$ and $D_0 := e(\mathsf{f})^2$. Then

$$\|\mathsf{f}\|_F \le Cn, \qquad D_0 = 2\rho^{-2} \sum_{A=U,V} \operatorname{tr}[T_A^3] \ge \sqrt{2/\varepsilon_{\text{scale}}}. \tag{147}$$

Indeed $T_U \succeq \rho U$ implies $\operatorname{tr}[T_U^2] \ge \rho \operatorname{tr}[T_U U] = \rho \operatorname{tr}[W]$. Summing both blocks and applying the power-mean inequality on $2n$ eigenvalues gives $\sum_A \operatorname{tr}[T_A^3] \ge \rho^{3/2}/\sqrt{2n}$.

On a fixed-radius neighborhood in the optimizer variables, the KKT map $\mathcal{E}_{\mathrm{KKT}}(y, \mathsf{s})$ has $\|\mathcal{E}_{\mathrm{KKT}\mathsf{s}}\| \le C$: inverse matrices, W, Z, and the coupling have bounded operator norms. Hence the inverse bound in Eq. (142) gives a Newton radius $r_m := \kappa_m n^{-2}$. Inside this radius,

$$\|\mathsf{s}_{j+1} - \mathsf{s}_*\| \le Cn^2 \|\mathsf{s}_j - \mathsf{s}_*\|^2. \tag{148}$$

Inexact evaluation adds a polynomially controlled error floor.

Write

$$\mathcal{C}(y, \mathsf{P}) := (U^{-1} + M + \mathcal{K}_y(V), V^{-1} - M + \mathcal{K}_y(U)), \qquad \mathcal{L} := b + \rho\langle E, \mathsf{P}\rangle + \langle \mathsf{Q}, \mathcal{C} - bE\rangle.$$

The optimizer satisfies $\mathcal{C} = bE$, $\mathscr{J}\mathsf{Q} = \rho E$, and $\operatorname{tr}[W + Z] = 1$. The symbol $\mathsf{s} := (b, \mathsf{P}, \mathsf{Q})$ denotes all optimizer variables. The matrix $\mathscr{E}$ is the Hessian of $\mathcal{L}$ in P.

Write $C_i = a_i a_i^*$ and

$$u_i := a_i^* U a_i, \quad v_i := a_i^* V a_i, \quad w_i := a_i^* W a_i, \quad z_i := a_i^* Z a_i, \quad t_i := v_i w_i + u_i z_i.$$

There is no numerical choice of phases for these vectors: the expressions can all be computed as matrix traces. Define

$$\varpi_i := (-\beta_i'')(\iota t_i + \gamma), \qquad G := \operatorname{diag}(\varpi_i). \tag{149}$$

At a light state, $u_i, v_i \le 2/(\iota\beta_i)$ and $w_i + z_i \le \operatorname{tr}[C_i] \le 1$. Consequently

$$\gamma/4 \le \varpi_i \le P_n. \tag{150}$$

For a Euclidean unit direction r, define

$$A_r := \iota \sum_i (-\beta_i'') t_i r_i^2, \qquad \theta := r^\top G r = A_r + \gamma \sum_i (-\beta_i'') r_i^2. \tag{151}$$

In particular $\theta \ge \gamma/4 \ge a/n^2$, where $a > 0$ is absolute.

5.5 A budget-weighted gradient/curvature dichotomy

The following dichotomy selects the next type of local move using thresholds scaled by the coordinate budgets.

Lemma 5.4. *At every light state, either some coordinate has*

$$|\mathcal{F}_i| \geq P_n^{-1}\sqrt{\varpi_i/n}, \tag{152}$$

or there is a Euclidean unit direction r with

$$r^\top \nabla^2 \mathcal{F} r \leq -P_n^{-1} r^\top G r. \tag{153}$$

Both alternatives can be implemented with fixed-factor separated numerical thresholds and inverse-polynomial margins.

Proof. Use the pointwise drift $h_i = -C_{2,i}M_iK_{ii}$, where $M_i := u_i - v_i + 2\iota\beta_i' u_i v_i$ and $C_{2,i} = (1 + \tau/2)k_i$. Lemma 4.4, the strict surplus in Eq. (106), and the budget descent estimate give

$$\mathscr{L}\mathcal{P} \leq -a\sum_i p_i q_i t_i K_{ii}, \qquad \mathscr{L}\mathcal{B} \leq -\frac{1}{20}\sum_i (-\beta_i'')K_{ii}.$$

Here $a > 0$ is absolute, although small. Since $p_i q_i \geq 1/2$ and $-\beta'' \leq P_n$,

$$\mathscr{L}\mathcal{F} \leq -P_n^{-1}\operatorname{tr}[GK]. \tag{154}$$

On the other hand lightness implies $|M_i| \lesssim u_i + v_i$. Define $\nu_i := \operatorname{tr}[C_i]$. Rank-one Cauchy–Schwarz and the dual lower bounds give

$$w_i \geq \rho u_i^2/\nu_i, \quad z_i \geq \rho v_i^2/\nu_i, \quad u_i, v_i \geq \kappa_0 \nu_i.$$

Therefore $t_i \geq \kappa_0\rho(u_i^2 + v_i^2)$, and hence

$$(C_{2,i}M_i)^2 \lesssim n\varpi_i. \tag{155}$$

If every $|\mathcal{F}_i|$ is smaller than a sufficiently small inverse-polylogarithmic multiple of $\sqrt{\varpi_i/n}$, equations (154)–(155) imply

$$\operatorname{tr}[\nabla^2 \mathcal{F} K] \leq -P_n^{-1}\operatorname{tr}[GK].$$

The covariance is nonzero. Conjugating by $G^{-1/2}$ proves the second alternative. Otherwise the first holds. Approximate ϖ_i to relative constant accuracy, with a tighter internal floor, using $\varpi_i \geq \gamma/4$. Compute a negative-curvature witness for the symmetric scaled Hessian $G^{-1/2}\nabla^2\mathcal{F}G^{-1/2}$, then normalize $G^{-1/2}\mathsf{h}$ in Euclidean norm and verify the resulting generalized Rayleigh quotient. Its norm, inverse accuracy, and all diagonal scalings are polynomially bounded. The curvature routine in Section 5.9, with the error bound in Section 7.9, applies. □

5.6 Directional estimates retaining the size of the budget

Derivatives in this section are along the straight original-coordinate line $y + tr$. Work at endpoint distance at least $\delta_{\mathrm{end}}/2$. Define $\mathsf{A} := \mathsf{P}'$, $\mathsf{B} := \mathsf{Q}'$, and $F_2 := \mathcal{A}'\mathsf{Q}$.

The next estimates retain the directional budget scale, which controls the optimizer response and the length of a local step.

Lemma 5.5. *For $1 \leq j \leq 4$,*

$$q(\mathcal{A}^{(j)}D) \lesssim \sqrt{A_r}\,e(D), \qquad |\langle F_2, D\rangle| \lesssim \sqrt{A_r}\,e(D). \tag{156}$$

For $j \geq 2$, $q(\mathcal{C}_{tj}) \lesssim A_r$. The second-order envelope identity is

$$\mathcal{P}'' = -A_r + e(\mathsf{A})^2 + 2\langle F_2, \mathsf{A}\rangle. \tag{157}$$

Consequently, in a region where $\mathcal{P}'' \leq C\theta_0$ and $\theta \asymp \theta_0$, one has

$$e(\mathsf{A}) \lesssim \sqrt{\theta_0}, \quad \|\mathsf{A}\|_{\mathrm{F}} \lesssim n^{3/4}\sqrt{\theta_0}, \quad \|\mathsf{B}\|_{\mathrm{F}} \lesssim n\sqrt{\theta_0}. \tag{158}$$

Proof. For one coupling block, use

$$|a_i^* D_V a_i|^2 \le v_i a_i^* D_V V^{-1} D_V a_i.$$

Weighted scalar Cauchy–Schwarz bounds its majorant cost by the product of

$$(\iota \sum_i \frac{|\beta_i^{(j)}|^2}{\beta_i} |r_i|^{2j} v_i w_i)^{1/2} \quad \text{and} \quad (\iota \sum_i \beta_i w_i \frac{|a_i^* D_V a_i|^2}{v_i})^{1/2}.$$

The first squared factor is at most $P_n A_r$. The second is at most $\operatorname{tr}[D_V V^{-1} D_V \mathcal{K}(W)] \le e(D_V)^2/2$, by the dual equation. Combine both blocks. This proves (156). For direct partial derivatives use $M^{(j)} = 0$ for $j \ge 2$ and (137); their majorant costs are at most $P_n A_r$. Differentiating the stationary Lagrangian gives (157) because $\langle \mathsf{Q}, \mathcal{C}_{tt} \rangle = -A_r$. It gives $\mathcal{P}'' \ge e(\mathsf{A})^2 - P_n \sqrt{A_r} e(\mathsf{A}) - A_r$, proving the energy estimate. The remaining estimates follow from Eqs. (143)–(145), $\mathsf{B} = \mathscr{J}^{-1}(\mathscr{E}\mathsf{A} + F_2)$, and the dual energy bound for F_2. □

5.7 Rank-one operator actions

The rank-one representation reduces coupling evaluations and their directional derivatives to rectangular matrix products.

Lemma 5.6. *For a fixed number of Hermitian arguments, the coupling and its fixed-order directional derivatives can be applied in $\widetilde{O}(n^\chi)$ arithmetic and bit operations. The same bound computes all endpoint scores, gradient coordinates, and weights ϖ_i at a given optimizer.*

Proof. Use the stored factors $\widehat{C}_i = \gamma_i b_i b_i^*$ from the short-data construction, with $\gamma_i > 0$ and $\|b_i\|_\infty \le 2$; zero inputs may be omitted. No square root of γ_i is required. Put all $k \le n^2$ columns in C and pad to an n-by-n^2 matrix. For an argument D_{aux}, compute $D_{\mathrm{aux}}\mathsf{C}$, and then its column inner products with C. This gives every $b_i^* D_{\mathrm{aux}} b_i$. A coupling with scalar weights κ_i is

$$\mathsf{C} \operatorname{diag}(\kappa_i \gamma_i^2 (b_i^* D_{\mathrm{aux}} b_i)) \mathsf{C}^*.$$

The two products have shapes (n, n, n^2) and (n, n^2, n). Permutation symmetry of the matrix multiplication tensor gives cost $\widetilde{O}(n^\chi)$ for each. The column inner products and diagonal scalings cost $O(nk) \le O(n^3)$. The same products compute all quadratic forms defining u_i, v_i, w_i, z_i in the potential notation. Directional differentiation changes the weights to $\iota \beta_i^{(j)} r_i^j$, so a fixed derivative order has the same cost. These products also form signed sums $\sum_i \kappa_i \widehat{C}_i$. Every fixed-order differentiated KKT right-hand side is a fixed sum of these expressions and matrix products, so its assembly has the same bound. The primal inverse-congruence and Lagrangian-Hessian maps require a fixed number of order-n products, whose cost is smaller.

The stable implementation and bit cost of these actions are proved in Section 7.4.

□

5.8 Regular endpoint moves

The next lemma bounds regular endpoint moves. Its rank-one feasibility adjustment is also used for the smoothed potential in Section 6.5.

Lemma 5.7. *There are $\widetilde{O}(n)$ regular endpoint moves.*

Proof. Process near-endpoint coordinates first. For a regular move to 1, define $C_i = a_i a_i^*$, $\nu_i := \operatorname{tr}[C_i] \le 1$, and $d := 1 - y_i \ge \delta_{\mathrm{end}}$. At an exact trigger, $\iota\beta_i v_i \ge d$. After changing M by dC_i and deleting this coupling slot, the two slacks at the old optimizer are $(\iota\beta_i v_i - d)C_i$ and $(\iota\beta_i u_i + d)C_i$. Define $a_0 := \iota\beta_i u_i + d$ and replace V by

$$V_+ := (V^{-1} + a_0 C_i)^{-1}.$$

Keep b, U fixed. The second constraint becomes tight, and the first remains feasible because $V_+ \preceq V$ and the remaining coupling preserves PSD order. Sherman–Morrison and Cauchy–Schwarz give

$$\mathcal{P}_{\mathrm{old}} - \mathcal{P}_{\mathrm{new}} \ge \rho \operatorname{tr}[V - V_+] = \rho \frac{a_0 a_i^* V^2 a_i}{1 + a_0 v_i} \ge \rho \frac{a_0 v_i^2}{\nu_i (1 + a_0 v_i)}. \tag{159}$$

The last fraction increases with a_0, v_i and decreases with ν_i.

For an approximate trigger of absolute error at most δ, the exact score is at least -2δ. Take $\delta \le \delta_{\mathrm{end}}/4$. Then $v_i \ge \delta_{\mathrm{end}}/(2\iota)$, since $\beta_i \le 1$. The feasibility repair increases b by at most $2\delta\nu_i$. Using $a_0 \ge \delta_{\mathrm{end}}$ and $\delta_{\mathrm{end}} \le 1$, the gain in Eq. (159), before that repair, is at least

$$\frac{\rho \delta_{\mathrm{end}}^3}{4\iota^2 + 2\iota}.$$

Choose δ smaller than a fixed fraction of this quantity. The repaired move decreases $\mathcal{P}$ by $a\rho\delta_{\mathrm{end}}^3$. The move to -1 is identical with U, V interchanged. The reservoir also decreases. Since the nonnegative potential starts at an absolute constant, and $\delta_{\mathrm{end}}^{-1} = O(\log(2n))$, there are $\widetilde{O}(\rho^{-1}) = \widetilde{O}(n)$ such moves. All additional comparison margins are inverse polynomial. □

5.9 Arithmetic cost per stage

We give the matrix algorithms and count their arithmetic operations to the inverse-polynomial accuracies required above. Section 7.3 proves that finite-precision evaluation preserves these accuracies.

Inversion by Newton–Schulz iteration. Here is an explicit substitute for rational elimination with growing intermediate denominators. Suppose a real square matrix A has polynomially bounded $\|A\|$ and $\|A^{-1}\|$. This applies to the KKT Jacobian throughout its Newton neighborhood by Eq. (142) and the Newton-neighborhood bound above. Choose a polynomial upper bound $L \ge \|A\|$, and initialize

$$\mathsf{Z}_0 := A^\top / L^2, \qquad \mathsf{Z}_{r_{\mathrm{aux}}+1} := \mathsf{Z}_{r_{\mathrm{aux}}}(2I - A\mathsf{Z}_{r_{\mathrm{aux}}}).$$

With $E_{r_{\mathrm{aux}}} := I - A\mathsf{Z}_{r_{\mathrm{aux}}}$, exact arithmetic gives $E_{r_{\mathrm{aux}}+1} = E_{r_{\mathrm{aux}}}^2$ and

$$\|E_0\| \le 1 - \frac{1}{L^2 \|A^{-1}\|^2}.$$

Thus an inverse-polynomial residual is reached in $O(\log(2kn))$ matrix multiplications. This works for nonsymmetric and indefinite matrices; it does not rely on positive eigenvalues of A.

For order $d_{\mathrm{sys}} = O(n^2)$, these products and multiplication by the right-hand sides cost $\widetilde{O}(n^{2\omega_0})$ arithmetic operations. The same construction handles the order-n primal inverses. Section 7.4 proves the residual bound for rounded iteration and the corresponding bit cost.

A curvature witness by normalized matrix squaring. This construction avoids a cubic-time Jacobi sweep and requires no separation between adjacent eigenvalues. Let A approximate the symmetric scaled Hessian $G^{-1/2}\nabla^2\mathcal{F}G^{-1/2}$. Choose a known polynomial bound $L \geq \max\{1, \|A\|\}$, and let δ be a sufficiently small inverse-polylogarithmic threshold. If the gradient test fails, the existing error margins ensure $\lambda_{\min}(A) \leq -2\delta$. Define

$$\mathsf{T}_0 := I - A/(2L), \qquad D_{\mathrm{pow}_0} := \mathsf{T}_0/\|\mathsf{T}_0\|_F, \qquad D_{\mathrm{pow}_{j+1}} := D_{\mathrm{pow}_j}^2/\|D_{\mathrm{pow}_j}^2\|_F.$$

In exact arithmetic $D_{\mathrm{pow}_{r_{\mathrm{aux}}}} = \mathsf{T}_0^q/\|\mathsf{T}_0^q\|_F$, where $q := 2^{r_{\mathrm{aux}}}$. Choose q as the least power of two satisfying

$$q \geq (4L/\delta)\log(8kL/\delta), \qquad r_{\mathrm{aux}} = O(\log(2kn)).$$

An eigenvalue of A at most -2δ gives an eigenvalue of T_0 at least $1+\delta/L$. For eigenvalues of A at least $-\delta$, the corresponding ratio is at most

$$\frac{1+\delta/(2L)}{1+\delta/L} \leq 1 - \frac{\delta}{4L}.$$

The total squared Frobenius weight of these latter eigenvectors in $D_{\mathrm{pow}_{r_{\mathrm{aux}}}}$ is therefore at most

$$k\exp(-q\delta/(2L)) \leq \delta/(8L).$$

It follows that

$$\mathrm{tr}[D_{\mathrm{pow}_{r_{\mathrm{aux}}}} A D_{\mathrm{pow}_{r_{\mathrm{aux}}}}] \leq -3\delta/4, \qquad \|D_{\mathrm{pow}_{r_{\mathrm{aux}}}}\|_F = 1.$$

Writing this trace as the sum of the column quadratic forms shows that a column has negative Rayleigh quotient. More quantitatively, discard columns of squared norm less than $\delta/(16kL)$; their total contribution in absolute value is at most $\delta/16$. Some retained column has Rayleigh quotient at most $-\delta/2$. Compute all column quadratic forms using one product $AD_{\mathrm{pow}_{r_{\mathrm{aux}}}}$, then scan them. This costs $\widetilde{O}(k^{\omega_0}) \leq \widetilde{O}(n^{2\omega_0})$.

Section 7.9 bounds the error of rounded normalized squaring and verifies a final Rayleigh quotient at most $-\delta/4$. The same arithmetic cost therefore suffices for a certified witness.

5.10 Trace-corrected local estimates and deterministic responses

We retain the local estimates and response constructions needed by the higher-rank deterministic algorithm. Use $\beta_{\mathrm{tr}}, \iota_{\mathrm{tr}}$ from Section 4.8, abbreviated by β, ι here, and the direct mixed covariance of Lemma 4.5. The two gradient corrections in Eq. (120) are not proportional to the covariance diagonal. We first give a direct implementation using the strict descent of $\mathcal{P}$. Entropy smoothing then improves its running time without changing the covariance certificate or constant.

Preprocessing and the maintained model. Assume the reference inputs have already been normalized and reduced by grouped short-precision partial signing. At most $k \leq n^2$ coordinates remain fractional and $\sum_i C_i^2 \preceq I$. Keep $\varepsilon_{\mathrm{scale}} := 10^{-8}$, $\rho := \varepsilon_{\mathrm{scale}}/n$, and $\varepsilon_{\mathrm{round}} = \varepsilon_{\mathrm{discard}} = \chi = e_{\mathrm{end}} := 10^{-6}$. Round the initial coefficients with Euclidean error at most $\varepsilon_{\mathrm{round}}/2$ and initialize the modeled sum at zero as before. Small-input rounding is applied in phases, as specified below. It preserves the modeled sum and uses the same total allowance χ. Write $t_i := \mathrm{tr}[C_i]$; throughout, $t_i \leq 1$. Choose δ_{end} by Eq. (136). At endpoint distance at most δ_{end}, record the fractional coefficient, preserve the current modeled contribution in F, and remove the coupling slot. At the end, apply Lemma 5.3 jointly to all recorded coefficients. This lemma uses only the variance

bound and endpoint distances, so it applies to the present budget and pure matrix potential. It is deterministic and costs less than $\varepsilon_{\mathrm{round}}/2$ in discrepancy. Together with initial rounding, the allowance remains $\varepsilon_{\mathrm{round}}$. Now $\delta_{\mathrm{end}}^{-1} = O(\log(2n))$; in particular, $\delta_{\mathrm{end}} \geq a/n$. The initial matrix potential is at most $2\sqrt{\iota + 2\varepsilon_{\mathrm{scale}}}$.

Lemma 5.8. *Let $k \leq n^2$, let $C_i \succeq 0$ have rank at most one, and let $y_i \in [-1,1]$ be the initial dyadic coefficients. Choose a common dyadic PSD rank-one cache with $\|\widehat{C}_i - C_i\|_F \leq u$, where $u \leq \chi/(64n^2)$, and define*

$$\tau_{\mathrm{small}} := \frac{\chi}{4n}, \qquad S := \{i : \mathrm{tr}[\widehat{C}_i] \leq \tau_{\mathrm{small}}\}.$$

There is a deterministic choice of signs on S such that

$$\|\sum_{i\in S}(\sigma_i - y_i)C_i\| < \chi/2.$$

Every remaining matrix satisfies $\mathrm{tr}[C_i] \geq \chi/(8n)$. For cache entries and coefficients of logarithmic binary length, the rounding costs $O(kn^2)$ arithmetic operations and $\widetilde{O}(kn^2)$ bit operations.

Proof. For independent signs with mean y_i, centering and rank one give

$$\mathbb{E}[\|\sum_{i\in S}(\sigma_i - y_i)\widehat{C}_i\|_F^2] = \sum_{i\in S}(1-y_i^2)\|\widehat{C}_i\|_F^2 \leq k\tau_{\mathrm{small}}^2.$$

Derandomize this identity by conditional expectation. Starting with $E = 0$, at each coordinate choose the sign minimizing $\|E + (\sigma_i - y_i)\widehat{C}_i\|_F^2$, and add its contribution to E. The two possible conditional expectations have weighted average equal to the current value, so choosing the smaller cannot increase it. The final cache error is at most $\sqrt{k}\tau_{\mathrm{small}} \leq \chi/4$ in Frobenius norm. Consequently

$$\|\sum_{i\in S}(\sigma_i - y_i)C_i\| \leq \chi/4 + 2ku \leq 9\chi/32 < \chi/2.$$

If $i \notin S$, then $\mathrm{tr}[C_i] > \tau_{\mathrm{small}} - nu \geq 15\chi/(64n) > \chi/(8n)$. The difference between the squared norms for signs 1 and -1 is $4(\mathrm{tr}[E\widehat{C}_i] - y_i\,\mathrm{tr}[\widehat{C}_i]^2)$. Each comparison and update therefore costs $O(n^2)$ operations. All comparisons are exact on the common dyadic grid. Summing at most k terms adds only $O(\log(2k))$ bits, which proves the bit bound. The PSD cache construction in Section 7.3 permits all its factors to be dyadic and supplies the stated error at this precision. □

Phases for small-input rounding. Set $L := 1 + \lceil \log_2(2n^2) \rceil$. At the start of a phase let K be the current number of active coordinates. Apply Lemma 5.8 with cutoff

$$\tau_K := \frac{\chi}{4L\sqrt{K}}, \qquad u \leq \frac{\chi}{64Ln^2}.$$

The cutoff can be rounded downward by a factor of at most two. Use one common cache of this accuracy throughout the algorithm. The same conditional-expectation calculation gives a discrepancy error of at most $\chi/(4L) + 2Ku < \chi/(2L)$ in this phase. Every retained trace is at least $\chi/(16L\sqrt{K})$: indeed the trace error is at most $nu \leq \chi/(64Ln)$ and $K \leq n^2$. Keep the removed fractional contributions in F, so this operation preserves the modeled matrix and only deletes nonnegative coupling terms.

Start another phase whenever the active count becomes at most $K/2$. If small-input rounding itself causes this, repeat before taking a local step. There are at most L phases, so the total

discrepancy error is less than $\chi/2$, as before. During local steps of a phase, $K/2 < k \le K$ and consequently

$$t_i \ge \frac{a}{L\sqrt{k}}, \qquad \varpi_i \ge \frac{a}{L^2 nk}. \tag{160}$$

The second bound follows from $u_i, v_i \ge at_i$, $w_i, z_i \ge at_i/n$, and $-\beta_i'' \ge a$. In particular the earlier lower bound becomes $\varpi_i \ge aP_n^{-1}n^{-3}$, which suffices below. Each phase costs $O(Kn^2)$ operations. The starting counts decrease geometrically, so all such rounding costs $\widetilde{O}(n^4)$ in both models. The signs are chosen by exact dyadic comparisons; these phases introduce no random bits.

Quantitative drift witnesses. In this paragraph, $X \lesssim Y$ means $X \le P_n Y$ for a fixed positive polynomial in $\log(2n)$; $X \gtrsim Y$ has the reverse meaning. Choose a single sufficiently large polynomial P_n to cover all such bounds below, enlarging it when finitely many factors are combined. On a bounded-potential region, the same KKT estimates give

$$aI \preceq U, V \preceq C\sqrt{n}I, \qquad an^{-1}I \preceq W, Z \preceq I.$$

Together with the trace cutoff, the budget bounds, and lightness, these imply

$$\begin{aligned} at_i \le u_i, v_i \le C\sqrt{n}t_i, \qquad & an^{-1}t_i \le w_i, z_i \le t_i, \\ \beta_i \gtrsim 1, \qquad & a \le p_i, q_i \lesssim 1, \qquad a \le p_i q_i \lesssim 1, \\ c_i^2, d_i^2 \gtrsim 1, \qquad & c_i^2, d_i^2 \lesssim 1, \qquad a \le \tau_i^2 \lesssim 1. \end{aligned}$$

Indeed, $\delta_{\mathrm{end}}^{-1} = O(\log(2n))$ and $p_i q_i \le C/(1-|y_i|)$, while $p_i, q_i \ge 1/4$. Lightness also gives the stronger upper bound

$$u_i, v_i \le \frac{2}{\iota\beta_i} \lesssim 1.$$

The movement estimate in Eq. (116) gives

$$ak_{\mathrm{act}} \le \mathrm{tr}[R_{\mathsf{A}}] \le Ck_{\mathrm{act}}, \qquad R_{\mathsf{A}} := (I - \widehat{T})\mathsf{A}(I - \widehat{T})^{\top}.$$

Rank-one Cauchy–Schwarz and $W \succeq \rho U^2$, $Z \succeq \rho V^2$ give

$$w_i \ge \rho u_i^2/t_i, \qquad z_i \ge \rho v_i^2/t_i.$$

Consequently, writing $\Delta_i := D_i\tau_i^2$,

$$\Delta_i = p_i^2 u_i w_i + q_i^2 v_i z_i \ge \frac{a(u_i^3 + v_i^3)}{nt_i} \ge an^{-1}t_i^2 \ge aP_n^{-1}n^{-3}. \tag{161}$$

The budget bounds also give

$$a \le k_i = \frac{\iota(-\beta_i'')}{2\eta p_i q_i} \le C\delta_{\mathrm{end}}^{-5/4} \lesssim 1.$$

Define the diagonal metric and the strict descent by

$$\varpi_i := \iota(-\beta_i'')(v_i w_i + u_i z_i), \qquad G := \mathrm{diag}(\varpi_i), \tag{162}$$

$$D_{\mathrm{tr}} := \kappa_{\mathrm{tr}} \sum_i k_i \rho_{\mathrm{rat}\,i} D_i \pi_i = \frac{\kappa_{\mathrm{tr}}}{2\eta}\,\mathrm{tr}[GK]. \tag{163}$$

The identity preceding Eq. (131) gives $\varpi_i/\Delta_i = 2\eta k_i \rho_{\mathrm{rat}\,i}$. At every light state, $aP_n^{-1}n^{-3} \le \varpi_i \le P_n$.

Suppose that

$$|\partial_i \mathcal{P}| < \kappa_1 P_n^{-1} n^{-1/2} \sqrt{\varpi_i} \quad \text{for every active } i. \tag{164}$$

We first show that $\widetilde{w}_i/\widetilde{z}_i \in [1/2, 2]$. The dual lower bounds and $u_i, v_i \ge at_i$ give

$$\frac{\varpi_i}{(\widetilde{w}_i + \widetilde{z}_i)^2} \lesssim \frac{u_i + v_i}{w_i + z_i} \lesssim nt_i \frac{u_i + v_i}{u_i^2 + v_i^2} \lesssim n.$$

Since $\partial_i \mathcal{P} = \tau_i(\widetilde{w}_i - \widetilde{z}_i)$ and $\tau_i \ge a$, Eq. (164) bounds $|\widetilde{w}_i - \widetilde{z}_i|/(\widetilde{w}_i + \widetilde{z}_i)$ by $1/3$, after enlarging P_n and choosing κ_1 sufficiently small. Consequently

$$\frac{1}{2} \le \rho_{\mathrm{rat}\,i} = \frac{\widetilde{u}_i \widetilde{z}_i + \widetilde{v}_i \widetilde{w}_i}{\widetilde{u}_i \widetilde{w}_i + \widetilde{v}_i \widetilde{z}_i} \le 2, \qquad \varpi_i \asymp \Delta_i$$

up to the polylogarithmic factors. The movement bound and $D_i \pi_i = (R_{\mathsf{A}})_{e_i,e_i}/\tau_i^2$ imply $D_{\mathrm{tr}} \gtrsim k_{\mathrm{act}}$.

We bound the drift in the same metric. For its ordinary part, the terms containing β_i' cancel: $\tau_i(\widetilde{u}_i + \widetilde{v}_i) = p_i u_i + q_i v_i = u_i + v_i$. Thus $|h_i^{(0)}| \lesssim (u_i + v_i) K_{ii}$. The lower bounds on w_i, z_i imply

$$\varpi_i \gtrsim \frac{u_i v_i (u_i + v_i)}{nt_i}, \qquad \frac{(u_i + v_i)^2}{\varpi_i} \lesssim nt_i (1/u_i + 1/v_i) \lesssim n.$$

It follows that $\sum_i |h_i^{(0)}| \sqrt{\varpi_i} \lesssim \sqrt{n}\,\mathrm{tr}[GK]$.

For the trace correction, the compact certificate and block weights give

$$|a_{*,i}| + |w_i^e| \le C, \qquad |w_i^{ef}| \le C\sqrt{\delta_i}, \qquad 0 < w_i^f \le C\delta_i,$$

$$\delta_i (b_{*,i}^2 + r_{*,i}^2) \le C, \qquad |\delta_i d_{*,i}| = |j_{0i}| \le C.$$

Hence $h_{*,i}, j_{*,i}, q_{*,i}, \delta_i p_{*,i}$ are bounded. Set $s_i := \min\{u_i, v_i\}$ and $r_i := \max\{u_i, v_i\}/s_i \ge 1$ within this paragraph. Then $s_i \ge at_i$, $\max\{u_i, v_i\} \lesssim 1$, and $|v_i| \lesssim \sqrt{r_i}$. For $r \ge 1$,

$$\frac{(1+r)^2(1+\sqrt{r})^2}{1+r^3} \le 16, \qquad \frac{r(1+\sqrt{r})^2}{1+r^3} \le 2.$$

Since $\delta_i = \iota \beta_i u_i v_i$, Eq. (161) gives

$$\frac{(u_i + v_i)(1 + |v_i|)}{\sqrt{\Delta_i}} \lesssim \sqrt{nt_i/s_i} \lesssim \sqrt{n}, \tag{165}$$

$$\frac{\delta_i (1 + |v_i|)^2}{\Delta_i} \lesssim \frac{nt_i}{s_i} \frac{r_i (1 + \sqrt{r_i})^2}{1 + r_i^3} \lesssim n. \tag{166}$$

The diagonal and lower off-diagonal correction is

$$\nu_i^{(0)} = \frac{\delta_i [h_{*,i} - j_{*,i} + p_{*,i} v_i] \tau_i (\widetilde{u}_i - \widetilde{v}_i)}{\Delta_i}.$$

Since $|\tau_i(\widetilde{u}_i - \widetilde{v}_i)| \le u_i + v_i$ and $\varpi_i \lesssim \Delta_i$, Eq. (165) gives $\sqrt{\varpi_i}|\nu_i^{(0)}| \lesssim \sqrt{n}$. The upper correction is $\nu_i^{(1)} := q_{*,i}\Omega_i$. The balance $\widetilde{w}_i/\widetilde{z}_i \in [1/2, 2]$ and Eq. (112) give $|\Omega_i| \lesssim (u_i v_i)^{3/2}/\Delta_i$. Since $|q_{*,i}| \lesssim 1$ and $\varpi_i \asymp \Delta_i$,

$$\sqrt{\varpi_i}|\nu_i^{(1)}| \lesssim \frac{(u_i v_i)^{3/2}}{\sqrt{\Delta_i}} \lesssim \sqrt{nt_i} \frac{(u_i v_i)^{3/2}}{\sqrt{u_i^3 + v_i^3}} \lesssim \sqrt{n}.$$

The last inequality uses lightness, $u_i, v_i \lesssim 1$, and $t_i \le 1$. Thus the weighted sum for $\nu^{(1)}$ is at most $P_n\sqrt{n}k_{\mathrm{act}}$. All helper functions are fixed polynomials in $|y_i|$ and $H_i := \iota^2\beta_i^2 u_i v_i/(1-y_i^2) \in [0,1]$. No derivatives of these helpers enter the descent argument.

For the mixed correction, retain the coefficient Γ_i in the coercivity inequality. The response condition gives $\mathsf{E}_i - \mathsf{Y}_i = e_i^\top \widehat{T}\xi$ and $\mathsf{F}_i = f_i^\top \widehat{T}\xi$. Since Γ is scalar on each paired plane and $\mathsf{A} \preceq CI$, Eq. (97) gives

$$\sum_i \Gamma_i \mathbb{E}[(\mathsf{E}_i - \mathsf{Y}_i)^2 + \mathsf{F}_i^2] = \operatorname{tr}[\widehat{T}^\top \Gamma \widehat{T}\mathsf{A}] \le C\operatorname{tr}[\widehat{D_{\mathrm{aux}}}] = 2Ck_{\mathrm{act}}. \tag{167}$$

Set $\mathsf{Z}_i := \mathsf{E}_i - \mathsf{Y}_i + \upsilon_i \mathsf{F}_i$. Eq. (166) and $\delta_i = 1/\Gamma_i$ imply

$$\sum_i \frac{\mathbb{E}[\mathsf{Z}_i^2]}{\Delta_i} \le \sum_i \frac{(1+|\upsilon_i|)^2}{\Delta_i}\mathbb{E}[(\mathsf{E}_i - \mathsf{Y}_i)^2 + \mathsf{F}_i^2] \lesssim nk_{\mathrm{act}}.$$

The definition of θ_i, $|g_i| \lesssim 1$, and $\varpi_i \asymp \Delta_i$ now give, by Cauchy–Schwarz,

$$\sum_i |\theta_i|\sqrt{\varpi_i} \lesssim \sum_i \frac{|\mathbb{E}[\mathsf{Y}_i\mathsf{Z}_i]|}{\sqrt{\Delta_i}} \le \sqrt{\sum_i \mathbb{E}[\mathsf{Y}_i^2]}\sqrt{\sum_i \frac{\mathbb{E}[\mathsf{Z}_i^2]}{\Delta_i}} \lesssim \sqrt{n}k_{\mathrm{act}}.$$

Here $\sum_i \mathbb{E}[\mathsf{Y}_i^2] = \operatorname{tr}[R_{\mathsf{A}}] \le Ck_{\mathrm{act}}$. Combining the contributions proves, under Eq. (164),

$$D_{\mathrm{tr}} \gtrsim k_{\mathrm{act}}, \qquad \mathscr{L}\mathcal{P} \le -D_{\mathrm{tr}}, \qquad \sum_i |h_i|\sqrt{\varpi_i} \le P_n\sqrt{n}D_{\mathrm{tr}}. \tag{168}$$

Thus, for a sufficiently small absolute $\kappa_1 > 0$,

$$\max_i \frac{|\partial_i \mathcal{P}|}{\sqrt{\varpi_i}} \ge \kappa_1 P_n^{-1} n^{-1/2} \quad \text{or} \quad \lambda_{\min}(G^{-1/2}\nabla^2\mathcal{P}G^{-1/2}) \le -\kappa_1. \tag{169}$$

Indeed, failure of the first inequality makes the drift term smaller than $D_{\mathrm{tr}}/4$ in absolute value. If the second also fails, then Eq. (163) gives $\operatorname{tr}[K\nabla^2\mathcal{P}]/2 > -D_{\mathrm{tr}}/4$ after decreasing κ_1. Their sum cannot be at most $-D_{\mathrm{tr}}$. The algorithm uses this alternative without computing the covariance, drift, or helper coefficients.

Multiplicity of negative curvature. The same covariance construction can be restricted to a prescribed subspace of physical directions. Let its codimension be $q \le k/8$. The response graph is bijective, so the restricted response space has dimension $\ell := k - q$. In Lemma 4.5, use an orthonormal basis E of this space and set

$$A := E(E^\top \mathsf{D}^{-1}E)^{-1}E^\top, \qquad B := \mathsf{D} - A.$$

All row completions and trace identities remain unchanged. The only rank-dependent equality is now $\operatorname{tr}[\mathsf{D}^{-1}A] = k - q$. Thus the right side of Eq. (115) acquires exactly the additive term q.

The movement lower bound remains proportional to k. To check this without using the smallest weight, the block bounds in that lemma give $A \preceq CI$ and $\mathsf{D}^{-1} \preceq C\Pi_0 + C\Gamma\Pi_1$. The response identity and energy dominance imply

$$\operatorname{tr}[E^\top \mathsf{D}^{-1}E] \le Ck, \qquad \|\widehat{T}\|_F^2 \le 2k/\iota.$$

Writing $\Pi := EE^\top$, we have

$$\operatorname{tr}[(I - \widehat{T})\Pi] \ge \ell - \sqrt{2k\ell/\iota} \ge k/16.$$

For the last inequality, the expression increases for $\ell \geq 7k/8$, and $(13/16)^2 > 7/(4\iota)$ for $\iota = 699/250$. Weighted Frobenius Cauchy–Schwarz, with $C_E := E^\top \mathsf{D}^{-1} E$, gives

$$(\operatorname{tr}[(I - \widehat{T})\Pi])^2 \leq \operatorname{tr}[(I - \widehat{T})A(I - \widehat{T})^\top] \operatorname{tr}[C_E].$$

Hence $ak \leq \operatorname{tr}[(I - \widehat{T})A(I - \widehat{T})^\top] \leq Ck$. The upper bound follows from $A \preceq CI$ and energy dominance.

All the preceding drift estimates hold for this restricted covariance: they used precisely these movement bounds, $A \preceq CI$, and Eq. (167), which still follows from $\widehat{T}^\top \Gamma \widehat{T} \preceq \widehat{D_{\text{aux}}}$. Thus, at a failed gradient test, its physical covariance K_q satisfies

$$D_q \geq a P_n^{-1} k, \qquad D_q = \frac{\kappa_{\text{tr}}}{2\eta} \operatorname{tr}[G K_q], \qquad \nabla \mathcal{P} \cdot h_q + \frac{1}{2} \operatorname{tr}[K_q \nabla^2 \mathcal{P}] \leq -D_q + q, \tag{170}$$

and $\sum_i |(h_q)_i| \sqrt{\varpi_i} \leq P_n \sqrt{n} D_q$. Choose the common gradient threshold small enough that its drift term has magnitude at most $D_q/4$, uniformly over the restricted spaces.

There are a fixed $\gamma > 0$ and an inverse-polylogarithmic $\pi_n > 0$ such that

$$\#\{\lambda_j(G^{-1/2} \nabla^2 \mathcal{P} G^{-1/2}) \leq -8\gamma\} \geq \pi_n k. \tag{171}$$

Indeed, suppose the number on the left is $q < \pi_n k$, and restrict to the inverse image of its orthogonal complement. Choose $\pi_n \leq 1/8$ and $\pi_n k \leq D_q/4$ using the uniform lower bound in Eq. (170). On this complement,

$$\frac{1}{2} \operatorname{tr}[K_q \nabla^2 \mathcal{P}] \geq -4\gamma \operatorname{tr}[G K_q] \geq -D_q/4$$

if $\gamma \leq \kappa_{\text{tr}}/(32\eta)$. The left side of Eq. (170) is then at least $-D_q/2$, whereas its right side is at most $-3D_q/4$, a contradiction. This also covers $\pi_n k < 1$, since q is an integer. These covariances are proof witnesses and are not computed.

Deterministic selection of a spread direction. We give the selection procedure explicitly. Suppose a real symmetric $k \times k$ matrix H has norm at most a polynomial in n, and at least $\pi_n k$ eigenvalues at most -4γ. Set $B := H + 2\gamma I$ and define the bounded filter

$$F_* := \frac{1}{2}(I - B(B^2 + \epsilon^2 I)^{-1/2}).$$

Choose a positive inverse-polynomial ϵ sufficiently small. The scalar formula shows that $0 \preceq F_* \preceq I$, $\operatorname{tr}[F_*^2] \geq \pi_n k/2$, and $\|\Pi_{\text{bad}} F_*\| \leq \epsilon^2/(4\gamma^2)$, where Π_{bad} is the spectral projection of H onto eigenvalues larger than $-\gamma$. Compute a symmetric dyadic approximation F such that

$$\|F\| \leq 2, \qquad \operatorname{tr}[F^2] \geq \pi_n k/2, \qquad \|\Pi_{\text{bad}} F\| \leq \eta_f, \quad \eta_f^2 \leq \frac{\gamma \pi_n}{16(\|H\| + \gamma)}. \tag{172}$$

There is slack for the trace guard: taking $\epsilon \leq \gamma/2$ makes $F_* \geq 3/4$ on the specified negative eigenspace, and decreasing ϵ further makes this eigenvalue arbitrarily close to one. No spectral projection is needed in the computation.

The inverse square root costs $\widetilde{O}(k^{\omega_0})$ with $O(\log n)$ working bits. We give an evaluation that preserves this precision. Define $A := B^2 + \epsilon^2 I$, and choose known bounds $a_0 I \preceq A \preceq b_0 I$, with a_0^{-1}, b_0 polynomial in n. For a prescribed inverse-polynomial error η_s, use

$$A^{-1/2} = \frac{2}{\pi} \int_0^\infty (A + t^2 I)^{-1} dt.$$

Choose an integer $J = O(\log n)$ so that $2^{-J}(a_0^{-1} + 1) \le \eta_s/8$. Truncating the integral to $[2^{-J}, 2^J]$ has operator error at most $\eta_s/8$ before the factor $2/\pi$. Split this interval into $[a, 2a]$ with $a = 2^j$. Define $R_a := (A/a^2 + (5/2)I)^{-1}$, so $\|R_a\| \le 2/5$. On each interval the geometric expansion gives

$$\int_a^{2a} (A + t^2 I)^{-1} dt = a^{-1} \sum_{\ell=0}^{s-1} c_\ell R_a^{\ell+1} + E_a, \qquad c_\ell := \int_1^2 (5/2 - u^2)^\ell du.$$

Since $|5/2 - u^2| \le 3/2$, the remainder summed over all intervals has norm at most $(3/5)^s \pi/(2\sqrt{a_0})$. Choose $s = O(\log n)$ to make this at most $\eta_s/8$. The coefficients are explicit:

$$c_\ell = \sum_{v=0}^{\ell} (-1)^v \binom{\ell}{v} (5/2)^{\ell - v} \frac{2^{2v+1} - 1}{2v + 1}.$$

The sum of the absolute values of these terms is at most $2(13/2)^\ell$, which is polynomial for $\ell \le s$. Evaluate them on a common dyadic grid, rounding after each operation. All matrix norms, inverse norms, scalar magnitudes, and error amplifications in these finite sums are polynomial. The powers of R_a are contractive. Thus the short-precision inversion and stable products in Section 7.4, with a sufficiently large fixed multiple of $\log n$ bits, give total error at most η_s. There are $O(\log n)$ intervals, and each uses $O(\log n)$ products, including its inverse; the cost remains $\widetilde{O}(k^{\omega_0})$ in both models. Choose η_s below a fixed fraction of the filter guards divided by $1 + \|B\|$, symmetrize the output, and use the trace slack above to obtain Eq. (172).

To obtain a flat vector, consider independent signs ξ_j only as a device for conditional expectation, and put

$$\Phi(\xi) := \|F\xi\|^2 - \lambda_f \sum_i \cosh((F\xi)_i), \qquad \lambda_f := \pi_n/100.$$

Each row has squared norm at most four. Since $\prod_j \cosh(F_{ij}) \le \exp(\sum_j F_{ij}^2/2) < 8$, Eq. (172) gives $\mathbb{E}[\Phi] \ge \pi_n k/2 - 8\lambda_f k > 2\pi_n k/5$. Fix the signs one at a time, choosing the larger conditional expectation, to additive error at most $\pi_n/100$ per decision. Then $z := F\xi$ satisfies

$$\Phi(\xi) \ge \pi_n k/4, \qquad \|z\|^2 \ge \pi_n k/4, \qquad \|z\|_\infty \le \log(800k/\pi_n).$$

For the last inequality use $\|z\|^2 \le 4k$ and $\lambda_f \cosh(z_i) \le \|z\|^2$. Also Eq. (172) gives

$$z^\top H z \le -\gamma \|z\|^2 + (\|H\| + \gamma)\eta_f^2 k \le -3\gamma \|z\|^2/4.$$

Maintain the prefix sum, the remaining squared column norms, and the row products in the conditional estimator. Each choice costs $\widetilde{O}(k)$ arithmetic operations, hence the total is $\widetilde{O}(k^2)$. Section 7.6 proves that logarithmic precision suffices even when the two choices nearly tie.

Apply this procedure to a symmetric approximation of $\widehat{G}^{-1/2} \nabla^2 \mathcal{P} \widehat{G}^{-1/2}$ with operator error at most $\gamma/4$. Eq. (171) and $\widehat{G} \preceq G \preceq 2\widehat{G}$ supply the required multiplicity with slack. Undo the diagonal scaling and normalize, $r := \widehat{G}^{-1/2} z / \|\widehat{G}^{-1/2} z\|$. The computed direction, after sufficiently accurate dyadic rounding and normalization, has certified bounds

$$r^\top \nabla^2 \mathcal{P} r \le -a r^\top G r, \qquad \|r\|_\infty^2 \le P_n n (r^\top G r). \tag{173}$$

The curvature margin follows from the preceding quotient and the operator error. For the second bound, before the harmless rounding,

$$\frac{\|r\|_\infty^2}{r^\top G r} \le \frac{\max_i (z_i^2/\widehat{\varpi}_i)}{\|z\|^2} \le \frac{P_n}{k \min_i \widehat{\varpi}_i} \le P_n n$$

by Eq. (160). All quantities have polynomial reciprocals where needed, so the rounding and the final quotient and spread checks require only $O(\log n)$ working bits. The procedure uses no random bits and does not assume a gap between adjacent eigenvalues of H.

Entropy smoothing for the smaller constant. We now reduce the dimension exponent by smoothing the matrix maximum. The budget, coupling, preprocessing, and rounding allowances remain unchanged. Set

$$L_{\log} := \lceil \log_2(4n) \rceil, \qquad \tau_{\mathrm{s}} := 10^{-8}/L_{\log},$$

and define, on matrix pairs,

$$\mathcal{C}(\mathsf{P}, y) := (U^{-1} + M(y) + \mathcal{K}_y(V), V^{-1} - M(y) + \mathcal{K}_y(U)), \qquad \mathsf{P} := (U, V).$$

Use the smoothed potential without a scalar reservoir:

$$\mathcal{P}_{\mathrm{s}}(y) := \min_{U,V \succ 0} \{\rho \operatorname{tr}[U + V] + \tau_{\mathrm{s}} \log \operatorname{tr}[\exp(\mathcal{C}(\mathsf{P}, y)/\tau_{\mathrm{s}})]\}. \tag{174}$$

The trace and exponential in this formula act on the block-diagonal matrix associated with the pair. The log-sum-exp inequalities give

$$\|M(y)\| \le \mathcal{P}(y) \le \mathcal{P}_{\mathrm{s}}(y) \le \mathcal{P}(y) + \tau_{\mathrm{s}} \log(2n) \le \mathcal{P}(y) + 10^{-8}. \tag{175}$$

Let $\mathsf{Q} := (W, Z)$ be its Gibbs state. Stationarity in U, V gives

$$\operatorname{tr}[\mathsf{Q}] = 1, \qquad \mathscr{J}\mathsf{Q} = \rho E, \qquad W = U\mathcal{K}_y(Z)U + \rho U^2, \quad Z = V\mathcal{K}_y(W)V + \rho V^2.$$

The bounded-sublevel, positive-inverse, and unique-root arguments in Section 6.1 use these identities and nonnegative coupling coefficients. They therefore apply to β_{tr} and $\iota = 699/250$. In particular all the primal and dual bounds used above remain valid. The budget inequalities needed there follow from the fixed even polynomial and its 3/4 endpoint power on the truncated interval. No reservoir is used in these root estimates.

Transfer of the trace-corrected covariance. It is essential to distinguish the comparison response from the derivative of the smoothed optimizer. At a smoothed root and for a direction r, the inverse function theorem supplies a local comparison path $\mathsf{P}_{\mathrm{c}}(t)$ satisfying

$$\mathcal{C}(\mathsf{P}_{\mathrm{c}}(t), y + tr) = \mathcal{C}(\mathsf{P}, y), \qquad \mathsf{P}_{\mathrm{c}}(0) = \mathsf{P}.$$

Indeed the derivative with respect to P is $-\mathscr{J}$, which is invertible by the strict dual identities. Its first derivative satisfies Eq. (87). The comparison value

$$\varphi_r(t) := \rho \operatorname{tr}[\mathsf{P}_{\mathrm{c}}(t)] + \tau_{\mathrm{s}} \log \operatorname{tr}[\exp(\mathcal{C}(\mathsf{P}, y)/\tau_{\mathrm{s}})]$$

majorizes $\mathcal{P}_{\mathrm{s}}(y + tr)$ and agrees with it at zero. The second term is constant along this path. Pairing the first and second differentiated constraint equations with Q gives

$$\partial_i \mathcal{P}_{\mathrm{s}} = p_i w_i - q_i z_i, \tag{176}$$

$$\frac{1}{2} r^\top \nabla^2 \mathcal{P}_{\mathrm{s}} r \le \mathcal{Q}(r) + \sum_i m_i(r) - \frac{1}{2} r^\top G r. \tag{177}$$

Here $G = \operatorname{diag}(\varpi_i)$ is exactly Eq. (162), evaluated at the smoothed root; $\mathcal{Q}(r)$ and $m_i(r)$ use the comparison derivatives $\dot{U}_{\mathrm{c}}, \dot{V}_{\mathrm{c}}$, as in Eq. (118). There is no entropy remainder in Eq. (177).

The response graph, energy dominance, and local trace identities of Lemma 4.5 require the response equation, the strict dual identities, rank one, and lightness. They do not require $\mathcal{C} = bE$. Thus that lemma, the exact mixed correction, and its scalar certificate apply verbatim to the comparison response, with $\nabla\mathcal{P}_{\mathrm{s}}$ in place of $\nabla\mathcal{P}$. The covariance restriction proof also applies. In particular the drift estimate in Eq. (168), the separated gradient test, and the multiplicity in Eq. (171) hold for $\mathcal{P}_{\mathrm{s}}$. The deterministic filter consequently supplies

$$r^\top \nabla^2 \mathcal{P}_{\mathrm{s}} r \le -a r^\top G r, \qquad \|r\|_\infty^2 \le P_n n (r^\top G r), \qquad \|r\|_2 = 1. \tag{178}$$

The phase trace floor in Eq. (160) is unchanged.

Local steps and optimizer correction. We now use derivatives of the actual smoothed optimizer. On the trace-zero dual space define the entropy energy

$$\mathcal{E}(B)^2 := \tau_{\mathrm{s}}\langle B, D\log_{\mathsf{Q}}[B]\rangle.$$

For a local line let $A := \mathsf{P}'$, $B := \mathsf{Q}'$, $F_2 := \mathcal{A}'\mathsf{Q}$, and $\theta := r^\top G r$. The exact envelope identity and the coupling bound are

$$\mathcal{P}_{\mathrm{s}}'' = -\theta + e(A)^2 + 2\langle F_2, A\rangle + \mathcal{E}(B)^2, \qquad |\langle F_2, D\rangle| \le P_n\sqrt{\theta}\,e(D). \tag{179}$$

The Gibbs variance and relative-dual estimates in Eqs. (200) and (201) apply here. Together with the forced energy identity they show, on a coordinate line with $\theta_0 = \varpi_i(y)$,

$$e(A) + \mathcal{E}(B) \le P_n\sqrt{\theta_0}, \qquad |\mathcal{P}_{\mathrm{s}}''| \le P_n\theta_0, \qquad |\varpi_i'| \le P_n(1+\sqrt{n\theta_0})\varpi_i$$

in its comparison tube. To check the forcing without a reservoir, write $\nu_i := \mathrm{tr}[C_i]$. Its signed-matrix term obeys

$$\mathrm{tr}[\mathsf{Q}F_1^2] = \nu_i(p_i^2 w_i + q_i^2 z_i) \le P_n\nu_i(w_i + z_i) \le P_n(v_i w_i + u_i z_i) \le P_n\varpi_i,$$

using $u_i, v_i \ge a\nu_i$ and the expanded light bounds. This is the coordinate estimate of Section 6.3 with the reservoir set to zero; its use does not require a floor of order $1/k$.

Face changes for the smoothed implementation. We retain the phase cutoff and deterministic small-input rounding. At a regular endpoint, choose the score accuracy to satisfy

$$\delta \le \min\{e_{\mathrm{end}}/(16n^2), \delta_{\mathrm{end}}/8, \rho\delta_{\mathrm{end}}^3/(16(4\iota^2+2\iota))\}.$$

For a move to 1, write $d_i := 1 - y_i \ge \delta_{\mathrm{end}}$, $\nu_i := \mathrm{tr}[C_i] \le 1$, and $h_i := \iota\beta_i u_i + d_i$. Test the new face at the old U and $V^+ := (V^{-1} + h_i C_i)^{-1}$. Its minus constraint block is unchanged; its plus block is at most the old block plus $2\delta C_i$. Order monotonicity and translation by a scalar identity give

$$\mathcal{P}_{\mathrm{s,old}} - \mathcal{P}_{\mathrm{s,new}} \ge \rho\frac{h_i v_i^2}{\nu_i(1+h_i v_i)} - 2\delta\nu_i \ge a\rho\delta_{\mathrm{end}}^3.$$

For the second bound, the conservative trigger implies $v_i \ge \delta_{\mathrm{end}}/(2\iota)$, while $h_i \ge \delta_{\mathrm{end}}$. The fraction is increasing in h_i, v_i and decreasing in ν_i. The move to -1 exchanges the two blocks. Thus only $\widetilde{O}(n)$ regular endpoints occur. Their charged continuation is verified below using Lemma 6.22.

For a deferred or small-input removal, preserve M and attenuate its coupling slot before deleting it. The argument in Lemma 6.3 also applies with no reservoir. Explicitly, write the slot as $e^{-s}\mathcal{K}_i$, let K_s be its pair map, and put $v(s) := \langle \mathsf{Q}, K_s\mathsf{P}\rangle \ge 0$. The envelope and root forcing give

$$\mathcal{P}_{\mathrm{s}}'(s) = -v(s), \quad e(\mathsf{P}') + \mathcal{E}(\mathsf{Q}') \le P_n\sqrt{v(s)}, \quad |v'(s)| \le P_n v(s).$$

These estimates only use positivity and inclusion of the slot in the full coupling. In particular they apply to small-input removals, not only coordinates close to an endpoint. A step of order $1/(P_n(1+\sqrt{nv(s)}))$ admits an anchored correction. Track to a positive slot weight below the root-cache floor divided by a fixed polynomial in n, k, then delete its remainder and correct once. Its negative logarithm is polylogarithmic. The root inverse has a uniform polynomial bound also at zero slot weight, so this last jump fits the Newton ball. The finite-precision construction in Lemma 6.3 supplies the same certified implementation.

For a removal with potential drop Δ_i, integration and Cauchy–Schwarz bound its corrected stages by $\widetilde{O}(1+\sqrt{n\Delta_i})$. We apply this bound to positive-slot batches below. The partial-signing residual, small-input signs, and final deferred rounding have exactly their previous discrepancy allowances.

Hessian assembly. For completeness the smoothed Hessian is explicitly constructible. Let $Lr := \mathcal{C}_y[r]$, let $Tr := \mathcal{A}'_y[r]\mathsf{Q}$, and let Q_{Q} denote the Gibbs covariance in Eq. (199). Then

$$J_{\mathrm{s}} := \mathscr{E} + \mathscr{J}\, Q_{\mathsf{Q}}\, \mathscr{J}, \qquad B := T - \mathscr{J}\, Q_{\mathsf{Q}} L,$$

and the ordinary Hessian of the optimized value is

$$\nabla^2 \mathcal{P}_{\mathrm{s}} = -G + L^\top Q_{\mathsf{Q}} L - B^\top J_{\mathrm{s}}^{-1} B. \tag{180}$$

This follows by differentiating the primal objective twice and taking the Schur complement in its minimizing variables. In real Hermitian coordinates all maps have order $O(n^2)$ and $k \le n^2$. The maps L, T have explicit rank-one columns, while Q_{Q} has logarithmic-mean coefficients in a basis diagonalizing W, Z. Forming their coordinate matrices costs $\widetilde{O}(n^4)$. Their products, the coupling blocks, and the inverse of J_{s} cost $\widetilde{O}(n^{2\omega_0})$, including short-precision inversion. Indeed $\mathscr{E} \preceq J_{\mathrm{s}} \preceq P_n \mathscr{E}$ and $\mathscr{E} \succeq a n^{-3/2} I$, so all inverse norms are polynomial. Symmetrize and compute sufficiently accurate entries to enclose the true scaled Hessian in operator norm. The deterministic filter already proved above has the same $\widetilde{O}(n^{2\omega_0})$ cost.

A small reservoir and two diagonal metrics. We control the matrix response of a curvature direction as well as its coordinates. The coupling and scalar certificate remain unchanged. In a phase starting with K coordinates, choose a positive dyadic λ_K within fixed factors of $a/(P_n n\sqrt{K})$ and set

$$\mathcal{F}_{\mathrm{s}} := \mathcal{P}_{\mathrm{s}} + \lambda_K \sum_{i \in I} \beta_i, \qquad G_+ := G + \lambda_K \operatorname{diag}(-\beta_i''). \tag{181}$$

Here and below a is a sufficiently small fixed constant and P_n is a sufficiently large fixed polylogarithm. Fix these choices once, including the finite number of inequalities below. The optimizer does not depend on λ_K. The total positive reservoir injection, including initialization and all phase changes, is at most

$$\sum_{\text{phases}} \lambda_K K \le \frac{Ca}{P_n n} \sum_{\text{phases}} \sqrt{K} \le Ca/P_n.$$

Choose it below $e_{\mathrm{end}}/4$. Reservoir removals are nonnegative. Thus the sum of all accepted potential decreases is still bounded by an absolute constant.

Gradient tests use G, not G_+. By Eq. (160) and the truncated budget bounds, we can choose λ_K so that

$$|\lambda_K \beta_i'| \le a P_n^{-1} \sqrt{\varpi_i / n}$$

with any prescribed fixed slack. A failed separated gradient test for $\mathcal{F}_{\mathrm{s}}$ therefore implies the hypothesis of Eq. (170) for $\mathcal{P}_{\mathrm{s}}$. For every restricted covariance K_q, adding the reservoir gives

$$\frac{1}{2} \operatorname{tr}[K_q \nabla^2 \mathcal{F}_{\mathrm{s}}] \le -\frac{3}{4} D_q + q - \frac{\lambda_K}{2} \operatorname{tr}[K_q \operatorname{diag}(-\beta_i'')].$$

Here the drift is paired with $\nabla \mathcal{P}_{\mathrm{s}}$, whose magnitude is at most $D_q/4$. Since D_q is a fixed positive multiple of $\operatorname{tr}[G K_q]$, the same restriction argument proves Eq. (171) with G_+ and $\nabla^2 \mathcal{F}_{\mathrm{s}}$, after decreasing its fixed constants. Moreover, during this phase,

$$G_+ \succeq \frac{a}{P_n n \sqrt{k}} I, \qquad K/2 < k \le K. \tag{182}$$

Deterministic control of the matrix responses. Let $\mathcal{V}$ denote the bounded symmetric filter constructed above, now for $H := G_+^{-1/2}\nabla^2\mathcal{F}_{\mathrm{s}}G_+^{-1/2}$. Its guards give $\|\mathcal{V}\| \le 2$ and $\mathrm{tr}[\mathcal{V}^2] \ge \pi_n k/2$. Its leakage outside the negative spectral subspace can be made any prescribed inverse polynomial. For a physical direction r, let $A(r), B(r)$ be the derivatives of the actual smoothed primal and dual optimizers. Introduce the four real-linear Hermitian-valued maps

$$\mathcal{T}_1 r := U^{-1/2}A_U(r)U^{-1/2}, \qquad \mathcal{T}_2 r := V^{-1/2}A_V(r)V^{-1/2},$$
$$\mathcal{T}_3 r := W^{-1/2}B_W(r)W^{-1/2}, \qquad \mathcal{T}_4 r := Z^{-1/2}B_Z(r)Z^{-1/2}.$$

For $r = G_+^{-1/2}\mathcal{V}x$, completing the square in Eq. (179), with the reservoir included, gives

$$e(A(r)) + \mathcal{E}(B(r)) \le P_n\|x\|_2, \qquad \|\mathcal{T}_\ell G_+^{-1/2}\mathcal{V}\|_{2\to F} \le P_n\sqrt{n}. \tag{183}$$

Indeed, the direct term is $-r^\top G_+ r$, and the coupling term is bounded by $P_n\sqrt{r^\top G r}e(A)$. On the negative subspace the Hessian quadratic form is nonpositive; the leakage contributes at most $\|x\|^2$ after tightening its error. The primal relative Frobenius bound follows from

$$U^{-1/2}WU^{-1/2} \succeq aI/n, \qquad V^{-1/2}ZV^{-1/2} \succeq aI/n.$$

For the dual bound, let q_i be the eigenvalues of Q. The entropy energy of B has coefficient

$$\tau_{\mathrm{s}} q_i q_j / L(q_i, q_j) \ge a\tau_{\mathrm{s}}/n$$

on the entries of $\mathsf{Q}^{-1/2}B\mathsf{Q}^{-1/2}$, where L is the logarithmic mean and $L(q_i, q_j) \le \max(q_i, q_j)$. This proves Eq. (183) for complex Hermitian matrices as well.

Put $D_{\ell j} := \mathcal{T}_\ell G_+^{-1/2}\mathcal{V}e_j$. For each ℓ,

$$\|\sum_j D_{\ell j}^2\| \le P_n n^2. \tag{184}$$

To see this, fix a unit vector v. The real-linear map $D \mapsto Dv$ from Hermitian matrices with Frobenius norm has norm at most one and rank at most $2n$. Its composition with the map in Eq. (183) therefore has squared Frobenius norm at most $2n(P_n\sqrt{n})^2$. This is precisely $\sum_j \|D_{\ell j}v\|^2$; enlarging P_n proves the claim.

We now choose signs ξ_j deterministically. The following conditional-expectation calculation uses independent uniform signs only as a proof device. Choose $s := 1/(P_n n)$ so that $s^2\|\sum_j D_{\ell j}^2\| \le 1$. Consider

$$\|\mathcal{V}\xi\|^2 - \frac{\pi_n}{100}\sum_i \cosh((\mathcal{V}\xi)_i) - \frac{\pi_n k}{1000n}\sum_{\ell=1}^{4}\sum_{\varepsilon=\pm 1}\mathrm{tr}[\exp(\varepsilon s\sum_j \xi_j D_{\ell j})]. \tag{185}$$

Use the exact conditional expectation for the first two terms. For a partial assignment with remaining set J, replace a matrix penalty by

$$\mathrm{tr}[\exp(\varepsilon s S_\ell + (s^2/2)\sum_{j\in J} D_{\ell j}^2)], \qquad S_\ell := \sum_{j\notin J}\xi_j D_{\ell j}.$$

This is a pessimistic estimator: its two children have average at most its current value. Apply the trace-exponential concavity inequality [Tro12, Corollary 3.3] and then $\log\cosh(sD) \preceq s^2D^2/2$. No commutativity of the different $D_{\ell j}$ is used. Initially each matrix penalty is at most $ne^{1/2}$, whereas each scalar penalty has expectation at most e^2, because every row of $\mathcal{V}$ has squared norm at most

four. The initial lower estimator is consequently at least $\pi_n k/3$. At each assignment choose the child with the larger lower estimator. Errors below $\pi_n/100$ per comparison leave a final value at least $\pi_n k/4$.

Writing $z := \mathcal{V}\xi$, positivity of all penalties and $\|z\|^2 \le 4k$ give

$$\|z\|^2 \ge \pi_n k/4, \qquad \|z\|_\infty \le P_n, \qquad \|\sum_j \xi_j D_{\ell j}\| \le P_n n.$$

Both signs in the exponential penalty are needed for the last operator-norm bound. Normalize $r := G_+^{-1/2} z / \|G_+^{-1/2} z\|_2$ and put $\theta := r^\top G_+ r$. The filter leakage, chosen smaller than a fixed inverse-polynomial multiple of π_n, gives

$$r^\top \nabla^2 \mathcal{F}_{\mathrm{s}} r \le -a\theta, \qquad R := \|r\|_\infty + \max_\ell \|\mathcal{T}_\ell r\| \le P_n d_k \sqrt{\theta}, \qquad d_k := \min\{\sqrt{n}, n/\sqrt{k}\}. \tag{186}$$

For the matrix term, combine the last display with Eq. (183). For the coordinate term, Eq. (182) gives $\|r\|_\infty^2/\theta \le P_n n/\sqrt{k}$. This is at most $P_n d_k^2$ because $1 \le k \le n^2$.

The sequential sign choices cost $\widetilde{O}(k^2 + kn^3)$ once the response columns are available. Scalar conditional expectations are maintained as in the earlier flat-vector construction. For the matrix terms maintain S_ℓ and the sum of the remaining squares, and evaluate the eight trace exponentials at each child by symmetric matrix-function evaluation. This costs $\widetilde{O}(n^3)$ per choice. Every retained lower estimator is positive, so its matrix penalties are bounded by $O(n/\pi_n)$. Their variance terms are PSD and bounded in norm. Including both signs then bounds all their exponent matrices in norm by $O(\log(2n/\pi_n))$. A trial child changes this norm by a polynomially bounded amount, in fact $s\|D_{\ell j}\| \le P_n/\sqrt{n}$. Thus the scalar and matrix functions, response columns, and all conditional comparisons can be enclosed at logarithmic working precision with any prescribed inverse-polynomial error. Choose P_n in s also to bound $s\|D_{\ell j}\|$ by one. The selection has no randomized fallback.

A finite-step estimate from linear primal and dual paths. Matrix response control at one root does not by itself control the actual optimizer along a longer step. We instead use linear comparison paths and bound their entropy residual. Write $A := A(r)$, $B := B(r)$, and, for this paragraph only, define

$$\mathsf{P}_t := \mathsf{P} + tA, \qquad \mathsf{Q}_t := \mathsf{Q} + tB, \qquad y_t := y + tr.$$

For $|t|R \le a/P_n$, these matrices remain positive and comparable to their base values, $\mathrm{tr}[\mathsf{Q}_t] = 1$, and all coordinates remain at endpoint distance at least $\delta_{\mathrm{end}}/2$. Define

$$S(t) := \rho \,\mathrm{tr}[\mathsf{P}_t] + \langle \mathsf{Q}_t, \mathcal{C}(\mathsf{P}_t, y_t)\rangle - \tau_{\mathrm{s}} \,\mathrm{tr}[\mathsf{Q}_t \log \mathsf{Q}_t] + \lambda_K \sum_i \beta(y_i + tr_i).$$

Stationarity and its first derivative imply

$$S(0) = \mathcal{F}_{\mathrm{s}}(y), \qquad S'(0) = \mathcal{F}_{\mathrm{s}}'[r], \qquad S''(0) = \mathcal{F}_{\mathrm{s}}''[r, r].$$

Let c_0, c_1 be the scalar multipliers in the Gibbs equation and its derivative. The Hermitian residual

$$E(t) := \mathcal{C}(\mathsf{P}_t, y_t) - \tau_{\mathrm{s}} \log \mathsf{Q}_t - (c_0 + c_1 t) I$$

satisfies $E(0) = E'(0) = 0$. We claim that, on this linear comparison interval, there is a PSD pair $J(t)$ with

$$-J(t) \preceq E(t) \preceq J(t), \qquad \|J(t)\| \le P_n R^2 t^2, \qquad \langle \mathsf{Q}_t, J(t)\rangle \le P_n \theta t^2, \qquad |S'''(t)| \le P_n R\theta. \tag{187}$$

Here the bounds use $e(A)^2+\mathcal{E}(B)^2 \le P_n\theta$, which follows from the envelope identity for the selected negative direction, and not from the norm of its sign seed.

We give the majorants explicitly enough to verify both different bounds in Eq. (187). Along the linear path, the second derivative of each constraint block is a sum of an inverse quadratic, a term $\mathcal{K}_{yy}[r,r]\mathsf{P}_t$, and $2\mathcal{K}_y'[r]A$. An inverse quadratic is PSD, has operator norm at most $P_n \max_\ell \|\mathcal{T}_\ell r\|^2$, and weighted trace at most $P_n e(A)^2$. The direct coupling term has a PSD absolute majorant with norm at most $P_n\|r\|_\infty^2$ and weighted trace at most $P_n r^\top G r$. Use $|\beta''|/\beta \le P_n$ and the bounded operator norms of $\mathcal{K}_y(U), \mathcal{K}_y(V)$.

For the mixed term, rank one and scalar Young's inequality majorize each absolute coefficient, up to a fixed factor, by

$$(-\beta_i'')r_i^2 C_i P C_i + \frac{(\beta_i')^2}{-\beta_i''} C_i A_P P^{-1} A_P C_i,$$

where P is the relevant primal block. This uses $|\operatorname{tr}[C_i A_P]|^2 \le \operatorname{tr}[C_i P]\operatorname{tr}[C_i A_P P^{-1} A_P]$. The operator norm of the sum is at most $P_n(\|r\|_\infty^2 + \|P^{-1/2}A_P P^{-1/2}\|^2)$. Its weighted trace is at most $P_n(\theta + e(A)^2)$: pair the second sum with the dual block, use $(\beta_i')^2/[(-\beta_i'')\beta_i] \le P_n$, and then $\mathcal{K}_y(Z) \preceq U^{-1}WU^{-1}$ and its exchanged identity. All these bounds remain valid on the linear paths by fixed-factor Loewner comparisons and scalar budget comparisons; no order comparison of noncommuting products is required.

Finally, $-D^2 \log_Q[B,B]$ is PSD, has norm at most $2\|Q^{-1/2}BQ^{-1/2}\|^2$, and satisfies

$$-\operatorname{tr}[QD^2\log_Q[B,B]] = \operatorname{tr}[BD\log_Q[B]].$$

For the norm assertion use the resolvent representation of log: if $T := Q^{-1/2}BQ^{-1/2}$ and $R_u := (Q+uI)^{-1}$, then $BR_uB \preceq \|T\|^2 Q$, so $2\int_0^\infty R_u B R_u B R_u du \preceq 2\|T\|^2 I$. The trace identity follows by differentiating $\operatorname{tr}[QD\log_Q[B]] = \operatorname{tr}[B]$ along $Q+tB$. Multiplication by τ_{s} gives weighted trace $\mathcal{E}(B)^2$. Entropy energies along comparable linear matrices are comparable by the same resolvent formula. Integrate these PSD absolute majorants for E'' twice. This proves the first three assertions of Eq. (187).

For the last assertion, differentiate S three times. Inverse cubic words cost at most $P_n R e(A)^2$; differentiating the direct coupling or its scalar budget costs at most $P_n R\theta$. The terms containing B paired with the second constraint derivative cost at most its PSD weighted majorant times $P_n \max_\ell \|\mathcal{T}_\ell r\|$, because $-R\mathsf{Q}_t \preceq B \preceq R\mathsf{Q}_t$ up to a fixed factor. The entropy cubic costs at most $P_n R\mathcal{E}(B)^2$, by the preceding positive resolvent majorant. The scalar reservoir cubic is bounded by $P_n\|r\|_\infty \lambda_K \sum_i(-\beta_i'')r_i^2$. These are all terms, since both matrix paths are linear, and prove the last assertion.

For completeness, the entropy residual converts this comparison into an upper bound on the optimized value. If $Q \succ 0$ has trace one and $\|E\| \le \tau/2$, Golden–Thompson and $e^x \le 1+x+x^2$ for $|x| \le 1/2$ give

$$\tau \log \operatorname{tr}[\exp(\log Q + E/\tau)] - \operatorname{tr}[QE] \le \tau^{-1}\operatorname{tr}[QE^2]. \tag{188}$$

Indeed the exponential trace is at most $\operatorname{tr}[Qe^{E/\tau}]$; apply the scalar inequality spectrally and then $\log(1+x) \le x$. If $-J \preceq E \preceq J$, factor $E = J^{1/2}TJ^{1/2}$ with $\|T\| \le 1$. It follows that $E^2 \preceq \|J\|J$. Apply this to Eq. (187) and use $\tau_{\mathrm{s}}^{-1} \le P_n$. Evaluating the primal objective at P_t, rather than its minimizer, yields the rigorous finite-step bound

$$\mathcal{F}_{\mathrm{s}}(y_t) \le \mathcal{F}_{\mathrm{s}}(y) + t\mathcal{F}_{\mathrm{s}}'[r] + \frac{t^2}{2}\mathcal{F}_{\mathrm{s}}''[r,r] + P_n R\theta|t|^3 + P_n R^2\theta t^4. \tag{189}$$

The smallness $\|E(t)\| \le \tau_{\mathrm{s}}/2$ is ensured by the same choice of P_n and step below. This argument uses no transport assumption about the optimizer between the two endpoints.

Orient r so that its directional gradient is nonpositive, allowing error at most $a\theta t$ in this comparison, and take a dyadic step within fixed factors of

$$t := \frac{a}{P_n^2 d_k \sqrt{\theta}}. \tag{190}$$

Eqs. (186) and (189) give a true decrease at least $a/(P_n d_k^2)$, after enlarging P_n. The chosen branch and the comparison value are bounded throughout the segment. We evaluate only this oriented trial. Its new optimizer is initialized by the explicit data homotopy in Section 6.4, at cost $\widetilde{O}(\sqrt{n})$ corrected stages; a single warm correction over the long step is not assumed. Inward rounding and certified value comparison reserve a fixed fraction of the decrease. All required thresholds have polynomial reciprocals by Eq. (182).

Gradient sweeps and batched removals for the trace budget. To avoid paying for a root after every short coordinate move, use the deterministic frozen-root sweeps of Lemmas 6.23 and 6.24. The necessary adaptation is to keep the original, smaller weight in the gradient test. At a light base put

$$\kappa_i^0 := \iota(v_i^0 w_i^0 + u_i^0 z_i^0), \qquad \varphi_i(t) := (w_i^0 - z_i^0)t + (\kappa_i^0 + \lambda_K)\beta(t), \qquad d_0 := \sum_i \sqrt{\kappa_i^0}|y_i - y_i^0|.$$

The proof of Lemma 6.23 bounds both coordinate forcings by $P_n\sqrt{\kappa_i^0}$ even without a reservoir. The scalar reservoir has zero derivative in root variables, so it cancels exactly from the gradient approximation error. Consequently, for $d_0 \le a/(P_n^2\sqrt{n})$, that proof gives

$$\|z(y) - z_0\|_{\mathrm{Met}_0} \le P_n d_0, \qquad |\partial_i \mathcal{F}_{\mathrm{s}}(y) - \varphi_i'(y_i)| \le P_n^2\sqrt{\kappa_i^0} d_0, \qquad \kappa_i(y) \asymp \kappa_i^0.$$

Here $z = (\mathsf{P}, \mathsf{Q})$ and Met_0 is the joint energy metric. These claims require only a small weighted parameter neighborhood, not lightness at every intermediate point. Choose high and low thresholds as in Lemma 6.24, with normalization $\sqrt{\kappa_i/n}$. Make the high threshold small enough for the failed-gradient condition proved above. Concavity of every φ_i then gives a saturated sweep decrease at least $1/(P_n n)$, or an unsaturated exit with every remaining normalized gradient below $h_n/(8\sqrt{n})$. The same proof gives at most k_0 complete coordinate updates and one partial update per saturated sweep. Its positive-decrease floor now uses $\kappa_i \ge a/(P_n n k)$ from the phase cutoff in place of the reservoir floor; it is still inverse polynomial.

All marked coordinates are removed in one attenuation batch, preserving the modeled signed sum. Small-input removals are likewise batched once per phase. The positive-submap calculation for face changes applies to the sum of any such slots: positivity, rather than its rank, was the only property used. If its matrix-potential drop is Δ, its joint energy arc is at most $P_n\sqrt{\Delta}$ and it costs $\widetilde{O}(1 + \sqrt{n\Delta})$ corrected stages. Hold λ_K fixed during the batch and remove its scalar reservoir contributions at the end.

For clarity the charging argument also transfers with these smaller weights. Along a batch, no remaining coefficient changes. For $\widetilde{g}_i := \partial_i \mathcal{F}_{\mathrm{s}}/\sqrt{\kappa_i}$, the forcing bounds and relative scalar comparisons give

$$|\widetilde{g}_i'| \le P_n a(s) + P_n\sqrt{n}|\widetilde{g}_i| a(s), \qquad a(s) := e(\mathsf{P}') + \mathcal{E}(\mathsf{Q}').$$

Until $|\widetilde{g}_i| = h_n/(2\sqrt{n})$, the right side is at most $P_n a(s)$. Thus a batch with drop below $1/(P_n' n)$, for a sufficiently large fixed polylogarithm P_n', cannot turn the low exit into a successful high test. It

also keeps the metric comparable by the same first-exit argument. The tiny-slot final correction is chosen below this separation margin. Larger batch drops can occur only $\widetilde{O}(n)$ times. A subsequent sweep must otherwise be preceded by a curvature move, a regular endpoint, or a phase change.

There are $\widetilde{O}(n)$ curvature moves by Eq. (190), $\widetilde{O}(n)$ regular endpoints by their unchanged matrix-potential decrease, and $O(\log n)$ phases. It follows that there are $\widetilde{O}(n)$ sweeps and batches, all deterministically. Summing the stage bounds by Cauchy–Schwarz gives $\widetilde{O}(n)$ batch stages. Regular endpoints and the long curvature trials use at most $\widetilde{O}(n^{3/2})$ initialization stages in total. Coordinate updates cost $\widetilde{O}(n^4)$. The trace cutoff, recorded coefficients, and final deterministic rounding have the same discrepancy allowances as before.

Root, Hessian, filter, matrix-exponential, and sign comparisons all have polynomial condition bounds and inverse-polynomial slack. Use the anchored enclosures of Section 7.5, with sufficiently fine common grids for the new comparisons and finite-step decrease. The additional response congruences have polynomial inverse norms by the primal and dual lower bounds. The extra reservoir injection is below $e_{\mathrm{end}}/4$, smoothing costs at most 10^{-8}, and all numerical and endpoint errors can still be kept below $e_{\mathrm{end}}/2$. Their sum is below the original $e_{\mathrm{end}} = 10^{-6}$ allowance. The discrepancy constant remains 3.3443, with partition constant 1.67215.

A deterministic grouped reduction with dimensional slack. The preceding construction can be accelerated without changing its budget, discrepancy allowances, or curvature step. We first remove the preprocessing bottleneck. The notation D in the next lemma denotes a real vector-space dimension.

Lemma 5.9 (Deterministic reduction with a factor-two slack)**.** *Under the short-data hypotheses of Lemma 5.2, a deterministic algorithm returns $x \in [-1,1]^m$ with at most $2D$ fractional coordinates and $\|\sum_i x_i b_i\|_2 \le \eta$ using $\widetilde{O}(mD + D^{\omega_0})$ arithmetic and bit operations. The long input length is additive after the one-time input conversion.*

Proof. Use the group tree in Lemma 6.19. We replace only the random direction in its small reducer. At each such call there are $2D < r \le r_0 \le 4D$ active columns. With its current slack matrix, form Π_δ from Eq. (278). Its proof gives

$$0 \prec \Pi_\delta \preceq I, \qquad \mathrm{tr}[\Pi_\delta^2] \ge r - D, \qquad \|\mathsf{B}\Pi_\delta\| \le \sqrt{\delta}/2.$$

Use precisely the parameters $\nu, \delta, \ell, H, T_*$ and coefficient grid in Eqs. (281) and (282). The residual-certified ridge inverse used there also constructs an explicit symmetric dyadic matrix V with $\|V - \Pi_\delta\| \le \nu/\sqrt{r}$. All its products have order $O(D)$, so this costs $\widetilde{O}(D^{\omega_0})$.

For independent uniform signs ξ, consider the scalar estimator

$$\Psi(\xi) := \|V\xi\|_2^2 - \frac{1}{100}\sum_{i=1}^{r} \cosh((V\xi)_i). \tag{191}$$

Since $\nu \le 1/64$, we have $\|V\| \le 65/64$ and $\mathrm{tr}[V^\top V] \ge r - D - r/16$. Each row of V has squared norm at most $(65/64)^2$. The elementary inequality $\cosh t \le e^{t^2/2}$ therefore gives

$$\mathbb{E}[\Psi] \ge r - D - r/16 - r/50 > \frac{3}{4}(r - D).$$

Here $r < 2(r - D)$, and $e^{(65/64)^2/2} < 2$. Fix the signs successively by conditional expectation. If u is the current prefix sum, the two conditional expectations of its quadratic term are

$$\|u \pm V_j\|_2^2 + \sum_{h>j} \|V_h\|_2^2,$$

and the conditional expectation of row i's hyperbolic cosine is

$$\cosh(u_i \pm V_{ij}) \prod_{h>j} \cosh(V_{ih}).$$

Maintain the prefix, the remaining squared norm, and these row products. Each of the r choices takes $\widetilde{O}(r)$ work. Evaluate a candidate to absolute error at most $1/(1000r)$ and choose the larger approximation. The total comparison loss is at most $1/500$, so the resulting signs satisfy $\Psi(\xi) \geq (r-D)/2$.

Put $\widetilde{v} := V\xi$, computed exactly on its dyadic grid. Then $\|\widetilde{v}\|_2^2 \geq (r-D)/2$ and $\|\widetilde{v}\|_2^2 \leq 2r$. Positivity of Ψ implies $\cosh(\widetilde{v}_i) \leq 200r$, whence

$$\|\widetilde{v}\|_\infty \leq \log(400r) < \log_2(64r) \leq \ell, \qquad \|\widetilde{v} - \Pi_\delta \xi\|_2 \leq \nu.$$

For the strict logarithmic inequality, use $\log 400 < 6$ and $\log r \leq \log_2 r$. Thus both exact tests in Eq. (279) always pass.

Section 7.6 gives the logarithmic-precision implementation of these conditional comparisons.

The deterministic progress, endpoint-freezing, residual, and grid proofs in Lemma 6.18 apply to every passing direction. They give fewer than T_* steps and residual below $\varepsilon/128$. Its short-precision barrier sign rule and final residual check apply without change. No trials or fallback are required. The cost of a small call is therefore deterministically $\widetilde{O}(D^{\omega_0})$ in both models. In the balanced group tree, retain its inherited coefficients and allocate $\varepsilon = \eta/(M_{\mathrm{grp}} R_{\mathrm{tree}})$ per level. The group-sum preparation costs $\widetilde{O}(mD)$, there are only logarithmically many levels, and their unnormalized residuals sum to at most η. The common grids have the same uniform logarithmic length proved in Lemma 6.19. This proves the claim. □

For the matrix problem apply this lemma with $D = n^2$ after the same input conversion. It leaves at most $2n^2$ fractional slots. Pad all retained matrices by zeros to order $N := 2^{\lceil \log_2(2n) \rceil} < 4n$, and run the matrix algorithm with this working order. Then $k \leq 2n^2 \leq N^2$. All variances, operator norms, residual allowances, and returned signs are preserved. In particular use $\rho = 10^{-8}/N$, so $N\rho = 10^{-8}$ is unchanged. The initial potential bound and every discrepancy allowance therefore remain the same. This deterministic grouped step costs $\widetilde{O}(mn^2 + n^{2\omega_0})$. Below we again write n for the working order; the original and working orders differ by less than a factor four.

Solving all response columns at one root. The primal and dual responses of all coordinate directions share their optimizer and their linear operator. We compute them together.

Lemma 5.10 (Block response construction)**.** *At a corrected root with $k \leq n^2$ active coordinates, the scaled Hessian and all four response maps in Eq. (183), composed with the deterministic filter, can be computed to any prescribed inverse-polynomial accuracy in*

$$\widetilde{O}(n^3 + kn^{\omega_0} + n^2 k^{\omega_0 - 1} + k^{\omega_0}) \tag{192}$$

arithmetic and bit operations. The same bound includes the matrix conditional selection producing Eq. (186).

Proof. Write Hermitian matrices in real orthonormal coordinates and let Φ have columns $\mathrm{vec}(C_i)$. Its row dimension is $O(n^2)$. Rank one gives $C_i X C_i = \mathrm{tr}[C_i X] C_i$, so the coupling in these coordinates is

$$\Phi \,\mathrm{diag}(\iota\beta_i) \Phi^\top.$$

An application to k right-hand sides consists of products of shapes $(k, O(n^2), k)$ and $(O(n^2), k, k)$. Dividing the long dimension into blocks of k gives cost $\widetilde{O}(n^2 k^{\omega_0 - 1})$. Explicitly forming Φ and all rank-one derivative columns costs $O(n^2 k)$, already covered by this bound.

Prepare the gap-free spectral decompositions for the two Lyapunov inverses and the Gibbs covariance exactly as in the proof of Eq. (204). Their total setup cost is $\widetilde{O}(n^3)$ at this root. For each of the k right-hand sides, the remaining operations are a fixed number of left and right order-n products, basis changes, entrywise multiplication by divided-difference coefficients, and trace centering. They cost $\widetilde{O}(kn^{\omega_0})$ together. The decompositions are reused for every column and every iteration at this root.

By Eq. (203), the Chebyshev inverse for $J_{\mathrm{s}} = \mathscr{E} + \mathscr{J} Q_{\mathsf{Q}} \mathscr{J}$ has polylogarithmic degree after $\mathscr{E}$ preconditioning. Apply this one fixed polynomial simultaneously to the columns of $B = T - \mathscr{J} Q_{\mathsf{Q}} L$ in Eq. (180). Each iteration has the preceding block cost. Recover the dual response columns from Q_{Q} and these primal responses. The products $L^\top Q_{\mathsf{Q}} L$ and $B^\top J_{\mathrm{s}}^{-1} B$ have the same $\widetilde{O}(n^2 k^{\omega_0 - 1})$ bound. Diagonal scaling then gives the required Hessian, including its reservoir term. The filter costs $\widetilde{O}(k^{\omega_0})$. Composing each response map with this filter again costs $\widetilde{O}(n^2 k^{\omega_0 - 1})$, and all relative congruences cost $\widetilde{O}(kn^{\omega_0})$.

Section 7.6 proves that all response columns can be evaluated to their prescribed tolerance with logarithmic precision.

Finally form all variance matrices $\sum_j D_{\ell j}^2$ in $\widetilde{O}(kn^{\omega_0})$ work. The matrix conditional estimator used above has order-n Hermitian arguments of norm $O(\log n)$ at every accepted prefix and either next candidate. Indeed its positive summands bound both signed exponentials, its remaining variance has norm at most $O(\log n)$, and a single signed increment has bounded norm by Eq. (184). Use the factorial Taylor evaluation in Lemma 6.20, with its error target tightened by the polynomial weights in the estimator. Each of the k sign choices costs $\widetilde{O}(n^{\omega_0})$. Scalar conditional quadratic and hyperbolic-cosine terms take $\widetilde{O}(k^2)$ additional work. All comparisons retain the existing slack, proving Eq. (192) in both models. □

6 Root continuation and randomized numerical tools

Sections 6.1–6.5 develop the smoothed potential, root continuation, and local moves. Sections 6.6–6.8 construct the randomized numerical routines and certify their outputs. The subsequent estimates give larger curvature moves, grouped reduction, and amortized sweeps. Finally, Sections 6.13–6.16 combine masked curvature with delayed deletion and bound the cost of matrix functions and recursive solves.

The primal matrices are U, V and the dual matrices are W, Z. Define their pairs by $\mathsf{P} := (U, V)$ and $\mathsf{Q} := (W, Z)$. All scalar products and traces on pairs sum over both blocks. Fix the multiplication parameters in Eq. (132). The notation P_n denotes a fixed polynomial in $\log(2n)$, enlarged a finite number of times; constants depend only on the fixed budget and regularization parameters. All numerical tolerances below are specified by these uniform bounds.

6.1 Smoothed potential and retained geometry

We analyze a retained instance of order n with $k \le n^2$ active coordinates. The grouped preprocessing in Lemma 5.9 leaves at most $2n^2$ coordinates before padding; its padded order $N < 4n$ satisfies $k \le N^2$. The local estimates below apply at that padded order. The retained matrices are PSD and rank one, and satisfy

$$\sum_i V_i^2 \preceq I, \qquad \|\sum_i a_i V_i\|_{\mathrm{F}} \le \|a\|_2.$$

The single preprocessing residual has operator norm at most $\eta_{\mathrm{ps}} = 10^{-7}$. The deterministic grouped routine in Lemma 5.2 is available for fallback. To see the compatibility with the new potential, let y^0 be the partial signing and $e_0 = \sum_i y_i^0 C_i$. Keep the true retained matrices $V_i = C_i$ and subtract e_0 from the frozen contribution. Then the modeled sum starts at zero, and its difference from the true signed sum remains exactly e_0. The residual is charged once at the end. The grouped procedure changes coefficients only; it preserves rank one, positivity, and the variance bound on the retained inputs. The Frobenius bound above follows by considering the nonnegative Gram matrix $J_{ij} = \mathrm{tr}[V_i V_j]$: with $\nu_i = \mathrm{tr}[V_i] > 0$, $J\nu \le \nu$ follows from $\sum_i V_i^2 \preceq I$, so $\|J\| \le 1$. The common-grid construction and its short-precision bounds are unchanged. Stop directly if $k = 0$. Fix

$$\iota = 14779/2500, \quad \varepsilon_{\mathrm{scale}} = 10^{-8}, \quad \rho = \varepsilon_{\mathrm{scale}}/n, \quad \gamma = 10^{-6}/k, \quad \varepsilon_{\mathrm{round}} = \varepsilon_{\mathrm{discard}} = 10^{-6}.$$

Use the degree-20 budget β_{fast} from Section 4, $\beta(y) = (1-y^2)^{97/200} f(|y|)$. These reference estimates take γ fixed. The final implementation uses the phase values in Eq. (268), keeping γ fixed within every move and attenuation path. Their additional logarithmic factors are absorbed in P_n. Define $L_{\log} = \lceil \log_2(4n) \rceil$ and retain a dyadic δ_{end} with

$$\frac{{\varepsilon_{\mathrm{round}}}^2}{8192 L_{\log}} < \delta_{\mathrm{end}} \le \frac{{\varepsilon_{\mathrm{round}}}^2}{4096 L_{\log}}, \qquad \tau_{\mathrm{s}} = 10^{-8}/L_{\log}.$$

The smoothing temperature τ_{s} is unrelated to the budget parameter τ in the scalar certificate. At endpoint distance at least $\delta_{\mathrm{end}}/2$,

$$-\beta'' > 1/2, \quad |\beta^{(j)}| \le P_n(-\beta''), \quad |\beta^{(j)}|^2/\beta \le P_n(-\beta'') \quad (1 \le j \le 3). \tag{193}$$

Second derivatives are continuous at zero; higher derivatives below are one-sided there.

On a face, let $M(y) = F + \sum_i y_i V_i$ and $\mathcal{K}_y(D) = \iota \sum_i \beta(y_i) V_i D V_i$. With $\mathsf{P} = (U, V)$, $E = (I, I)$, define

$$\mathcal{C}(\mathsf{P}, y) = (U^{-1} + M + \mathcal{K}_y(V),\ V^{-1} - M + \mathcal{K}_y(U)).$$

Pairs are identified with block-diagonal matrices, and traces and Frobenius inner products sum over the two blocks. Define

$$\begin{aligned} f_{\tau_{\mathrm{s}}}(C) &= \tau_{\mathrm{s}} \log \mathrm{tr}[\exp(C/\tau_{\mathrm{s}})], \\ \mathcal{P}_{\mathrm{s}}(y) &= \min_{U,V \succ 0} \{\rho\, \mathrm{tr}[U+V] + f_{\tau_{\mathrm{s}}}(\mathcal{C}(\mathsf{P}, y))\}, \\ \mathcal{F}_{\mathrm{s}}(y) &= \mathcal{P}_{\mathrm{s}}(y) + \gamma \sum_i \beta(y_i). \end{aligned} \tag{194}$$

For the original potential $\mathcal{P}$,

$$\mathcal{P} \le \mathcal{P}_{\mathrm{s}} \le \mathcal{P} + \tau_{\mathrm{s}} \log(2n) \le \mathcal{P} + 10^{-8}, \qquad \mathcal{P}_{\mathrm{s}} \ge \|M\|. \tag{195}$$

Indeed these are the eigenvalue/log-sum-exp inequalities before taking the infimum over P.

The objective is convex: inversion is operator convex, and $f_{\tau_{\mathrm{s}}}$ is convex and PSD-order increasing. Bounded sublevels are compact inside the positive cone, so a minimizer exists. Its Gibbs state

$$\mathsf{Q} = (W, Z) = \exp(\mathcal{C}/\tau_{\mathrm{s}}) / \mathrm{tr}[\exp(\mathcal{C}/\tau_{\mathrm{s}})]$$

satisfies

$$\mathrm{tr}[\mathsf{Q}] = 1, \quad \mathscr{J}\mathsf{Q} = \rho E, \quad W = U\mathcal{K}(Z)U + \rho U^2, \quad Z = V\mathcal{K}(W)V + \rho V^2, \tag{196}$$

where

$$\mathscr{J}(D_U, D_V) = (U^{-1}D_U U^{-1} - \mathcal{K}(D_V),\ V^{-1}D_V V^{-1} - \mathcal{K}(D_U)).$$

The positive-map argument preceding Eq. (87) makes $\mathscr{J}$ invertible with positive inverse. It uses the strict dual identities, not hard tight constraints. The objective Hessian in P is positive definite, as shown below; thus the optimizer is unique and is C^2 on every open face, including coordinate zeros.

Uniformly at roots on bounded-model sublevels, scalar feasible matrices and Eq. (196) give

$$\begin{aligned} U, V \succeq a_0 I, \quad \|U^{-1}\|, \|V^{-1}\|, \|\mathcal{K}(U)\|, \|\mathcal{K}(V)\| \le C, \\ W \succeq \rho U^2, \quad Z \succeq \rho V^2, \quad \rho \operatorname{tr}[U^2 + V^2] \le 1, \quad W, Z \succeq a_0^2 \rho I. \end{aligned} \tag{197}$$

Here and below zero inputs may be omitted. These bounds also hold along the data homotopy in Section 6.4.

Let $\mathscr{E}$ be the Hessian in P of $\langle \mathsf{Q}, \mathcal{C} \rangle$ with Q fixed. Define

$$\widehat{D}_U = U^{-1/2} D_U U^{-1/2}, \quad T_U = U^{-1/2} W U^{-1/2}, \quad e(D)^2 = \langle D, \mathscr{E} D \rangle = 2 \sum_{A=U,V} \operatorname{tr}[T_A \widehat{D}_A^2].$$

The dual identities and the bounds above give

$$\begin{aligned} \mathscr{E} \succeq a n^{-3/2} I, \quad \mathscr{E} \preceq C I, \quad \|D\|_{\mathrm{F}} \le C n^{3/4} e(D), \\ \|\widehat{D}_A\| \le C\sqrt{n} e(D), \quad a/n \le \lambda_{\min}(T_A), \quad \sum_A \operatorname{tr}[T_A] \le C. \end{aligned} \tag{198}$$

The weighted order norm is

$$q(B) = \inf\{\langle \mathsf{Q}, A \rangle : A \succeq 0,\ -A \preceq B \preceq A\} = \sum_A \|\mathsf{Q}_A^{1/2} B_A \mathsf{Q}_A^{1/2}\|_*.$$

The equality follows by congruence and minimizing the trace of a PSD majorant. In particular $q(\mathcal{C}_{\mathsf{PP}}[D, D]) = e(D)^2$.

Entropy metric and an inexpensive root inverse On the trace-zero space $\mathcal{T} = \{\mathsf{B} : \operatorname{tr}[\mathsf{B}] = 0\}$ define

$$B_{\mathsf{Q}} = \tau_{\mathrm{s}} \Pi_{\mathcal{T}} D \log_{\mathsf{Q}} |_{\mathcal{T}}, \qquad \mathcal{E}(\mathsf{B})^2 = \langle \mathsf{B}, B_{\mathsf{Q}} \mathsf{B} \rangle.$$

Its inverse is the Gibbs covariance operator

$$Q_{\mathsf{Q}}(B) = \frac{1}{\tau_{\mathrm{s}}} \Big(\int_0^1 \mathsf{Q}^t B \mathsf{Q}^{1-t} dt - \operatorname{tr}[\mathsf{Q}B]\mathsf{Q} \Big). \tag{199}$$

Adding a scalar multiple of E to its argument has no effect. In a basis diagonalizing Q, the integral has logarithmic-mean coefficients $\mathsf{L}(z_i, z_j) \le (z_i + z_j)/2$. Hence

$$\langle B, Q_{\mathsf{Q}} B \rangle \le \tau_{\mathrm{s}}^{-1} \operatorname{tr}[\mathsf{Q}B^2], \qquad |\langle B, \mathsf{B} \rangle| \le \tau_{\mathrm{s}}^{-1/2} \sqrt{\operatorname{tr}[\mathsf{Q}B^2]}\, \mathcal{E}(\mathsf{B}). \tag{200}$$

The inverse relation also follows by differentiating the normalized matrix exponential. At roots, B_{Q} has eigenvalues between τ_{s} and $P_n n$.

The following stronger relative bound will be used repeatedly:

$$\|\mathsf{Q}^{-1/2} \mathsf{B} \mathsf{Q}^{-1/2}\| \le (\tau_{\mathrm{s}} \lambda_{\min}(\mathsf{Q}))^{-1/2} \mathcal{E}(\mathsf{B}) \le P_n \sqrt{n}\, \mathcal{E}(\mathsf{B}). \tag{201}$$

Indeed, if $\mathsf{B} = \mathsf{Q}^{1/2}\widehat{\mathsf{B}}\mathsf{Q}^{1/2}$, the energy coefficient of $|\widehat{\mathsf{B}}_{ij}|^2$ is $\tau_{\mathrm{s}} z_i z_j/\mathsf{L}(z_i, z_j)$, which is at least $\tau_{\mathrm{s}} \min(z_i, z_j)$.

For completeness the root preconditioner has a dimension-free bound apart from τ_{s}^{-1}. Complete positivity applied to $\left(\begin{smallmatrix} D_V V^{-1} D_V & D_V \\ D_V & V \end{smallmatrix}\right) \succeq 0$ gives $\mathcal{K}(D_V)^2 \preceq \|\mathcal{K}(V)\|\mathcal{K}(D_V V^{-1} D_V)$. Pair with W and use $\mathcal{K}(W) \preceq V^{-1} Z V^{-1}$. Together with $U^{-2} \preceq \|U^{-1}\| U^{-1}$ and $(A-B)^2 \preceq 2A^2 + 2B^2$, this proves

$$\operatorname{tr}[\mathsf{Q}(\mathscr{J} D)^2] \le K e(D)^2, \qquad K = \max\{\|U^{-1}\| + \|\mathcal{K}(U)\|,\ \|V^{-1}\| + \|\mathcal{K}(V)\|\} \le C. \tag{202}$$

Consequently the primal Schur complement obeys

$$\mathscr{E} \preceq J_{\mathrm{s}} := \mathscr{E} + \mathscr{J} Q_{\mathsf{Q}} \mathscr{J} \preceq (1 + K/\tau_{\mathrm{s}})\mathscr{E}. \tag{203}$$

This also proves the strict convexity and regularity asserted earlier.

To apply $\mathscr{E}^{-1}$, its U equation becomes the Lyapunov equation

$$\widehat{D}_U T_U + T_U \widehat{D}_U = U^{1/2} F_U U^{1/2}.$$

Diagonalize the order-n T_U and divide entry (i, j) by $t_i + t_j$; all denominators are inverse-polynomial. The V block is identical. A gap-free approximate Hermitian diagonalization, with residual and orthogonality bounds, costs $\widetilde{O}(n^3)$. It requires no accurate choice of individual eigenvectors at a repeated eigenvalue. Evaluating Q_{Q} uses the same kind of decomposition and scalar logarithmic means. For nearby eigenvalues use the divided exponential series, not cancellation of nearly equal logarithms.

The rectangular products of Lemma 5.6, or square products on groups of at most n factors, give cost

$$F(n, k) := \widetilde{O}(\min\{n^{\chi}, n^3 + (k+n)n^{\omega_0 - 1}\}). \tag{204}$$

This covers coupling actions, signed sums and complete scalar score scans. In particular, an n-by-n^2 factor matrix reduces these actions and their adjoints to products of shapes (n, n, n^2) and (n, n^2, n). Tensor permutation preserves the exponent χ and the fixed recursive error bounds. The $\widetilde{O}(n^3)$ spectral work is smaller than $\widetilde{O}(n^{\chi})$. It applies to a rational representation $V_i = c_i a_i a_i^*$ without extracting exact square roots. By Eq. (203), a Chebyshev solve for J_{s} uses only poly-logarithmically many $\mathscr{E}^{-1}$-preconditioned actions. Similarity to its symmetric normalization has polynomial condition number by Eq. (198). Second-kind Chebyshev bounds control accumulated application errors; the numerical details are specified in Sections 7.5 and 7.8.

Covariance and endpoint moves A regular move to $+1$ changes the two $\mathcal{C}$ blocks, at fixed P, by

$$(1 - y_i - \iota\beta_i v_i)V_i, \qquad (-(1 - y_i) - \iota\beta_i u_i)V_i.$$

Both are PSD-nonpositive when the endpoint test holds. Order monotonicity of $f_{\tau_{\mathrm{s}}}$ proves nonincrease. The other sign is symmetric. A deferred removal preserves M and removes a PSD coupling term. Conservative score errors are repaired by a uniform eE, whose cost is exactly e because $f_{\tau_{\mathrm{s}}}(C + eE) = f_{\tau_{\mathrm{s}}}(C) + e$. The same reservoir margin therefore pays for the endpoint-score errors.

We do not impose $\mathcal{C} = bE$ on the new root. Instead the inverse function theorem for $\mathscr{J}$ supplies a local comparison path with $\mathcal{C}(\mathsf{P}(t), y + tr) = \mathcal{C}(\mathsf{P}(0), y)$ fixed. Its first-order response satisfies Eq. (87). Along it $f_{\tau_{\mathrm{s}}}$ is constant and $\rho \operatorname{tr}[\mathsf{P}(t)] + f_{\tau_{\mathrm{s}}}(\mathcal{C}(\mathsf{P}(0), y))$ majorizes $\mathcal{P}_{\mathrm{s}}$ with first-order equality. Its second variation is exactly the inversion-energy and mixed-coupling expression preceding Eq. (89). Lemma 4.3 uses that response equation, the dual identities, rank one and

lightness, all valid here. The unchanged exact scalar certificate therefore yields the same strict generator surplus.

Define $u_i := \operatorname{tr}[UV_i]$, $v_i := \operatorname{tr}[VV_i]$, $w_i := \operatorname{tr}[WV_i]$, $z_i := \operatorname{tr}[ZV_i]$, and $t_i := v_i w_i + u_i z_i$. Thus $p_i = 1 + \iota\beta_i' v_i$ and $q_i = 1 - \iota\beta_i' u_i$ are the two response coefficients. Define

$$\varpi_i = (-\beta_i'')(\iota t_i + \gamma), \quad G = \operatorname{diag}(\varpi_i), \quad A_r = \iota \sum_i (-\beta_i'') t_i r_i^2, \quad \theta = r^\top G r.$$

At a light state $\gamma/4 \le \varpi_i \le P_n$ and $\theta \ge 1/(P_n k)$. The drift estimate in Lemma 5.4 $(C_{2,i} M_i)^2 \le P_n n \varpi_i$ follows from $w_i \ge \rho u_i^2 / \operatorname{tr}[V_i]$ and its partner, so remains valid. Dividing the strict generator estimate by its nonzero covariance proves the numerically separated alternative

$$|(\mathcal{F}_{\mathrm{s}})_i| \ge P_n^{-1} \sqrt{\varpi_i / n} \quad \text{for some } i, \quad \text{or} \quad r^\top \nabla^2 \mathcal{F}_{\mathrm{s}} r \le -P_n^{-1} \theta. \tag{205}$$

This is the proof of Lemma 5.4 applied to the fixed-constraint comparison path for $\mathcal{P}_{\mathrm{s}}$.

6.2 Joint-energy Newton continuation

We use the joint primal and entropy metrics to control a Newton correction, and keep the inverse at the base root fixed during each short continuation step.

Work with $\operatorname{tr}[\mathsf{Q}] = 1$ and residual

$$\mathcal{E}_{\mathrm{KKT}y}(\mathsf{P}, \mathsf{Q}) = (\rho E - \mathscr{J}_{\mathsf{P}} \mathsf{Q},\ \Pi_{\mathcal{T}}(\tau_{\mathrm{s}} \log \mathsf{Q} - \mathcal{C}(\mathsf{P}, y))). \tag{206}$$

Its derivative on matrix pairs A and trace-zero pairs B is

$$\mathbb{A} = \begin{pmatrix} \mathscr{E} & -\mathscr{J}|_{\mathcal{T}} \\ \Pi_{\mathcal{T}} \mathscr{J} & B_{\mathsf{Q}} \end{pmatrix}. \tag{207}$$

The cross terms are skew-adjoint. At a root let $\mathbb{D} = \operatorname{diag}(\mathscr{E}, B_{\mathsf{Q}})$ and $\|(\mathsf{A}, \mathsf{B})\|_{\mathbb{D}}^2 = e(\mathsf{A})^2 + \mathcal{E}(\mathsf{B})^2$. Then

$$\mathbb{A} + \mathbb{A}^* = 2\mathbb{D}, \qquad \|\mathbb{D}^{1/2} \mathbb{A}^{-1} \mathbb{D}^{1/2}\| \le 1. \tag{208}$$

The second assertion follows by pairing $\mathbb{A} y_{\mathrm{aux}} = b_{\mathrm{loc}}$ with y_{aux} and Cauchy–Schwarz. Norms of $\mathbb{D}^{1/2}$ and its inverse are polynomial.

A uniform Newton neighborhood ensures that approximate optimizers can be corrected during continuation.

Write $\mathsf{s} = (\mathsf{P}, \mathsf{Q})$ for the optimization variables.

Lemma 6.1 (Joint-energy Newton ball)**.** *At every bounded-model root there is a Newton ball*

$$\|\mathsf{s} - \mathsf{s}_*\|_{\mathbb{D}_*} \le a / (P_n \sqrt{n})$$

inside the positive cones and the trace-one affine space. On a smaller such ball, exact Newton iteration satisfies

$$\|\mathsf{s}_{\mathrm{new}} - \mathsf{s}_*\|_{\mathbb{D}_*} \le P_n \sqrt{n} \|\mathsf{s} - \mathsf{s}_*\|_{\mathbb{D}_*}^2. \tag{209}$$

Moreover

$$\|\mathbb{D}_*^{-1/2} (\mathbb{A}(\mathsf{s}) - \mathbb{A}(\mathsf{s}_*)) \mathbb{D}_*^{-1/2}\| \le P_n \sqrt{n} \|\mathsf{s} - \mathsf{s}_*\|_{\mathbb{D}_*}. \tag{210}$$

All comparisons and the Newton radius are uniform in the face and in real or complex Hermitian inputs.

Proof. The primal relative bound in Eq. (198) and the dual one in Eq. (201) keep U, V, Q within fixed Loewner factors of their root values in the stated ball. In particular $\mathsf{Q} \succeq a/n$, and $W \succeq (a/n)U$, $Z \succeq (a/n)V$, throughout it: these last inequalities follow by comparing W, U separately to their root values. Hence the bounds $T_A \succeq a/n$ and $\sum \operatorname{tr}[T_A] \le C$ remain valid there.

We detail the differential estimates giving metric comparison, rather than assuming it. For fixed test D, variation in P gives $\langle \mathsf{Q}, \mathcal{C}_{\mathsf{PPP}}[\Delta\mathsf{P}, D, D]\rangle$. Its three normalized inverse words are bounded by

$$C\sqrt{n}e(\Delta\mathsf{P})e(D)^2.$$

For an exterior $\widehat{\Delta\mathsf{P}}$ use weighted Cauchy–Schwarz and $\operatorname{tr}[T\widehat{D}^4] \le \|\widehat{D}\|^2 \operatorname{tr}[T\widehat{D}^2]$. For the middle word use $\widehat{D}T\widehat{D} \succeq 0$. Variation in Q contributes $\langle \Delta\mathsf{Q}, \mathcal{C}_{\mathsf{PP}}[D, D]\rangle$, of absolute value at most $\|\mathsf{Q}^{-1/2}\Delta\mathsf{Q}\mathsf{Q}^{-1/2}\| e(D)^2$. Thus $|d\mathscr{E}[D, D]| \le P_n\sqrt{n}\|(\Delta\mathsf{P}, \Delta\mathsf{Q})\|_{\mathbb{D}} e(D)^2$.

For the entropy block, define $R_q = (\mathsf{Q} + qI)^{-1}$. The resolvent formula gives

$$\langle \mathsf{B}, D\log_{\mathsf{Q}}[\mathsf{B}]\rangle = \int_0^\infty \operatorname{tr}[R_q\mathsf{B}R_q\mathsf{B}]dq,$$

and differentiation in $\Delta\mathsf{Q}$ bounds the derivative of this form by

$$2\|\mathsf{Q}^{-1/2}\Delta\mathsf{Q}\mathsf{Q}^{-1/2}\|\langle \mathsf{B}, D\log_{\mathsf{Q}}[\mathsf{B}]\rangle. \tag{211}$$

Indeed the differentiated integrand is minus twice the real trace of $(R_q^{1/2}\Delta\mathsf{Q}R_q^{1/2})(R_q^{1/2}\mathsf{B}R_q^{1/2})^2$. The first factor has operator norm at most the displayed relative norm, and the squared factor is PSD. No commutation is used.

Finally, the derivative of the cross block $\mathscr{J}$ in $\Delta\mathsf{P}$ pairs with B as $\langle \mathsf{B}, \mathcal{C}_{\mathsf{PP}}[\Delta\mathsf{P}, D]\rangle$. Its q-majorant is at most $Ce(\Delta\mathsf{P})e(D)$, by the matrix Young inequality. Equation Eq. (201) bounds the pairing by $P_n\sqrt{n}\mathcal{E}(\mathsf{B})e(\Delta\mathsf{P})e(D)$. These prove the normalized differential bound for $\mathbb{A}$, and the relative derivative bounds for both diagonal metric blocks.

A first-exit argument along the straight segment from s_* keeps $\mathbb{D}_{\mathsf{s}}$ within constant factors of $\mathbb{D}_*$ in the stated ball: until exit, integrate the preceding differential bounds, and choose a small. The direct Loewner comparisons already ensure positivity and the T_A bounds needed along the segment. Integrating the cross estimates now gives Eq. (210). By Eq. (208) and a Neumann perturbation, the inverse Jacobian in the fixed root metric has norm at most two. The integral Taylor remainder for $\mathcal{E}_{\mathrm{KKT}_y}$ proves Eq. (209). These arguments use only two derivatives of $\mathcal{C}$ in P and the smooth log map; coordinate zeros do not enter this fixed-parameter Newton argument. □

Let $\mathsf{A} = \mathsf{P}'$ and $\mathsf{B} = \mathsf{Q}'$ on a path of exact roots. Differentiating the two root equations gives

$$\mathscr{J}\mathsf{A} + \tau_{\mathrm{s}}D\log_{\mathsf{Q}}[\mathsf{B}] + d_{\mathrm{norm}}E = F_1, \quad \mathscr{J}\mathsf{B} - \mathscr{E}\mathsf{A} = F_2, \quad \operatorname{tr}[\mathsf{B}] = 0, \qquad F_1 = \mathcal{C}_t, \quad F_2 = \mathcal{A}'\mathsf{Q}.$$

Pairing the equations gives

$$e(\mathsf{A})^2 + \mathcal{E}(\mathsf{B})^2 = \langle F_1, \mathsf{B}\rangle - \langle F_2, \mathsf{A}\rangle. \tag{212}$$

The scalar d_{norm} is only the log-normalization derivative and need not be stored as an unknown.

Controlling metric and Jacobian variation allows correction systems along a short root path to reuse a fixed approximate inverse.

Lemma 6.2 (Anchored transport and executable correction)**.** *Along a unit-direction local root path, write $a(t) = e(\mathsf{A}) + \mathcal{E}(\mathsf{B})$. The metric and joint Jacobian satisfy*

$$\|\mathbb{D}^{-1/2}\mathbb{D}'\mathbb{D}^{-1/2}\|, \ \|\mathbb{D}^{-1/2}\mathbb{A}'\mathbb{D}^{-1/2}\| \le P_n\sqrt{n}(a(t) + \sqrt{\theta(t)}). \tag{213}$$

For the logarithmic data homotopy of Section 6.4, replace $\sqrt{\theta}$ by 1. If the integral of the right side over a path is sufficiently small, its base root is a valid Newton start at the end. A fixed approximate inverse at that base solves every correction system along the path in $\widetilde{O}(F(n,k))$ work. It is an implicitly applied inverse, not a stored dense matrix.

Proof. The metric derivatives and the optimizer-dependent part of $\mathbb{A}'$ are bounded in Lemma 6.1. The additional explicit coupling term in $\mathscr{J}'$ pairs with $\mathsf{B}_0 \in \mathcal{T}$ and D as $\langle \mathsf{B}_0, \mathcal{A}'D\rangle$. Lemma 5.5, specifically Eq. (156), gives $q(\mathcal{A}'D) \le P_n\sqrt{\theta}e(D)$; use Eq. (201) on B_0. This proves Eq. (213). On the homotopy, the relative coupling bound and the dual identities instead give $q(\mathcal{A}'D) \le Ce(D)$.

If a_{int} is the accumulated bound, integration gives $e^{-a_{\mathrm{int}}}\mathbb{D}_0 \preceq \mathbb{D}_t \preceq e^{a_{\mathrm{int}}}\mathbb{D}_0$ and

$$\|\mathbb{D}_0^{-1/2}(\mathbb{A}_t - \mathbb{A}_0)\mathbb{D}_0^{-1/2}\| \le e^{a_{\mathrm{int}}}a_{\mathrm{int}}.$$

Use Eq. (208) to get constant contraction for refinement with $\mathbb{A}_0^{-1}$. A sufficiently short energy arc, measured in the successive root metrics, also bounds the endpoint distance in its root metric. Take it below the radius in Lemma 6.1. That lemma adds a small perturbation for each off-root Newton iterate, so the same base inverse still gives contraction at most, say, 1/2.

At the base, solving $\mathbb{A}_0(\mathsf{A},\mathsf{B}) = (r_{\mathrm{aux}}, t)$ reduces to

$$J_{\mathrm{s},0}\mathsf{A} = r_{\mathrm{aux}} + \mathscr{J}_0 Q_0 t, \qquad \mathsf{B} = Q_0(t - \mathscr{J}_0\mathsf{A}). \tag{214}$$

Eq. (203) and the explicit Lyapunov inverse implement this in $\widetilde{O}(F(n,k))$. Build it from a sufficiently accurate base cache. For a current Jacobian $\widetilde{\mathbb{A}}$, use $y_{\mathrm{aux}j+1} = y_{\mathrm{aux}j} + \widetilde{A_{\mathrm{inv}0}}(b_{\mathrm{loc}} - \widetilde{\mathbb{A}}y_{\mathrm{aux}j})$, where $\widetilde{A_{\mathrm{inv}0}}$ means applying Eq. (214) to its prescribed accuracy. The exact base inverse has contraction bounded away from one; polynomially small root, coefficient, and application errors preserve it. All changes of norm have polynomial condition. The error recurrence in the base norm is at most one half the preceding error plus a polynomial multiple of the additive evaluation tolerance. Polylogarithmically many refinements therefore attain the required precision. In particular we do not assume that the off-root current Lyapunov preconditioner satisfies Eq. (203) merely from constant Loewner proximity. The use of the base inverse and the joint transport estimate avoids that inference. □

6.3 Local moves, curvature tubes, and zeros

The proof of Eq. (156) uses only Eq. (196) and therefore still gives, on a unit direction,

$$q(\mathcal{A}^{(j)}D) \le P_n\sqrt{A_r}e(D), \qquad |\langle F_2, D\rangle| \le P_n\sqrt{A_r}e(D), \qquad q(\mathcal{C}_{t^j}) \le P_nA_r \ (j \ge 2). \tag{215}$$

Only $j \le 3$ is needed. The last bound uses $M^{(j)} = 0$ for $j \ge 2$. The second-order envelope identity now reads

$$\mathcal{P}_{\mathrm{s}}'' = -A_r + e(\mathsf{A})^2 + 2\langle F_2, \mathsf{A}\rangle + \mathcal{E}(\mathsf{B})^2. \tag{216}$$

The entropy term is positive, not an uncontrolled remainder.

Coordinates For $r = e_i$, $F_1 = (p_iV_i, -q_iV_i)$. On expanded light bounds, $|p_i|, |q_i| \le P_n$ and

$$\mathrm{tr}[\mathsf{Q}F_1^2] = \nu_i(p_i^2w_i + q_i^2z_i) \le P_n\nu_i(w_i + z_i) \le P_nt_i \le P_n\varpi_i, \qquad \nu_i = \mathrm{tr}[V_i].$$

Here $u_i, v_i \ge a_0\nu_i$. Apply Eq. (200), Eq. (215), and Eq. (212); then use Eq. (216):

$$e(\mathsf{A}) + \mathcal{E}(\mathsf{B}) \le P_n\sqrt{\varpi_i}, \qquad |\partial_i^2\mathcal{F}_{\mathrm{s}}| \le P_n\varpi_i. \tag{217}$$

Relative variation satisfies $|u_i'|/u_i, |v_i'|/v_i \le C\sqrt{n}e(\mathsf{A})$. For the dual scalars Eq. (200) and $w_i \ge a_0^2\rho\nu_i$ give $|w_i'|/w_i, |z_i'|/z_i \le P_n\sqrt{n}\mathcal{E}(\mathsf{B})$. Thus

$$|\varpi_i'| \le P_n(1+\sqrt{n\varpi_{i,0}})\varpi_i$$

until a fixed-factor exit. A first-exit argument, also for the expanded light bounds and endpoint distance, yields the tube

$$|t| \le a_n \min\{\delta_{\mathrm{end}}, (n\varpi_{i,0})^{-1/2}\}, \qquad \varpi_i(t) \asymp \varpi_{i,0}. \tag{218}$$

At zero the second derivatives are continuous, so this argument integrates through it. A zero input needs only the scalar reservoir and can also be rounded directly.

Curvature directions Consider a certified base with $\mathcal{F}_{\mathrm{s}}''(0) \le -P_n^{-1}\theta_0$. In a temporary region $\mathcal{P}_{\mathrm{s}}'' \le D_{\mathrm{aux}}\theta_0$ and $\theta \asymp \theta_0$, completing the square in Eq. (216) gives

$$e(\mathsf{A}) + \mathcal{E}(\mathsf{B}) \le P_n\sqrt{\theta_0}. \tag{219}$$

The needed third derivative is proved from the stationary saddle $\rho\,\mathrm{tr}[\mathsf{P}] + \langle \mathsf{Q}, \mathcal{C}\rangle - \tau_{\mathrm{s}}\,\mathrm{tr}[\mathsf{Q}\log\mathsf{Q}]$ on $\mathrm{tr}[\mathsf{Q}] = 1$. Let

$$B_2 = \mathcal{C}_{tt} + 2\mathcal{A}'\mathsf{A} + \mathcal{C}_{\mathsf{PP}}[\mathsf{A},\mathsf{A}], \qquad B_3 = \mathcal{C}_{ttt} + 3\mathcal{A}''\mathsf{A} + \mathcal{C}_{\mathsf{PPP}}[\mathsf{A},\mathsf{A},\mathsf{A}].$$

The envelope identity is

$$\mathcal{P}_{\mathrm{s}}''' = \langle \mathsf{Q}, B_3\rangle + 3\langle \mathsf{B}, B_2\rangle - \tau_{\mathrm{s}}D^3\,\mathrm{tr}[\mathsf{Q}\log\mathsf{Q}][\mathsf{B},\mathsf{B},\mathsf{B}]. \tag{220}$$

Terms containing second optimizer derivatives cancel by stationarity. Eq. (215) and normalized inverse words give $q(B_2) \le P_n\theta_0$ and $|\langle \mathsf{Q}, B_3\rangle| \le P_n(\theta_0 + \sqrt{n}\theta_0^{3/2})$. Eq. (201) bounds $|\langle \mathsf{B}, B_2\rangle|$ by $P_n\sqrt{n}\theta_0^{3/2}$. The resolvent estimate Eq. (211), with $\Delta\mathsf{Q} = \mathsf{B}$, bounds the last term in Eq. (220) by the same quantity. Finally differentiating $A_r = -\langle \mathsf{Q}, \mathcal{C}_{tt}\rangle$ and using Eq. (215) proves

$$|\mathcal{P}_{\mathrm{s}}'''|,\ |\mathcal{F}_{\mathrm{s}}'''|,\ |\theta'| \le P_n(\theta_0 + \sqrt{n}\theta_0^{3/2}). \tag{221}$$

These are one-sided inequalities at zeros.

At the base, $\mathcal{P}_{\mathrm{s}}'' \le \theta_0$. A simultaneous first-exit argument using Eq. (221) proves Eq. (219) and weight comparison throughout

$$|t| \le a_n \min\{\delta_{\mathrm{end}}, (n\theta_0)^{-1/2}\}. \tag{222}$$

This length is at most $\delta_{\mathrm{end}}/4$ after fixing its prefactor. It also keeps the model in a bounded-sum neighborhood. Coordinate-zero crossings cause no problem: $\mathcal{P}_{\mathrm{s}}''$ and θ are continuous, and their one-sided derivatives have the asserted absolute bound. We use a third-order remainder, not a fourth derivative or a Taylor jet crossing a jump.

The better of the two signs of Eq. (222) has decrease

$$\Delta_c \ge P_n^{-1}\min\{{\delta_{\mathrm{end}}}^2\theta_0, n^{-1}\}. \tag{223}$$

Indeed, in the symmetric average the linear terms cancel and the absolute third-order remainder is at most a small fraction of $P_n^{-1}\theta_0 t^2$. Since $\theta_0 \ge a/k$, the accepted count is

$$I_c \le \widetilde{O}(k+n). \tag{224}$$

Eq. (219) and Eq. (222) give an energy arc small enough for Lemma 6.2 over each whole candidate. The same base root starts both branches; each needs only a constant number of corrected stages, including final grid rounding. No $\widetilde{O}(n)$ tracking mesh remains.

For later use, Eq. (216) also gives a global lower bound $\nabla^2\mathcal{F}_{\mathrm{s}} \succeq -P_nG$ at a light state, by completing the square in $e(\mathsf{A})$ and retaining $\mathcal{E}(\mathsf{B})^2$. Eq. (217) gives bounded diagonal entries of the scaled Hessian. Therefore

$$A := G^{-1/2}\nabla^2\mathcal{F}_{\mathrm{s}}G^{-1/2} \succeq -P_nI, \qquad |A_{ii}| \le P_n, \qquad \|A\| \le P_nk. \tag{225}$$

The last implication uses positivity and the trace bound of $A + P_nI$.

6.4 Explicit initialization

For a face with modeled M_0 and coupling $\mathcal{K}_0$, use

$$M_{r_{\mathrm{hom}}} = r_{\mathrm{hom}}M_0, \qquad \mathcal{K}_{r_{\mathrm{hom}}} = r_{\mathrm{hom}}\mathcal{K}_0, \qquad 0 \le r_{\mathrm{hom}} \le 1,$$

in the smoothed problem, with the same ρ, τ_{s}. This is only an optimizer-evaluation homotopy; it never changes the signing or its potential comparisons. At $r_{\mathrm{hom}} = 0$ the exact root is

$$U_0 = V_0 = (2n\rho)^{-1/2}I, \qquad W_0 = Z_0 = I/(2n). \tag{226}$$

The constraints are scalar and the normalized exponential uniform; substitution in Eq. (196) proves stationarity and thus optimality. Scalar root evaluation approximates this to any prescribed precision.

The true and cached data satisfy $\|M_0\| \le 20$, $\mathcal{K}_0(I) \preceq 32I$. Taking $U = V = I$ bounds the objective uniformly along the homotopy. Thus Eq. (197) holds throughout it. In particular the Euclidean inverse of Eq. (207) is polynomially bounded, by Eq. (208); a conservative bound on $\mathsf{s}'_{r_{\mathrm{hom}}}$ is P_nn^2. The r_{hom}-forcing has Frobenius norm $O(\sqrt{n})$ and the inverse has norm $O(n^{3/2} + \tau_{\mathrm{s}}^{-1})$. An initial positive dyadic $r_{\mathrm{hom}*} \asymp a_nn^{-5}$ therefore makes Eq. (226) a valid Newton start at $r_{\mathrm{hom}*}$. To apply the anchored solve across this first step, one may also bound the explicit-data cross variation by P_nndr_{hom}: $\|\mathcal{A}_0D\|_{\mathrm{F}} \le C\|D\|_{\mathrm{F}}$, Eq. (198), and $\|\mathsf{B}\|_{\mathrm{F}} \le \tau_{\mathrm{s}}^{-1/2}\mathcal{E}(\mathsf{B})$ are more than sufficient. The stated $r_{\mathrm{hom}*}$ absorbs this bound.

For $r_{\mathrm{hom}} \ge r_{\mathrm{hom}*}$ parameterize by $t = \log r_{\mathrm{hom}}$. Then $M'_t = M_t$, $\mathcal{K}'_t = \mathcal{K}_t$. At a root, $F_1 = (M_{r_{\mathrm{hom}}} + \mathcal{K}_{r_{\mathrm{hom}}}(V), -M_{r_{\mathrm{hom}}} + \mathcal{K}_{r_{\mathrm{hom}}}(U))$ has operator norm bounded by a constant. Eq. (200) gives $|\langle F_1, \mathsf{B}\rangle| \le P_n\mathcal{E}(\mathsf{B})$. The dual identities give $|\langle \mathcal{A}_{r_{\mathrm{hom}}}\mathsf{Q}, D\rangle| \le Ce(D)$ by weighted Cauchy–Schwarz. The energy identity Eq. (212) proves

$$e(\mathsf{P}'_t) + \mathcal{E}(\mathsf{Q}'_t) \le P_n. \tag{227}$$

The explicit-data bound for transport is $q(\mathcal{A}_{r_{\mathrm{hom}}}D) \le Ce(D)$. Choose logarithmic parameter increments between fixed fractions of $a/(P_n\sqrt{n})$. Lemma 6.2 supplies one corrected stage per increment, at cost Eq. (204). The whole face requires

$$O(P_n\sqrt{n}\log(1/r_{\mathrm{hom}*})) = \widetilde{O}(\sqrt{n}) \tag{228}$$

stages. This supplies the first root. The final implementation uses Lemma 6.22 for regular endpoints and charged attenuation for deferred batches.

Store r_{hom} on a common dyadic grid, with mesh smaller than a fixed inverse-polynomial fraction of $r_{\mathrm{hom}*}/(P_n\sqrt{n})$. Round multiplicative increments so their logarithms stay between the chosen fractions, except for the final truncated increment. No exact product of all previous rational increments is retained. The parameter error is below the same root-transport margin. The initial face starts from Eq. (226). Later face changes use the separately proved continuation paths.

6.5 Charging face changes to potential decrease

The explicit homotopy above is used initially and after regular endpoint moves. Deferred removals are instead followed to their new face. We give a quantitative bound for these paths because merely accelerating the curvature search would leave the old initialization cost unchanged.

Lemma 6.3 (Regular endpoints and charged attenuation)**.** *There are $\widetilde{O}(n)$ regular endpoint moves. All deferred removals, including final corrections, use $\widetilde{O}(k+\sqrt{nk})$ corrected stages. All face work, including explicit initializations, therefore uses $\widetilde{O}(k+\sqrt{nk}+n^{3/2})$ stages of cost $\widetilde{O}(n^{\chi})$ in both operation models. These bounds hold for every sequence of certified local moves and for Hermitian inputs.*

Proof. For a regular move to 1, put $d_i = 1 - y_i \geq \delta_{\mathrm{end}}$, $\nu_i = \operatorname{tr}[V_i] \leq 1$, and $h_i = \iota\beta_i u_i + d_i$. After changing M by $d_i V_i$ and deleting the coupling slot, test the new problem at the old U and

$$V^{+} = (V^{-1} + h_i V_i)^{-1}.$$

The new minus block of $\mathcal{C}$ equals its old value. At an exact trigger the plus block does not increase: $\iota\beta_i v_i \geq d_i$, $V^{+} \preceq V$, and the remaining coupling is positive. Order monotonicity of $f_{\tau_{\mathrm{s}}}$ and Sherman–Morrison give

$$\mathcal{P}_{\mathrm{s,old}} - \mathcal{P}_{\mathrm{s,new}} \geq \rho \operatorname{tr}[V - V^{+}] \geq \rho \frac{h_i v_i^2}{\nu_i(1 + h_i v_i)}. \tag{229}$$

Thus the rank-one adjustment in Lemma 5.7 also applies to the smoothed maximum; it does not require hard tight constraints. More explicitly, at a conservative trigger of absolute score error δ_e, the exact score is at least $-2\delta_e$. For $\delta_e \leq \delta_{\mathrm{end}}/4$, we have $v_i \geq \delta_{\mathrm{end}}/(2\iota)$, since $\beta_i \leq 1$. A uniform $2\delta_e \nu_i E$ repairs the comparison, at cost exactly $2\delta_e \nu_i$. Before repair the last fraction in Eq. (229) is at least $\rho {\delta_{\mathrm{end}}}^3/(4\iota^2 + 2\iota)$. Taking the tolerance smaller than a fixed fraction of this and of the original reservoir allowance leaves decrease $a\rho {\delta_{\mathrm{end}}}^3$. The other endpoint exchanges U, V. All other accepted moves decrease the same nonnegative, initially bounded potential, so the regular count is $\widetilde{O}(n)$. Explicit initialization thus costs $\widetilde{O}(n^{3/2})$ stages in total.

For a deferred coordinate, absorb its unchanged fractional coefficient into F, preserving M. Write its old slot as $\mathcal{K}_i(D) = \iota\beta_i V_i D V_i$. Hold the other slots fixed and track

$$\mathcal{K}_s = \mathcal{K}_{\mathrm{rest}} + r\mathcal{K}_i, \qquad r = e^{-s}, \qquad 0 \leq s \leq L_*.$$

The reservoir is constant during this evaluation path and is removed at its end. All bounded-model root estimates hold, because the coupling only decreases. Define the positive map on pairs

$$K_s(D_U, D_V) = (r\mathcal{K}_i(D_V), r\mathcal{K}_i(D_U)), \qquad v(s) = \langle \mathsf{Q}, K_s \mathsf{P} \rangle \geq 0.$$

The envelope gives $\mathcal{P}_{\mathrm{s}}'(s) = -v(s)$, and $v(s) \leq C$. The forcing is $F_1 = -K_s\mathsf{P}$, $F_2 = -K_s\mathsf{Q}$. The pair $K_s\mathsf{P}$ is a PSD subpair of the full coupling. Hence the entropy variance bound and weighted Cauchy–Schwarz give

$$\operatorname{tr}[\mathsf{Q}F_1^2] \leq Cv, \quad |\langle F_2, D\rangle| \leq C\sqrt{v}e(D), \quad q(K_s D) \leq C\sqrt{v}e(D).$$

For the last two inequalities, apply the proof of Eq. (215) to the single nonnegative coefficient $r\iota\beta_i$: the first squared factor is $r\iota\beta_i(v_i w_i + u_i z_i) = v$, and the second is bounded by the full dual energy since this slot is included in $\mathcal{K}_s$. Eq. (212) now yields

$$e(\mathsf{P}') + \mathcal{E}(\mathsf{Q}') \leq P_n\sqrt{v}. \tag{230}$$

The proof of Lemma 6.2, with the displayed $q(K_sD)$ bound as its explicit-data term, gives normalized metric and Jacobian variation rate at most $P_n\sqrt{nv}$. Lightness is not needed on this auxiliary path.

There is a relative scalar bound as well. Since $K_s' = -K_s$,

$$v' = -v + \langle \mathsf{Q}', K_s\mathsf{P}\rangle + \langle K_s\mathsf{Q}, \mathsf{P}'\rangle.$$

The two pairings are at most $P_n\sqrt{v}\mathcal{E}(\mathsf{Q}')$ and $C\sqrt{v}e(\mathsf{P}')$ in absolute value. Consequently

$$|v'| \le P_n v. \tag{231}$$

A zero slot can be omitted directly. At a corrected root compute an upper approximation $v \le \bar{v} \le C(v + 1/n)$, using absolute error at most a/n, and take a dyadic increment within fixed factors of

$$h = \frac{a}{P_n(1 + \sqrt{n\bar{v}})},$$

except for the final truncation. All other calculations still use the finer root cache floor. Eq. (231) keeps v comparable over the increment. Its joint energy arc is at most $a/(P_n\sqrt{n})$ and its normalized variation is small. Thus Lemmas 6.1 and 6.2 justify one corrected stage, with positive matrices and trace exactly one.

Choose $r_* = e^{-L_*}$ below the root-cache floor divided by a fixed polynomial in n, k. With the logarithmic cache precision, r_* is inverse polynomial and $L_* = O(\log(2kn))$. At r_* drop the remaining slot exactly and correct once. The direct derivative with respect to r is uniformly polynomially bounded even at zero: the fixed slot has bounded Frobenius operator norm, $\|\mathsf{P}\|_F \le C\sqrt{n}$, and the joint inverse and metric changes have polynomial norm bounds. Extra fixed guard powers in the choice of r_* put this last jump inside the joint Newton ball and the required error floor. The endpoint is precisely the optimizer of the new face, not a presumed continuation across a large face change.

For $\Delta_i = \int_0^{L_*} v(s)ds$, the step count is at most

$$P_n \int_0^{L_*} (1 + \sqrt{nv(s)})ds + 1 \le P_n(L_* + \sqrt{nL_*\Delta_i}) + 1.$$

This follows by integrating over each interval, where v is comparable to its starting value. The nonnegative Δ_i sum to at most the initial potential, since all other moves decrease it and each removal also drops its reservoir term. Summing over at most k removals gives $\widetilde{O}(k + \sqrt{nk})$ stages.

Store s on a common dyadic grid and evaluate $r = e^{-s}$ from the base, not by accumulating rounded products. The smallest step is inverse polynomial, the horizon is $O(\log(2kn))$, and r_* has $O(\log(2kn))$ bits. A fixed multiple of this guard precision puts coefficient, parameter, and solve errors below the preceding margins. The numerical slot stays nonnegative; the true matrices and symbolic signed contribution are unchanged. Process simultaneous deferred hits successively, charging each coordinate once. This introduces no extra rounding or discrepancy allowance. □

6.6 Gibbs-feature preconditioning of curvature search

The random component still returns only directly verified directions. The new sketch is not obtained by applying the Hessian to random vectors. Instead it samples a positive Gram matrix which is comparable to a positive shift of that Hessian. Its features have rank-two matrix representations and admit batched construction and implicit application.

At a failed gradient test, work first at the exact root and put $A = G^{-1/2}\nabla^2\mathcal{F}_{\mathrm{s}}G^{-1/2}$. Later we freeze a sufficiently accurate structured symmetric surrogate. With fixed slack in the tests, choose rational bounds

$$\lambda_{\min}(A) \le -2\delta, \quad -RI \preceq A \preceq M_{\mathrm{aux}}I, \quad 1 \le R \le P_n, \quad \delta^{-1} \le P_n, \quad 0 < \delta \le R/4, \quad M_{\mathrm{aux}} \le P_n k, \tag{232}$$

where $M_{\mathrm{aux}} \ge R$. All bounds may be replaced by rational upper bounds obtained from the fixed budget coefficient sums, the cutoff, and the stated spectral intervals; no norm-estimation oracle is used to choose them. The current face dimension is denoted by k inside a filter; it never exceeds the initial k used in total counts.

A comparable Gram matrix. Retain the exact symmetric Schur formula

$$\nabla^2\mathcal{F}_{\mathrm{s}} = -G + F_1^* Q_{\mathsf{Q}} F_1 - (\mathscr{J} Q_{\mathsf{Q}} F_1 - F_2)^* J_s^{-1} (\mathscr{J} Q_{\mathsf{Q}} F_1 - F_2). \tag{233}$$

One action costs $F(n,k) \le \widetilde{O}(n^{\chi})$, including the root solves, matrix functions, and adjoints. This identity follows by substituting the differentiated root equations into the second-order envelope, so involves the same root and the same potential throughout.

Lemma 6.4 (Gibbs Gram comparison)**.** *Set $D = G^{-1/2}F_1^* Q_{\mathsf{Q}} F_1 G^{-1/2}$. There are known rational bounds $\mathsf{T}_0 \le P_n k$, $\ell_0, b_0 \le P_n$ such that*

$$D \succeq 0, \quad \mathrm{tr}[D] \le \mathsf{T}_0, \quad a_{\mathrm{dom}} D - (1+2b_0)I \preceq A \preceq D - I, \qquad a_{\mathrm{dom}} = (1+2\ell_0)^{-1}.$$

In particular enlarge R so $R \ge 1 + 2b_0$, and put $B = A + 2RI$. Then

$$\tfrac{a_{\mathrm{dom}}}{2}(D + 2RI) \preceq B \preceq D + 2RI. \tag{234}$$

All these bounds survive a sufficiently small inverse-polynomial root/coefficient perturbation, with fixed slack in $a_{\mathrm{dom}}, R, \mathsf{T}_0$.

Proof. The following square roots are for analysis, not computation:

$$\mathsf{A} = Q_{\mathsf{Q}}^{1/2} F_1 G^{-1/2}, \qquad \mathsf{B} = Q_{\mathsf{Q}}^{1/2} \mathscr{J} \mathscr{E}^{-1/2}, \qquad T = \mathscr{E}^{-1/2} F_2 G^{-1/2}.$$

Eq. (203) gives $\|\mathsf{B}\|^2 \le K/\tau_{\mathrm{s}} \le \ell_0$. The weighted coupling estimate gives $\|T\|^2 \le b_0$. Coordinate forcing and the Gibbs variance bound give $D_{ii} \le P_n$, hence the trace bound, without computing the trace. The Schur formula becomes

$$A = -I + \mathsf{A}^*\mathsf{A} - (\mathsf{B}^*\mathsf{A} - T)^*(I + \mathsf{B}^*\mathsf{B})^{-1}(\mathsf{B}^*\mathsf{A} - T).$$

For any y, its quadratic form plus $\|y\|^2$ is

$$\min_{y_{\mathrm{aux}}} \{ \|\mathsf{A}y - \mathsf{B}y_{\mathrm{aux}}\|^2 + \|y_{\mathrm{aux}}\|^2 + 2\langle Ty, y_{\mathrm{aux}}\rangle \}.$$

Use $\|y_{\mathrm{aux}}\|^2 + 2\langle Ty, y_{\mathrm{aux}}\rangle \ge \frac{1}{2}\|y_{\mathrm{aux}}\|^2 - 2b_0\|y\|^2$. The remaining minimum equals

$$\langle \mathsf{A}y, (I + 2\mathsf{B}\mathsf{B}^*)^{-1}\mathsf{A}y\rangle \ge a_{\mathrm{dom}}\|\mathsf{A}y\|^2.$$

The upper bound follows by dropping the nonnegative Schur term. This proves all order comparisons. At a root all relevant smallest eigenvalues are inverse-polynomial; polynomial perturbation bounds therefore preserve the displayed estimates with their slack. In the numerical construction keep the exact algebraic Schur form, as specified below, rather than adding an unrelated Hessian approximation. □

Features whose covariance is D. For the fixed-depth precursor, first fix an integer $h_{\mathrm{sk}} \geq 2$ independently of the input dimension. Section 6.8 separately proves the uniform bounds needed to use a growing depth; it does not substitute a variable depth into a fixed-depth estimate. The features are real coordinate vectors even when the input is complex. For one sample choose t uniformly in $[0,1]$. Independently for each block Q_b of Q choose standard circular complex Gaussian vectors g_b, h_b (each real and imaginary part has variance $1/2$), and put

$$r_b = \mathsf{Q}_b^{t/2} g_b, \quad v_b = \mathsf{Q}_b^{(1-t)/2} h_b, \quad T_b = (r_b v_b^* + v_b r_b^*)/\sqrt{2}, \quad \widetilde{T} = T - \mathrm{tr}[T]\mathsf{Q}, \quad z = {\tau_{\mathrm{s}}}^{-1/2} G^{-1/2} F_1^* \widetilde{T}. \tag{235}$$

These definitions refer to a single fixed root; samples are independent.

Lemma 6.5 (Feature identity and a ridge covariance sample)**.** *The feature has mean zero and $\mathbb{E} z z^\top = D$. Let $1 \leq q \leq 2n^2$, $\mu = 2R + \mathsf{T}_0/q$, and let ℓ_{Q} be an integer upper bound for $1 + \log(\lambda_{\max}(\mathsf{Q})/\lambda_{\min}(\mathsf{Q}))$. Thus $\ell_{\mathsf{Q}} = O(\log(2n))$ with constants depending only on the fixed regularization. Put*

$$p = \lceil 128 \log_2(2kn\ell_{\mathsf{Q}}(h_{\mathrm{sk}}+1)) \rceil, \qquad s = \lceil 2^{20} q \ell_{\mathsf{Q}} p^3 \rceil, \qquad V = [z_1 \ \cdots \ z_s], \qquad W = \mu I + s^{-1} V V^\top.$$

With probability at least 0.98 for ideal samples,

$$\tfrac{3}{4}(D + \mu I) \preceq W \preceq \tfrac{5}{4}(D + \mu I). \tag{236}$$

There is a bounded finite-random-bit construction for which the comparisons with factors $1/2$ and 2 hold with probability at least $9/10$. It uses only polylogarithmic word lengths. Its matrix V is defined by its stored rank-two factors and the displayed centering; it is not perturbed by an arbitrary final full-rank rounding.

Proof. For Hermitian test pairs B, C, the conditional covariance of the uncentered T is the self-adjoint operator

$$\mathcal{T}_t(C) = \tfrac{1}{2}(\mathsf{Q}^t C \mathsf{Q}^{1-t} + \mathsf{Q}^{1-t} C \mathsf{Q}^t).$$

For one block this follows from $\mathrm{tr}[BT] = \sqrt{2}\,\mathrm{Re}(v^* B r)$ and independence of the two circular Gaussians; cross-block covariances are zero. Integrating in t gives $\mathcal{T}(C) = \int_0^1 \mathsf{Q}^t C \mathsf{Q}^{1-t} dt$. For every t, $\mathcal{T}_t(E) = \mathsf{Q}$ and $\mathrm{tr}[\mathsf{Q}] = 1$. Centering thus has covariance $\mathcal{T} - \mathsf{Q} \otimes \mathsf{Q} = \tau_{\mathrm{s}} Q_{\mathsf{Q}}$, proving the feature identity.

In a basis diagonalizing Q, the coefficient of $\mathcal{T}_t$ is no larger than $(z_i + z_j)/2$. The logarithmic-mean coefficient of $\mathcal{T}$ obeys

$$\frac{(z_i + z_j)/2}{(z_i - z_j)/(\log z_i - \log z_j)} = \frac{|\log(z_i/z_j)|}{2} \coth\!\left(\frac{|\log(z_i/z_j)|}{2}\right) \leq 1 + |\log(z_i/z_j)| \leq \ell_{\mathsf{Q}}.$$

Use the continuous value at equality; $y \coth y \leq 1 + y$ follows from $e^{2y} - 1 \geq 2y$. Therefore $\mathcal{T}_t \preceq \ell_{\mathsf{Q}} \mathcal{T}$ as quadratic forms. Apply the same centering congruence and coordinate map on both sides to obtain

$$\mathbb{E}[z z^\top \mid t] \preceq \ell_{\mathsf{Q}} D. \tag{237}$$

Here is an elementary concentration proof, including the unbounded Gaussian tail. Define $w = (D + \mu I)^{-1/2} z$ only for analysis. Its mean covariance is $J = (D + \mu I)^{-1/2} D (D + \mu I)^{-1/2} \preceq I$, with $\mathrm{tr}[J] \leq \mathsf{T}_0/\mu \leq q$. Conditional on t, every real coordinate of w is a bilinear form in two independent real Gaussian vectors (realify the circular pairs). For such a scalar form U, conditioning and Minkowski's inequality give

$$\mathbb{E}|U|^{2p} \leq ((2p-1)!!)^2 (\mathbb{E} U^2)^p.$$

Indeed condition on the second vector, apply the scalar Gaussian moment, and diagonalize the nonnegative covariance of its conditional variance; Minkowski bounds the p-norm of that weighted sum of squared Gaussians. Apply Minkowski once more to the sum of coordinate squares, and use Eq. (237). This proves

$$\mathbb{E}\|w\|^{2p} \le (4p^2\ell_{\mathsf{Q}}q)^p. \tag{238}$$

Let $L_* = 16p^2\ell_{\mathsf{Q}}q$ and truncate, only in the proof, $U = ww^\top \mathbf{1}_{\{\|w\|^2 \le L_*\}}$. The failure probability is at most 4^{-p} and

$$\|J - \mathbb{E}U\| \le L_* 4^{-p}, \quad \|U - \mathbb{E}U\| \le L_*, \quad \mathbb{E}(U - \mathbb{E}U)^2 \preceq L_* I.$$

For clarity, the scalar-variance matrix Bernstein bound needed here has a short proof. If a centered Hermitian B satisfies $\|\mathsf{B}\| \le L_*$ and $\mathbb{E}\mathsf{B}^2 \preceq L_* I$, its exponential series and $j! \ge 23^{j-2}$ give

$$\mathbb{E}e^{\theta\mathsf{B}} \preceq \exp\Big(\frac{\theta^2 L_*}{2(1-\theta L_*/3)}\Big) I \quad (0 \le \theta L_* < 3).$$

The Golden–Thompson inequality and independence apply this scalar matrix bound successively inside an expected trace exponential. Markov's inequality for the largest eigenvalue and then for its negative, with the optimizing scalar θ, yield

$$\Pr\{\|s^{-1}\sum_j (U_j - \mathbb{E}U_j)\| > 1/8\} \le 2k\exp[-s/(256L_*)].$$

This uses a scalar variance bound, not a product of matrix-valued moment bounds; it is the standard trace-exponential argument, cf. [Tro12].

Here $s/(256L_*) \ge 256p$. Moreover $s4^{-p}$ and L_*4^{-p} are below $1/100$ and $1/100$, respectively, with ample slack. Use $q \le 2n^2$ and the defining lower bound on p. Also $p^5 \le 2^{p/2}$ for every $p \ge 64$. The last inequality follows at 64 and then by monotonicity of $p/2 - 5\log_2 p$. A union bound ensures that no sample was truncated, except with the stated small probability. The total deviation is less than $1/4$, proving Eq. (236). No truncation test is required by the algorithm.

Section 7.7 gives the finite sampler, its coupling to these ideal samples, and the certified feature perturbation bound. It preserves the two-sided covariance bound with the stated slack.

□

Entry bounds and a uniform ridge sampler. The trace bound in Lemma 6.4 may be enlarged to $\mathsf{T}_0 = kd_0$, where $1 \le d_0 \le P_n$ is a known rational upper bound on every diagonal entry of D. This choice is used below. The sample matrix is never formed in its entirety.

Lemma 6.6 (Bounded entries without materialization)**.** *For the finite features in Lemma 6.5, with $1 \le q \le 2n^2$ and the enlarged p specified there, the covariance comparison with factors $1/2, 2$ and the entry bound*

$$|V_{ij}|^2 \le \mathscr{J}_e := 64p^2\ell_{\mathsf{Q}}d_0 \quad \textit{for all } i,j \tag{239}$$

hold simultaneously with probability at least $4/5$. This is an analytical event, not a test requiring all entries of V.

Proof. The scalar bilinear Gaussian moment calculation in the feature lemma, now applied to coordinate i before whitening, gives $\mathbb{E}|z_i|^{2p} \le (4p^2\ell_{\mathsf{Q}}D_{ii})^p$. Thus the probability that an ideal feature entry exceeds $16p^2\ell_{\mathsf{Q}}d_0$ in squared magnitude is at most 4^{-p}. Union over ks entries. The proof of the covariance lemma and its truncation bounds is unchanged for $q \le 2n^2$: its polynomial

factors are dominated by 4^{-p} with the enlarged p. For example $p^5 \le 2^{p/2}$ for $p \ge 64$, and the definition of p bounds $k, n, \ell_{\mathsf{Q}}, h_{\mathrm{sk}}+1$ by fixed small powers of 2^p. The same coupling of dyadic Box–Muller samples to ideal uniforms can make every entry error at most $\sqrt{\mathscr{J}_e}/4$, as well as keeping the covariance perturbation below $\mu/16$. All required errors are inverse polynomial. Unioning these failure events leaves probability greater than $4/5$, with the stated slack. Every finite sample, including the excluded event, still has the polynomial magnitude bound in that lemma. □

Lemma 6.7 (Uniform ridge sampling)**.** *Let $F \in \mathbb{R}^{d\times e}$ satisfy $|F_{ij}|^2 \le \mathscr{J}_e$, let $\iota > 0$, $\gamma \ge 1$, and put $\mathscr{E}_\gamma = \gamma I_d + \iota F F^\top$. Sample r column indices independently and uniformly with replacement, and let J be their list. For*

$$\mathcal{L} = \max\{1, \iota e d\, \mathscr{J}_e/\gamma\}, \qquad r \ge 2^{16}\mathcal{L}\log(200d/\zeta),$$

one has, with probability at least $1-\zeta$,

$$\tfrac{3}{4}\mathscr{E}_\gamma \preceq \widehat{\mathscr{E}}_\gamma := \gamma I_d + \iota(e/r)F_{:,J}F_{:,J}^\top \preceq \tfrac{5}{4}\mathscr{E}_\gamma. \tag{240}$$

The statement holds conditionally on any previous choices that fix F. Sampling with replacement, including repeated or zero columns, is allowed.

Proof. Write f_i for a column and define only for analysis $U = e\iota\mathscr{E}_\gamma^{-1/2} f_i f_i^\top \mathscr{E}_\gamma^{-1/2}$ for a uniform i. Then $0 \preceq U \preceq \mathcal{L}I$ and $\Gamma = \mathbb{E}U = \mathscr{E}_\gamma^{-1/2}(\mathscr{E}_\gamma - \gamma I)\mathscr{E}_\gamma^{-1/2} \preceq I$. For $\mathsf{B} = U - \Gamma$, $\|\mathsf{B}\| \le \mathcal{L}$ and $\mathbb{E}\mathsf{B}^2 \preceq \mathcal{L}I$. The exponential-series and Golden–Thompson argument in Lemma 6.5 gives

$$\Pr\{\|r^{-1}\textstyle\sum_j \mathsf{B}_j\| > 1/4\} \le 2d\exp[-r/(256\mathcal{L})] \le \zeta.$$

Congruence by $\mathscr{E}_\gamma^{1/2}$ proves the result. The sampling decisions need no leverage scores, eigendecomposition, or access to all entries. □

Fixed-depth alternating sketches. Fix $h = h_{\mathrm{sk}} \ge 2$ and rational exponents

$$2 \ge a_1 \ge a_2 \ge \cdots \ge a_h \ge 1, \qquad a_{h-1} \le \tfrac{3}{2}a_h, \qquad q_j = \lceil n^{a_j} \rceil. \tag{241}$$

The depth and all these exponents are fixed independently of n, m. Use $q = q_1$ in the Gibbs feature lemma and $\mu = 2R + kd_0/q_1$. Pad the coordinate rows of V by zeros to $\bar{k} = 2^{\lceil \log_2 k \rceil}$. For $k = 1$ this means $\bar{k} = 1$. Define the positive system in feature space

$$K_0 = I_s + (s\mu)^{-1}V^\top V. \tag{242}$$

This is an operator, not a materialized Gram matrix. The exact identity

$$W^{-1} = \mu^{-1}I - (s\mu^2)^{-1}V K_0^{-1} V^\top \tag{243}$$

will be evaluated by solving K_0 accurately; a constant-factor inverse approximation is not substituted into this subtractive formula.

Choose each d_j to be the least power of two at least $2^{32}q_j\ell_{\mathsf{Q}}p^3$. First sample d_1 coordinate indices from the padded $\bar{k}$ rows, obtaining a list I_1, and set

$$\Phi_1 = V_{I_1,:} \in \mathbb{R}^{d_1\times s}, \quad e_1 = s, \quad \iota_1 = \frac{\bar{k}}{d_1 s\mu}, \quad K_1 = I_{d_1} + \iota_1\Phi_1\Phi_1^\top. \tag{244}$$

Only the indices and the feature factors are stored. Lemma 6.7 applied to $V^\top$ shows that

$$W_0 = I_s + \iota_1 \Phi_1^\top \Phi_1, \qquad \tfrac{3}{4} K_0 \preceq W_0 \preceq \tfrac{5}{4} K_0, \qquad W_0^{-1} = I_s - \iota_1 \Phi_1^\top K_1^{-1} \Phi_1 \tag{245}$$

on its sampling event.

For $j = 1, \dots, h-1$, suppose $\Phi_j \in \mathbb{R}^{d_j \times e_j}$ and $K_j = I + \iota_j \Phi_j \Phi_j^\top$ have been defined. Put $\lambda_j = q_j / q_{j+1} \geq 1$. Select independently a list J_{j+1} of d_{j+1} columns from the e_j columns, and define

$$\begin{aligned} \Phi_{j+1} &= (\Phi_{j,:,J_{j+1}})^\top, \\ e_{j+1} &= d_j, \\ \iota_{j+1} &= \frac{\iota_j e_j}{d_{j+1} \lambda_j}, \\ K_{j+1} &= I_{d_{j+1}} + \iota_{j+1} \Phi_{j+1} \Phi_{j+1}^\top. \end{aligned} \tag{246}$$

Here $\Phi_{j,:,J}$ means the indicated column submatrix of Φ_j. The associated preconditioner and its exact inverse are

$$\begin{aligned} W_j &= \lambda_j (I_{d_j} + \iota_{j+1} \Phi_{j+1}^\top \Phi_{j+1}), \\ W_j^{-1} &= \lambda_j^{-1} (I_{d_j} - \iota_{j+1} \Phi_{j+1}^\top K_{j+1}^{-1} \Phi_{j+1}). \end{aligned} \tag{247}$$

Thus each inverse is reduced to the next, smaller system. At the last level form K_h and compute a residual-certified short inverse. No other feature Gram matrix is formed.

An index population of non-power-of-two size is padded by zero columns to the next power of two before sampling, and its padded size is used in the scale e_j / d_{j+1}. Such padding multiplies the coefficient invariants below by at most two per level. Alternatively use a fixed cap of dyadic rejection proposals for exact uniform indices from the unpadded population; reject the whole trial on exhaustion. To avoid these factors, use the latter convention for the first feature-column population s; every later population is a power of two. A cap of $\lceil \log_2(10^6 (h+1) \sum_j d_j) \rceil$ proposals per draw fails with probability at most 10^{-6} by a union bound. All conditional accepted indices are exactly uniform, and all samplers terminate after their caps.

Lemma 6.8 (Recursive comparisons and invariants)**.** *Conditionally on the entry event, with probability at least* 99/100 *all the uniform sampling comparisons above and below hold. For every j the exact coefficient invariant is*

$$\iota_j d_j e_j = \frac{\bar{k}}{\mu} \frac{q_j}{q_1} \leq \frac{2 q_j}{d_0}. \tag{248}$$

Moreover

$$\tfrac{3}{4} K_j \preceq W_j \preceq \tfrac{5}{4} \lambda_j K_j \quad (1 \leq j < h). \tag{249}$$

All finite-sample K_j and W_j are positive definite, even on failure of the comparisons, and have polynomially bounded norms and inverse norms for fixed h.

Proof. The invariant follows at $j = 1$ from Eq. (244) and is preserved by Eq. (246). Its inequality uses $\mu \geq k d_0 / q_1$ and $\bar{k} \leq 2k$. Every Φ_j is a submatrix of V or $V^\top$, with possible repetitions. Hence its entries still obey Lemma 6.6 and Eq. (239). The sampling lemma at level j has

$$\mathcal{L} \leq \max\{1, 2 q_{j+1} \mathscr{J}_e / d_0\} \leq 129 p^2 \ell_{\mathsf{Q}} q_{j+1}.$$

The initial sample has the same bound with q_1. The chosen d_j therefore makes each conditional failure probability at most $1/(1000(h+1))$. To check the scalar constants, all dimensions are at

most $2^{40}(n+1)^2(\ell_{\mathsf{Q}}+1)(p+1)^4(h+1)$, whose logarithm is at most $4p$; the exponential bound in the preceding proof has ample margin. Union the finitely many conditional failures and the finite-sampling cap failure. No independence between different levels is asserted or needed.

The sampling comparison is with $\lambda_j I+\iota_j\Phi_j\Phi_j^\top$. This matrix lies between K_j and $\lambda_j K_j$, proving Eq. (249). The identity for W_j^{-1} is Woodbury, with no rank or independence requirement on the selected columns. It holds for signed feature entries and repeated columns. Every $K_j \succeq I$ and every $W_j \succeq \lambda_j I$. The finite feature bounds, prescribed sample counts, and explicit positive rational scales bound their norms by a fixed polynomial, also off the good event. This proves the last assertion without a random conditioning assumption. □

Applying submatrices, not materializing them. Define the piecewise affine multiplication upper bound

$$\vartheta(a)=\begin{cases}\omega_0+2(a-1)(\xi_0-\omega_0), & 1\le a\le 3/2,\\ \xi_0+2(a-3/2)(\chi-\xi_0), & 3/2\le a\le 2,\end{cases}\qquad \xi_0=2.794634. \tag{250}$$

Both pieces are justified by tensor interpolation and permutation from the strict square bound and the bounds for shapes $(1,3/2,1)$ and $(1,2,1)$ in Table 1 of [ADW+25]. They are upper bounds, not assertions about the exact multiplication exponent.

Lemma 6.9 (Implicit restricted feature actions)**.** *A submatrix of V, or its transpose, having at most $\widetilde{O}(n^a)$ selected coordinate rows and selected feature columns can be applied, together with its adjoint, in $\widetilde{O}(n^{\vartheta(a)})$ arithmetic and bit operations. The lists may contain repetitions. At one fixed root all initial sample vectors cost $\widetilde{O}(n^3+n^{\vartheta(a_1)})$ to prepare. For the final factor Φ_h, the factor itself, its Gram matrix, and its short inverse cost*

$$\widetilde{O}(n^{\mathfrak{g}(a_{h-1},a_h)}+n^{a_h\omega_0}),\quad \mathfrak{g}(r,v)=2(\xi_0-\omega_0)r+(3\omega_0-2\xi_0)v. \tag{251}$$

Proof. For selected sample columns, a weighted feature sum is first represented as a pair of weighted rank-two outer-product sums, with the single centering correction in Eq. (235). Compute it with products of shapes (n,n^a,n), and evaluate only the selected original input quadratic forms with shapes (n,n,n^a). Response coefficients and diagonal $G^{-1/2}$ are applied before or after these maps as appropriate. For the adjoint, form the forcing pair from the selected original coordinates and test it against only the selected samples. The same products suffice. Repeated indices are treated as repeated factor columns, or their coefficients can be summed; neither method requires materializing the feature entries. Padded coordinate indices denote zero functionals and are skipped; no nonexistent weight is inverted. All these maps are real coordinate maps even for complex Hermitian data. There is no new spectral decomposition per action.

The two order-n root spectral decompositions are performed once. To form the initial s sample vectors, apply the eigenvector matrices to the arrays of diagonally scaled Gaussians by the same rectangular products. Scalar powers, sampling, diagonal scalings and lists cost $\widetilde{O}(n^{1+a_1})$, which is smaller. Thus the preparation cost is as asserted, rather than square blocking at exponent ω_0+a_1-1.

Let $r=a_{h-1}$ and $v=a_h$. The final factor has dimensions $\widetilde{O}(n^v)\times\widetilde{O}(n^r)$; its entries are computed by a fixed number of products of shape (n^v,n,n^r) and entrywise operations. Since $v\ge 1$, these fit the cost for the terminal Gram, whose shape is (n^v,n^r,n^v). For $v\le r\le 3v/2$, tensor interpolation of the square and $3/2$ shapes gives $\mathfrak{g}(r,v)$. The terminal system has smallest eigenvalue at least one. The existing normwise-stable Newton–Schulz construction forms its inverse with $\widetilde{O}(n^{v\omega_0})$ work, including residual certification.

The represented factors are fixed exactly by their stored dyadic rank-two data; a materialized terminal factor and its implicit actions are approximations to the same matrix. Normwise errors are additive. Use strict-exponent fixed recursive products, whose logarithmic depth and fixed coefficients give polynomial normwise amplification. Smaller polylogarithmic unit roundoff supplies all stated action and inverse accuracies. This proves the bit bounds and preserves the real/complex adjoint identities without an irrational Hermitian coordinate basis. □

Executable nested inverse solves. Use W_0 to solve K_0 on a fixed interval $[1/2, 2]$, and W_j to solve K_j on

$$[1/(4\lambda_j), 4] \quad (1 \le j < h). \tag{252}$$

The good-event comparisons leave fixed slack in these intervals. For any of these systems $Kv = b$ and preconditioner W, define $M := W^{-1}K$, $\iota := (a_{\mathrm{spec}} + \beta)/2$, $t := (\beta - a_{\mathrm{spec}})/2$, $D := (\iota I - M)/t$, and $\varrho := \iota/t$. With $b' := W^{-1}b$, use

$$\begin{aligned} a_1 &:= 1/\varrho, & a_{j+1} &:= (2\varrho - a_j)^{-1}, \\ v_0 &:= 0, & v_1 &:= b'/\iota, \\ v_{j+1} &:= 2a_{j+1}Dv_j - a_j a_{j+1} v_{j-1} + 2a_{j+1}b'/t. \end{aligned}$$

Its exact error identity is

$$v_* - v_j = T_j(D)v_*/T_j(\varrho). \tag{253}$$

All exact ratios lie in $[0, 1]$; clip computed ratios to this interval. No exponentially large recurrence numerator is stored. Thus K_0 needs polylogarithmically many iterations; K_j needs $\widetilde{O}(\sqrt{\lambda_j})$ iterations. Apply W_j^{-1} using Eq. (247) and a recursive solve of K_{j+1}. Use the residual corrections of Lemma 6.10 at each level, requesting the same inverse-polynomial relative-to-right-hand-side accuracy from every child. At the leaf use the stored short inverse. All recursion depths and iteration lengths are prescribed before the trial. This is an iterative inverse application, not an exact inverse oracle.

For the outer shifted system $B = A + 2RI$, apply W^{-1} by Eq. (243). The feature comparison gives the slack interval

$$\frac{a_{\mathrm{dom}}R}{4\mu} \le \gamma(W^{-1}B) \le 4,$$

so one accurate B solve uses $\widetilde{O}(\sqrt{1 + k/q_1})$ applications. Use a bounded outer polynomial. Define $T := RB^{-1}$, so $0 \preceq T \preceq I$, and define

$$t_- := R/(2R - \delta), \qquad t_+ := R/(2R - 2\delta), \qquad t_0 := (t_- + t_+)/2, \qquad \Delta := (t_+ - t_-)/2.$$

Here $\Delta \ge \delta/(8R)$. Choose a positive dyadic ϵ_f with $\epsilon_f \le 1/4$ and $\epsilon_f^2 \le \delta/(64k(M_{\mathrm{aux}} + \delta))$. Choose an integer $d_f \ge \max\{2/\Delta, 8\Delta^{-2}\log(2/\epsilon_f)\}$ and define $j_* := \lceil d_f t_0 \rceil$. Since R/δ is polylogarithmic, so is d_f. The filter is

$$p_f(t) := \sum_{j=j_*}^{d_f} \binom{d_f}{j} t^j (1-t)^{d_f - j}, \qquad y_{\mathrm{aux}} := p_f(T)\xi, \tag{254}$$

where $\xi \in \{\pm 1\}^k$ is independent of the sketches. The Bernoulli logarithmic moment bound, whose second derivative is at most $1/4$, gives $0 \le p_f \le 1$ on $[0, 1]$, $p_f \le \epsilon_f$ on $[0, t_-]$, and $p_f \ge 1 - \epsilon_f$ on $[t_+, 1]$. Evaluate the polynomial by de Casteljau's recurrence, starting with $v_j^{(0)} := \mathbf{1}_{\{j \ge j_*\}}\xi$ and setting $v_j^{(r)} := (I - T)v_j^{(r-1)} + Tv_{j+1}^{(r-1)}$. The $O(d_f^2)$ inverse actions change only logarithmic factors in the arithmetic cost. Section 7.9 proves the array contraction and the error bound for finite-precision evaluation.

6.7 Certified states and discrepancy

No hidden oracle at the Newton stages. At every face or accepted state an optimizer cache has been certified inside its joint-energy ball and then corrected to its specified floor. Use that base inverse for the next candidate or homotopy increment, as in Lemma 6.2. The current residual uses logarithms of positive Q, not a Gibbs state presumed known in advance. Traceless updates keep $\operatorname{tr}[\mathsf{Q}] = 1$ exactly; rational projection subtracts $\operatorname{tr}[\mathsf{B}]E/(2n)$. Positive definiteness at intermediate points follows from the proved ball and errors smaller than a fixed fraction of its radius. Residual enclosures translate to root-error enclosures by the local inverse bound. No analytic-center, SDP, or exact spectral oracle is invoked.

Section 7.5 constructs the short caches, matrix-function actions, and residual enclosures that maintain this invariant. The higher-rank endpoint and discrepancy accounting is given in the main proof.

6.8 An endpoint bound with dimension-dependent sketch depth

We use a dimension-dependent sketch depth $O(\log\log(2n))$. The following uniform precision and cost bounds give a complete verified curvature search in $\widetilde{O}(n^{\chi})$ operations.

We may pad the matrices with zero rows and columns to order $N = 2^{\lceil \log_2 \max\{2,n\} \rceil}$, where $N \le 2\max\{2,n\}$. Run the same normalized algorithm at that order, including its regularization $\rho = \varepsilon_{\mathrm{scale}}/N$. Padding preserves the matrix variance, rank one and the output discrepancy. Its input work remains $\widetilde{O}(mn^2)$, with additive L_{in} in bits. In the remainder of this subsection, dimension-dependent bounds refer to N. Put

$$r = \log_2 N, \qquad \mathscr{J} = r + 1.$$

All logarithmic bounds are in $\mathscr{J}$ because $k \le N^2$ in the nonlinear phase. Original long input coefficients are read and converted only in the unchanged preprocessing.

A uniform rectangular action bound. We will use a deliberately weaker slope than the interpolation bound in Eq. (250). Set $\sigma_{\mathrm{mm}} = 5/6$. For every power of two q with $N^{4/3} \le q \le N^2$, restricted feature actions, their adjoints, and the corresponding factor products have uniform cost

$$\widetilde{O}(N^{\chi}(q/N^2)^{5/6}). \tag{255}$$

Here the constants and logarithmic degree are independent of q; this is important when the shapes are chosen from the input dimension. Indeed $\chi - \omega_0 = 0.878859 > 5/6$. Fix once and for all a common power-of-two base large enough to implement the square and $(1,2,1)$ bilinear algorithms at strict exponents below ω_0 and χ. Choose their slack small enough that the difference of those exponents still exceeds $5/6$. Tensoring $t - r$ square levels and r rectangular levels gives shape (b^t, b^t, b^{t+r}), with cost at most a constant times

$$(b^t)^{\chi}(b^{t+r}/b^{2t})^{5/6}.$$

Rounding the two dimensions up to adjacent base powers changes them by fixed factors. Tensor permutation supplies both orientations used in Lemma 6.9. Thus one finite collection of fixed bilinear algorithms implements all the required shapes. Their depth is logarithmic; fixed coefficients give polynomial normwise error amplification uniformly in the mix of levels. Lists enlarged by polynomial factors in $\mathscr{J}$ can be split into that many blocks, contributing only a fixed logarithmic factor. No new or input-dependent matrix multiplication exponent is used.

Avoiding exponentially large recurrence numerators. Consider one of the positive systems $Kv = b$ with the stated preconditioner W, on a slack interval $[a_{\mathrm{spec}}, \beta]$. Retain $M = W^{-1}K$, $\iota = (a_{\mathrm{spec}} + \beta)/2$, $t = (\beta - a_{\mathrm{spec}})/2$, $D = (\iota I - M)/t$, and $\varrho = \iota/t > 1$. Instead of storing the numerators and $T_j(\varrho)$ in Eq. (253), store

$$a_j = T_{j-1}(\varrho)/T_j(\varrho).$$

Use the normalized recurrence

$$\begin{aligned} a_1 &= 1/\varrho, \qquad a_{j+1} = (2\varrho - a_j)^{-1}, \\ v_0 &= 0, \qquad v_1 = b'/\iota, \qquad b' = W^{-1}b, \\ v_{j+1} &= 2a_{j+1}Dv_j - a_j a_{j+1} v_{j-1} + 2a_{j+1}b'/t. \end{aligned} \tag{256}$$

It has exactly the same error polynomial:

$$v_* - v_j = T_j(D)v_*/T_j(\varrho). \tag{257}$$

In exact arithmetic $0 < a_j \le 1$. The update of a_j is Lipschitz with constant at most one on $[0, 1]$, since $2\varrho - a_j \ge 1$. The computed ratio can be clipped to $[0, 1]$; clipping cannot increase its distance from the exact ratio. Divisions and the clipping test are short rational operations.

Lemma 6.10 (Uniform precision for nested normalized solves)**.** *Let $2 \le h \le \mathscr{J}^2$ and let $N^{4/3} \le q_h \le \cdots \le q_1 \le N^2$ be powers of two. Use the feature and index sample counts in Lemmas 6.5 and 6.8, with this h. The entry/covariance and all conditional sampling events have the same constant joint success lower bound as before. On their intersection, all recursive solves, shifted solves, and the outer curvature filter can be evaluated with at most $D_0 \mathscr{J}$ accuracy and magnitude bits, for a fixed integer D_0 independent of h, N and the q_j. There is a fixed integer $D \ge 1$, independent of these parameters, such that the full capped cost of one trial is at most*

$$\begin{aligned} &\mathscr{J}^{4D}(N^3 + \mathcal{M}_{\mathrm{aux}}(N, q_1) + q_h^{\omega_0}) + \mathscr{J}^{4D}\sqrt{1 + k/q_1}\, N^{\chi} \\ &+ \sum_{j=1}^{h-1} \mathscr{J}^{D(j+4)} \sqrt{(k + q_1)/q_{j+1}}\, \mathcal{M}_{\mathrm{aux}}(N, q_{j-1}) \\ &+ \mathscr{J}^{D(h+4)} \sqrt{(k + q_1)/q_h}\, q_h^2, \end{aligned} \tag{258}$$

where $q_0 = q_1$ and $\mathcal{M}_{\mathrm{aux}}(N, q) = N^{\chi}(q/N^2)^{5/6}$. The setup term presumes $q_{h-1} = q_h$; a fixed factor in their ratio would give the same conclusion. Each returned direction is verified against the true model. All samples, including failed sampling events, have the displayed deterministic cost cap.

Proof. Section 7.8 proves the uniform precision assertion, the conditional success probability, and the deterministic caps. Here we count the recursive calls.

With this uniform precision guarantee, choose a fixed D large enough to absorb all fixed constants, sample oversampling, matrix-product and spectral overhead, and the accuracy factor in a Chebyshev length into $\mathscr{J}^D$ at a single level. This D does not depend on the depth. One recursive level therefore has cost at most $\mathscr{J}^D\sqrt{q_j/q_{j+1}}$ times the sum of its restricted action cost and one child solve. K_0 costs only another $\mathscr{J}^D$ factor, and the shifted solve costs $\mathscr{J}^D\sqrt{1 + k/q_1}$ such applications. Expanding this recurrence gives Eq. (258), with conservative extra factors $\mathscr{J}^{4D}$. At the leaf, $q_{h-1} = q_h$ allows the factor, its Gram, and its inverse to be formed by square products of order $\widetilde{O}(q_h)$, costing $\mathscr{J}^D q_h^{\omega_0}$; its inverse action costs $\mathscr{J}^D q_h^2$. Preparing all sampled vectors costs $\mathscr{J}^D(N^3 + \mathcal{M}_{\mathrm{aux}}(N, q_1))$. Index lists and every random bit are smaller. This proves the cost in both operation models without suppressing a depth-dependent logarithmic power. □

A sketch schedule absorbing the nesting overhead. Fix the integer D from Lemma 6.10, then fix

$$C_* = 1000(D+1), \qquad b_{\max} = \lfloor 2r/3 \rfloor, \qquad b_0 = \min\{b_{\max}, \lceil C_* \log_2 \mathscr{J} \rceil\}.$$

The first ceiling is computed as the least integer b for which $2^b \geq \mathscr{J}^{C_*}$; its exponent is a fixed integer. If $b_0 = b_{\max}$, use the two deficits (b_0, b_0). Otherwise form

$$b_{t+1} = \min\{b_{\max}, \lceil 3b_t/2 \rceil\} \tag{259}$$

until $b_T = b_{\max}$ and use the list

$$(\ell_1, \ldots, \ell_h) = (b_0, b_1, b_1, b_2, b_2, \ldots, b_T, b_T), \qquad q_j = 2^{2r-\ell_j}, \quad \ell_0 = \ell_1. \tag{260}$$

This is a completely integer schedule. The sample sizes are nonincreasing, and

$$N^{4/3} \leq q_h = q_{h-1} < 2N^{4/3}, \qquad 2^{b_0} \leq 2\mathscr{J}^{C_*}, \qquad h = O(\log \mathscr{J}).$$

More explicitly, $h \leq 6 + 4\log_2 \mathscr{J}$, apart from the two-level capped case; our large fixed C_* makes $h \leq \mathscr{J}^2$ in every case. Neither the matrix algorithms nor their coefficients depend on these variable exponents; Eq. (255) was uniform in the sizes.

Lemma 6.11 (Summing all recursive work)**.** *For the schedule in Eq. (260), the full capped cost in Eq. (258) is $\widetilde{O}(N^\chi)$, with a fixed logarithmic degree independent of N. In particular the factor $\mathscr{J}^{Dh}$ is not silently treated as a polylogarithm.*

Proof. The full top-level action term is bounded by

$$\mathscr{J}^{4D}\sqrt{1 + k/q_1}N^\chi \leq 2\mathscr{J}^{4D+C_*/2}N^\chi.$$

Setup costs at most a fixed logarithmic factor times $N^3 + N^\chi + N^{4\omega_0/3}$, and $4\omega_0/3 < \chi$.

Suppose first $b_0 < b_{\max}$. In the jth summand in Eq. (258), $k + q_1 \leq 2N^2$ gives, up to a fixed factor,

$$N^\chi \mathscr{J}^{D(j+4)} 2^{-5\ell_{j-1}/6 + \ell_{j+1}/2} \leq 2N^\chi \mathscr{J}^{D(j+4)} 2^{-\ell_{j-1}/12}. \tag{261}$$

The schedule satisfies $\ell_{j+1} \leq 3\ell_{j-1}/2 + 1$. Writing $t = \lfloor (j-1)/2 \rfloor$, the deficit $\ell_{j-1} = b_t$ has not yet been capped, so

$$\ell_{j-1} \geq (3/2)^t b_0 \geq C_*(3/2)^t \log_2 \mathscr{J}.$$

For every integer $j \geq 1$, $j + 4 \leq 8(3/2)^t$ and $j \leq 4(3/2)^t$; induction on the pairs $(j = 2t+1, 2t+2)$ proves both inequalities. Our choice of C_* thus gives

$$D(j+4) - C_*(3/2)^t/12 \leq -2j.$$

Summing Eq. (261) costs at most $2N^\chi \sum_{j \geq 1} \mathscr{J}^{-2j} = O(N^\chi)$.

The remaining leaf inverse-action term is at most a fixed factor times

$$\mathscr{J}^{D(h+4)} N^3,$$

since $q_h < 2N^{4/3}$. This term requires its own bound. Before the final truncation, each increase in b_t is at least $b_0/2$. It follows that $h + 4 \leq 12b_{\max}/b_0$. Therefore

$$D(h+4)\log_2 \mathscr{J} \leq (12D/C_*)b_{\max} \leq (8D/C_*)r < r/100.$$

The term is at most a constant times $N^{3.01}$, smaller than N^χ. This establishes the bound without invoking an unquantified subpolynomial-overhead argument.

If $b_0 = b_{\max}$, the depth is exactly two. The logarithmic factors are already fixed powers. The single restricted-action exponent has factor $2^{-b_0/3} \leq 1$, and the leaf action is $O(N^3)$ up to its fixed logarithmic factor. The top term obeys the same bound because $2^{b_0} \leq 2\mathscr{J}^{C_*}$. All cases are covered. □

6.9 A negative spectral subspace and larger curvature moves

Lemmas 6.10 and 6.11 make a complete verified curvature search cost $\widetilde{O}(n^{\chi})$. We retain the dimension of the negative spectral subspace to bound the number of moves. A bounded spectral filter, rather than a filter concentrating only on the most negative eigenvector, produces a direction with a useful coordinate spread. A logarithmic number of reservoir reallocations and simultaneous attenuation of deferred coordinates complete the implementation.

We retain the input conversion, variance scaling, batched grouped partial signing, smoothed potential, root equations, and numerical root machinery of the preceding subsections. In particular $\sum_i V_i^2 \preceq I$, $\iota = 14779/2500 > 5$, $\rho = 10^{-8}/n$, $\tau_{\mathrm{s}} = 10^{-8}/\lceil \log_2(4n) \rceil$, and the cutoff δ_{end} is unchanged. The initial retained count is $k_0 \le n^2$; below k means the current count. As in the endpoint implementation, zero padding to a power-of-two matrix order costs only constant factors. All bounds may be read at that padded order. This does not pad the active coordinate space in the spectral arguments below.

The symbol P_n continues to denote a fixed positive polynomial in $\log(2n)$, enlarged finitely many times. All additional constants below can be replaced by rational upper or lower bounds obtained from the fixed budget, its derivatives on the truncated interval, and the root bounds already proved. None requires estimating an unknown spectrum. The reservoir coefficient will vary between phases; it always satisfies $\gamma \ge 10^{-6}/(L_0 k_0)$, where $L_0 = 1 + \lceil \log_2 \max\{2, k_0\} \rceil$. Root geometry is independent of γ. All weight and numerical bounds used previously remain valid after this logarithmic change of the lower bound.

A rank-defect version of the covariance construction. We first record the linear algebra that supplies multiplicity, not merely one negative eigenvalue. The normalized response matrices in Lemma 4.3 act on a real coefficient space of dimension $2k$. In this paragraph call them $T, M_{\mathrm{aux}}, \Gamma$, to avoid confusion with the current scaled Hessian. In the paired orthonormal basis e_i, f_i they satisfy

$$M_{\mathrm{aux}} \succeq T^* \Gamma T, \quad M_{\mathrm{aux}} \succeq 0, \quad \langle e_i, M_{\mathrm{aux}} e_i \rangle = \langle f_i, M_{\mathrm{aux}} f_i \rangle = 1, \quad \Gamma e_i = \Gamma_i e_i, \quad \Gamma f_i = \Gamma_i f_i,$$

with $\Gamma_i \ge \iota > 5$, and

$$\langle e_i, T e_i \rangle + \langle f_i, T f_i \rangle = 0, \qquad \langle e_i, T e_i \rangle \le 0.$$

Let Π_0 project onto the span of the e_i.

Lemma 6.12 (Projection with a rank defect)**.** *Let Π have rank $k-r$ and satisfy $(I-\Pi_0)(I-T)\Pi = 0$. With*

$$\theta_i = \frac{2}{\sqrt{1+4\Gamma_i}-1}, \qquad \kappa_i = \frac{1+\theta_i}{1-\theta_i^3},$$

there is a PSD matrix A, supported on Π, such that

$$\operatorname{tr}[M_{\mathrm{aux}} \mathsf{A}] \le \sum_i \kappa_i \langle e_i, (I-T)\mathsf{A}(I-T)^* e_i \rangle + r. \tag{262}$$

If $0 \le r \le k/4$, the same matrix satisfies

$$\operatorname{tr}[(I-T)\mathsf{A}(I-T)^*] \ge k/175. \tag{263}$$

These estimates do not depend on the smallest covariance weight.

Proof. Write $u_i = 1 + \theta_i^2 + \theta_i^3$ and put

$$\varpi_i^e = \frac{1}{(1+\theta_i)u_i}, \quad \varpi_i^f = \frac{\theta_i}{(1+\theta_i)(2+\theta_i)}, \quad h_i = \frac{2}{u_i}, \quad j_i = \frac{2}{2+\theta_i}.$$

Define $D_* = \sum_i (\varpi_i^e e_i e_i^* + \varpi_i^f f_i f_i^*)$. For an orthonormal basis matrix E_* of range Π, set

$$C_* = E_*^* D_*^{-1} E_*, \quad \mathsf{A} = E_* C_*^{-1} E_*^*, \quad \mathsf{B} = D_* - \mathsf{A}.$$

Both A, B are PSD and $\mathsf{A}D_*^{-1}\mathsf{A} = \mathsf{A}$, $\mathsf{A}D_*^{-1}\mathsf{B} = 0$, and $\operatorname{tr}[D_*^{-1}\mathsf{A}] = k - r$.

For completeness, the row completions in Lemma 4.2 apply to this rank without alteration. For rows $t_i = e_i^* T$ and $s_i = f_i^* T$, they give

$$\Gamma_i t_i \mathsf{B} t_i^* + \kappa_i (e_i^* - t_i)\mathsf{A}(e_i - t_i^*) + h_i t_i e_i \geq \frac{\langle e_i, \mathsf{A} e_i\rangle}{\varpi_i^e} - \frac{h_i^2}{4\Gamma_i \varpi_i^e},$$
$$\Gamma_i s_i \mathsf{B} s_i^* + j_i s_i f_i \geq \frac{\langle f_i, \mathsf{A} f_i\rangle}{\varpi_i^f} - \frac{j_i^2}{4\Gamma_i \varpi_i^f}.$$

The first is completion with $\Gamma_i \mathsf{B} + \kappa_i \mathsf{A}$, whose inverse is

$$\Gamma_i^{-1} D_*^{-1} + (\kappa_i^{-1} - \Gamma_i^{-1}) D_*^{-1} \mathsf{A} D_*^{-1}.$$

The second uses $s_i \mathsf{A} = f_i^* \mathsf{A}$ and a PSD square with B; its invertibility is not needed. The scalar identity

$$\frac{h_i^2}{4\Gamma_i \varpi_i^e} + \frac{j_i^2}{4\Gamma_i \varpi_i^f} = 1 - \varpi_i^e - \varpi_i^f$$

and $h_i > j_i$ show, on summing, that

$$\operatorname{tr}[\Gamma T \mathsf{B} T^*] + \sum_i \kappa_i \langle e_i, (I-T)\mathsf{A}(I-T)^* e_i\rangle \geq \operatorname{tr}[M_{\mathrm{aux}} D_*] - r.$$

Here the sum of the added diagonal terms is nonpositive. Since $M_{\mathrm{aux}} \succeq T^* \Gamma T$, $\operatorname{tr}[M_{\mathrm{aux}}\mathsf{A}] \leq \operatorname{tr}[M_{\mathrm{aux}} D_*] - \operatorname{tr}[\Gamma T \mathsf{B} T^*]$; this proves Eq. (262). The loss is exactly the rank defect r, not an unquantified dimension factor.

For movement, $0 < \theta_i \leq 101/200$ implies $(\varpi_i^e)^{-1} \leq 3$ and $(\varpi_i^f)^{-1} \leq 2\Gamma_i$. Thus $D_*^{-1} \preceq 3\Pi_0 + 2\Gamma\Pi_1$, where $\Pi_1 = I - \Pi_0$. The response identity gives $\Pi_1 E_* = \Pi_1 T E_*$. Since Γ commutes with Π_1,

$$\operatorname{tr}[C_*] \leq 3(k-r) + 2\|\Gamma^{1/2}\Pi_1 T E_*\|_F^2 \leq 3k + 2\operatorname{tr}[M_{\mathrm{aux}}] \leq 7k.$$

Also $\|T\|_F^2 \leq 2k/\iota$. Write $\ell = k - r \geq 3k/4$. Then

$$\operatorname{tr}[(I-T)\Pi] \geq \ell - \sqrt{2k\ell/\iota} \geq k/5.$$

The last inequality follows from $\iota > 5$: the expression increases on $\ell \geq 3k/4$, and $(3/4 - 1/5)^2 = 121/400 > 3/10$. Weighted Frobenius Cauchy–Schwarz now gives

$$(\operatorname{tr}[(I-T)\Pi])^2 \leq \|(I-T)E_* C_*^{-1/2}\|_F^2 \|E_* C_*^{1/2}\|_F^2 = \operatorname{tr}[(I-T)\mathsf{A}(I-T)^*]\operatorname{tr}[C_*].$$

This proves Eq. (263). □

We spell out how the rank defect transfers to physical directions. Retain $u_i, v_i, w_i, z_i, t_i, p_i, q_i$ from the smoothed section and define

$$z_i^{\mathrm{mov}} = p_i^2 u_i w_i + q_i^2 v_i z_i > 0, \qquad M_{\mathrm{aux}\,i} = p_i u_i - q_i v_i.$$

In the response construction put $\iota_i = (\beta_i p_i/q_i)^{1/2}$, $d_i = (\beta_i q_i/p_i)^{1/2}$, and $\tau_i = (p_i q_i/\beta_i)^{1/2}$. The two response-coordinate energy diagonals are $E_{W,i} = \iota_i^2 u_i w_i$ and $E_{Z,i} = d_i^2 v_i z_i$; thus $\tau_i^2(E_{W,i} + E_{Z,i}) = z_i^{\mathrm{mov}}$. The response graph in Eq. (96) maps bijectively to physical $h \in \mathbb{R}^k$. Indeed the two equations $p - Bq = J$ and $Ap - q = J$ have a unique solution for every J, since $I - BA$ is invertible by the strict dual inequalities; $h_i = J_i/\tau_i$. Consequently imposing any r independent linear conditions on h reduces its normalized response subspace dimension from k to $k - r$. All other hypotheses of Lemma 6.12 are unchanged.

Corollary 6.13 (Covariance on a prescribed subspace)**.** *For every subspace $\mathscr{E} \subseteq \mathbb{R}^k$ of codimension $r \le k/4$ at a light state, there is a finitely supported centered $h \in \mathscr{E}$, with covariance Σ, such that*

$$\mathbb{E}Z(h) \le \sum_i \kappa_i z_i^{\mathrm{mov}} \Sigma_{ii} + r, \qquad \sum_i z_i^{\mathrm{mov}} \Sigma_{ii} \ge k/175. \tag{264}$$

Here $Z(h)$ denotes the fixed-constraint inversion energy, not the entropy covariance operator Q_{Q}.

Proof. Use the restricted response projection in Lemma 6.12. The normalized energy matrix has trace $2k$ and the same unit diagonal. The physical covariance is obtained by applying the fixed response map to $\mathsf{A}^{1/2}\xi$, where ξ consists of independent signs. The movement matrix is supported on the e_i coordinates and its ith diagonal is $z_i^{\mathrm{mov}}\Sigma_{ii}$. The two conclusions therefore follow directly. This distribution is only an existence certificate in the proof; the algorithm never forms the response projection, its covariance, or its square root. □

Many negative eigenvalues after a failed gradient test. For any covariance in Corollary 6.13, use the preceding drift $\mu_i = -C_{2,i} M_{\mathrm{aux}\,i} \Sigma_{ii}$, with $C_{2,i} = (1 + \tau/2)\kappa_i$. The same exact cancellation and scalar certificate as in Eq. (106) now yield

$$\mathcal{D}\mathcal{F}_{\mathrm{s}} \le -\upsilon_n \sum_i \varpi_i \Sigma_{ii} + C_\tau r, \qquad C_\tau = 1 + \tau/2, \tag{265}$$

where $\upsilon_n^{-1} \le P_n$ is a known rational bound. For example, if $\mathfrak{b}_n \ge \max(-\beta'')$ on the truncated interval, take

$$\upsilon_n = \min\{1/40, \kappa_{\mathrm{fast}}/(4\iota\mathfrak{b}_n)\}.$$

In fact the matrix term is at most $-\kappa_{\mathrm{fast}} \sum_i C_{2,i} p_i q_i t_i \Sigma_{ii} + C_\tau r$, and the reservoir term is at most $-\gamma \sum_i (-\beta_i'')\Sigma_{ii}/20$. Here $C_{2,i} \ge 1$ and $p_i q_i \ge 1/2$. This verifies Eq. (265) with the stated choice. The fixed-constraint majorization for $\mathcal{P}_{\mathrm{s}}$ is exactly the one already proved in Section 6.1; it does not assume hard tight equations for the smoothed root.

Choose a known $\mathfrak{m}_n \le P_n$, enlarged so that

$$|C_{2,i} M_{\mathrm{aux}\,i}| \le \mathfrak{m}_n \sqrt{n\varpi_i}, \qquad |M_{\mathrm{aux}\,i}|,\ |\beta_i'|,\ p_i q_i \le \mathfrak{m}_n.$$

Use separated gradient tests so that failure implies

$$|(\mathcal{F}_{\mathrm{s}})_i| \le \frac{\upsilon_n}{8\mathfrak{m}_n} \sqrt{\varpi_i/n} \quad \text{for all } i. \tag{266}$$

A successful test certifies a lower bound half this size, which changes only the logarithmic factors of the retained coordinate-step analysis. At such a failed state there is a known $\mathfrak{z}_n \le P_n$ with

$$z_i^{\mathrm{mov}} \le \mathfrak{z}_n \varpi_i. \tag{267}$$

Indeed the exact identity is

$$z_i^{\mathrm{mov}} = p_i q_i t_i + M_{\mathrm{aux}i}((\mathcal{F}_{\mathrm{s}})_i - \gamma \beta_i').$$

Since $\varpi_i \ge (\iota t_i + \gamma)/2$, the right side is bounded by $(2\mathfrak{m}_n/\iota + 1 + 2\mathfrak{m}_n^2)\varpi_i$ using Eq. (266). This value may be used for $\mathfrak{z}_n$.

Lemma 6.14 (Negative spectral multiplicity)**.** *Put*

$$\delta_n = \upsilon_n/16, \qquad \pi_n = \min\{1/8, \upsilon_n/(1024 C_\tau \mathfrak{z}_n)\}.$$

At a light state satisfying Eq. (266), the matrix $A = G^{-1/2}\nabla^2 \mathcal{F}_{\mathrm{s}} G^{-1/2}$ has at least $\pi_n k$ eigenvalues smaller than $-4\delta_n$. In particular $\pi_n^{-1}, \delta_n^{-1} \le P_n$. No covariance computation is part of the algorithmic assertion.

Proof. Let r count eigenvalues below $-4\delta_n$. If $r > k/4$ the result is immediate. Otherwise take $\mathscr{E}$ to be the preimage, under $G^{1/2}$, of their orthogonal complement, and apply Corollary 6.13. Write $\mathsf{P}_\Sigma = \sum_i \varpi_i \Sigma_{ii}$. Then $\mathsf{P}_\Sigma \ge k/(175\mathfrak{z}_n)$ and

$$\tfrac{1}{2}\operatorname{tr}[\nabla^2 \mathcal{F}_{\mathrm{s}} \Sigma] \ge -2\delta_n \mathsf{P}_\Sigma = -\upsilon_n \mathsf{P}_\Sigma/8.$$

On the other hand $|\nabla \mathcal{F}_{\mathrm{s}} \cdot \mu| \le \upsilon_n \mathsf{P}_\Sigma/8$, so Eq. (265) implies

$$\tfrac{1}{2}\operatorname{tr}[\nabla^2 \mathcal{F}_{\mathrm{s}} \Sigma] \le -7\upsilon_n \mathsf{P}_\Sigma/8 + C_\tau r.$$

It follows that $r \ge 3\upsilon_n k/(700 C_\tau \mathfrak{z}_n) \ge \pi_n k$. The argument also covers small k: the positive lower bound and integrality force at least one such eigenvalue. □

Reservoir phases without a new discrepancy allowance. Let $\Lambda = 10^{-6}$ and $L_0 = 1 + \lceil \log_2 \max\{2, k_0\} \rceil$. At the start of a phase record the current active count K and set

$$\gamma = \frac{\Lambda}{L_0 K}. \tag{268}$$

Keep this value fixed until the active count is at most $K/2$, then start the next phase at the new count. Stop if it is zero. Thus at each light state in a phase, $K/2 < k \le K$. There are at most L_0 phases, and $\gamma \ge \Lambda/(L_0 k_0)$ throughout.

A phase change modifies only the scalar reservoir, not $\mathcal{P}_{\mathrm{s}}$, its optimizer, the signed sum, or any coupling coefficient. Since $0 \le \beta \le 1$, each increase in the potential is at most $\gamma_{\mathrm{new}} k \le \Lambda/L_0$. The initial reservoir is also at most Λ/L_0. Therefore

$$\text{initial reservoir} + \text{all positive phase jumps} \le \Lambda. \tag{269}$$

This replaces, rather than adds to, the original initial reservoir allowance 10^{-6}. All other moves decrease their current phase potential. Their total certified decrease is bounded by the same constant as before, with Eq. (269) accounting for all jumps. Phase counters and updates cost only linear scalar work.

A bounded spectral filter with a spread guarantee. At a failed gradient test, freeze the numerical root and the exact symmetric Schur surrogate as before. Choose its accuracy so the scaled surrogate is within $\delta_n/100$ of the true scaled Hessian. Denote this surrogate by A and write $\delta = \delta_n$, $\pi = \pi_n$. By Lemma 6.14 and Weyl's inequality it has at least πk eigenvalues at most -3δ. Enlarge the known rational spectral bounds to

$$-RI \preceq A \preceq M_{\mathrm{aux}} I, \qquad M_{\mathrm{aux}} \le P_n k, \qquad 1 \le R \le P_n, \qquad 0 < \delta \le R/8.$$

The PSD shift is $B = A + 2RI \succeq RI$. The endpoint construction of Section 6.8 supplies accurate applications of B^{-1} with complete cost $\widetilde{O}(n^{\chi})$ per invocation, including sketch preparation, provided the number of outer applications and their accuracy and magnitude lengths are $O(\log(2kn))$. We change only its outer polynomial.

Lemma 6.15 (A checked flat-filter direction)**.** *Given the preceding spectral bounds, a finite random-sign filter, using polylogarithmically many applications of B^{-1}, has constant probability of producing a true generalized-curvature witness r with $\|r\|_2 = 1$ and*

$$r^{\top}\nabla^2 \mathcal{F}_{\mathrm{s}} r \le -\tfrac{1}{2}\delta r^{\top} G r, \qquad \|r\|_{\infty}^2 \le \mathsf{s}_n r^{\top} G r, \qquad \mathsf{s}_n \le P_n. \tag{270}$$

Both inequalities are verified with enclosures. Combined with the finite recursive feature and index sampler, a capped trial succeeds with conditional probability at least $1/64$ and costs $\widetilde{O}(n^{\chi})$ in arithmetic and bit operations, including unsuccessful trials.

Proof. Set $T = RB^{-1}$, so $0 \preceq T \preceq I$, and define

$$t_- = \frac{R}{2R-\delta}, \qquad t_+ = \frac{R}{2R-3\delta}, \qquad t_0 = (t_- + t_+)/2, \qquad \Delta = (t_+ - t_-)/2.$$

Here $\Delta \ge \delta/(8R)$. Choose a positive dyadic ε_f with

$$\varepsilon_f \le 1/16, \qquad \varepsilon_f^2 \le \frac{\delta\pi}{128(M_{\mathrm{aux}} + \delta)},$$

and an integer $d_f \ge \max\{2/\Delta, 8\Delta^{-2}\log(2/\varepsilon_f)\}$. All these quantities have polynomial or polylogarithmic magnitude and encoding bounds; d_f is polylogarithmic. For $j_* = \lceil d_f t_0 \rceil$ use the Bernstein polynomial

$$p_f(t) = \sum_{j=j_*}^{d_f} \binom{d_f}{j} t^j (1-t)^{d_f - j}. \tag{271}$$

It lies in $[0,1]$ on $[0,1]$. For completeness, the logarithmic moment function of a centered Bernoulli variable has value and derivative zero at zero and second derivative at most $1/4$. Twice integrating gives $\mathbb{E}\exp(t(U - \mathbb{E}U)) \le \exp(t^2/8)$. Independence and optimization of the exponential Markov bound therefore give $\Pr\{|\operatorname{Bin}(d_f, t)/d_f - t| \ge a\} \le 2e^{-2d_f a^2}$; each one-sided tail has the same bound without the factor two. It follows that

$$p_f(t) \le \varepsilon_f \ (t \le t_-), \qquad p_f(t) \ge 1 - \varepsilon_f \ (t \ge t_+).$$

For the latter bound, rounding $d_f t_0$ costs at most $1/d_f \le \Delta/2$; the two binomial tail bounds are at most $\exp(-d_f\Delta^2/2)$. Thus $F = p_f(T)$ obeys $0 \preceq F \preceq I$, is at least $1 - \varepsilon_f$ on eigenvalues of A at most -3δ, and is at most ε_f on eigenvalues of A at least $-\delta$.

Draw independent real signs $\xi \in \{\pm 1\}^k$ and let $y_{\mathrm{aux}} = F\xi$. Writing $M_f = \mathrm{tr}[F^2]$, we have

$$M_f \geq \pi k/2, \qquad \mathbb{E}\|y_{\mathrm{aux}}\|^2 = M_f, \qquad \mathbb{E}\|y_{\mathrm{aux}}\|^4 \leq 3M_f^2.$$

The fourth-moment inequality follows by expanding the quadratic form $\xi^\top F^2 \xi$; its second moment is at most $(\mathrm{tr}[F^2])^2 + 2\,\mathrm{tr}[F^4] \leq 3M_f^2$. Consequently $\Pr\{\|y_{\mathrm{aux}}\|^2 \geq M_f/2\} \geq 1/12$. Each coordinate of $F\xi$ has the sign moment-generating-function bound $\mathbb{E}e^{t y_{\mathrm{aux}\,i}} \leq e^{t^2/2}$, since its coefficient row has Euclidean norm at most one. A union bound yields

$$\Pr\{\|y_{\mathrm{aux}}\|_\infty^2 > 2\log(384k)\} \leq 1/192.$$

With probability greater than $1/16$ both estimates hold, so

$$\|y_{\mathrm{aux}}\|^2 \geq \pi k/4, \qquad \|y_{\mathrm{aux}}\|_\infty^2 \leq 2\log(384k). \tag{272}$$

The squared mass on eigenvalues at least $-\delta$ is at most $\varepsilon_f^2 k$. On this event

$$\frac{{y_{\mathrm{aux}}}^\top A y_{\mathrm{aux}}}{\|y_{\mathrm{aux}}\|^2} \leq -\delta + \frac{4(M_{\mathrm{aux}} + \delta)\varepsilon_f^2}{\pi} \leq -7\delta/8.$$

Unscale by the diagonal budget and normalize: $r = G^{-1/2} y_{\mathrm{aux}} / \|G^{-1/2} y_{\mathrm{aux}}\|$. In exact arithmetic,

$$\frac{\|r\|_\infty^2}{r^\top G r} = \frac{\max_i {y_{\mathrm{aux}}}_i^2/\varpi_i}{\|y_{\mathrm{aux}}\|^2}.$$

The weight lower bound, Eq. (268), and $k > K/2$ show that Eq. (272) makes this at most a fixed multiple of $L_0 \log(384(k_0+1))/(\Lambda\pi)$. For example one may take the known rational upper bound

$$\mathsf{s}_n \geq \frac{2^{10} L_0 \log(384(k_0+1))}{\Lambda\pi}.$$

The slack also covers numerical weights within a fixed small relative factor of the true weights. This proves the ideal spread and negative quotient with more margin than Eq. (270) requires.

Here is an explicit evaluation and precision argument. Set $v_j^{(0)} = \mathbf{1}_{\{j \geq j_*\}}\xi$ for $0 \leq j \leq d_f$ and use

$$v_j^{(r)} = (I - T)v_j^{(r-1)} + T v_{j+1}^{(r-1)}, \qquad 1 \leq r \leq d_f, \quad 0 \leq j \leq d_f - r. \tag{273}$$

Its final vector is $F\xi$. This requires $O(d_f^2)$ inverse actions, which is still a fixed polylogarithm. Section 7.9 proves the array error bound, the finite-precision acceptance tests, and the same conditional success probability. All trials have the prescribed finite caps.

□

A curvature tube retaining the maximum coordinate. The previous proof bounded every $|r_i|$ by one. Keeping its actual maximum makes the spread certificate useful.

Lemma 6.16 (Spread-sensitive curvature tube)**.** *Let r be a certified Euclidean unit curvature direction at a failed gradient state, put $\theta_0 = r^\top G r$ and $b = \|r\|_\infty$. A sufficiently small fixed inverse-polylogarithmic $a_n > 0$ makes*

$$|s| \leq a_n \min\{\delta_{\mathrm{end}}/b, (n\theta_0)^{-1/2}\} \tag{274}$$

a valid two-sided trial tube. Both candidates can be evaluated from the base optimizer with a constant number of corrected stages. The better sign decreases the true current phase potential by at least

$$P_n^{-1} \min\{{\delta_{\mathrm{end}}}^2 \theta_0 / b^2, n^{-1}\}. \tag{275}$$

If Eq. (270) holds, this is at least $1/(P_n n)$. The conclusion is valid across every coordinate-zero crossing.

Proof. Keep the directional maximum in the weighted coupling estimates. For $1 \le j \le 3$ and $j \ge 2$, respectively, they become

$$q(\mathcal{A}^{(j)} D) \le P_n b^{j-1} \sqrt{A_r}\, e(D), \qquad q(\mathcal{C}_{s^j}) \le P_n b^{j-2} A_r.$$

Indeed $|r_i|^{2j} \le b^{2j-2} r_i^2$ in weighted Cauchy–Schwarz, and $|r_i|^j \le b^{j-2} r_i^2$ in the direct term. The bound for $F_2 = \mathcal{A}'\mathsf{Q}$ is unchanged. In the temporary small-energy region of the curvature proof, completing the square in the second-order identity still gives $e(\mathsf{P}')+\mathcal{E}(\mathsf{Q}') \le P_n\sqrt{\theta_0}$. The third-order saddle identity now yields

$$|\mathcal{P}_{\mathrm{s}}'''|,\ |\mathcal{F}_{\mathrm{s}}'''|,\ |\theta'| \le P_n(b\theta_0 + \sqrt{n}\theta_0^{3/2}). \tag{276}$$

In detail the direct cubic and differentiated-coupling terms acquire $b\theta_0$; the inverse words, transfer of the dual derivative, and entropy cubic have the unchanged $\sqrt{n}\theta_0^{3/2}$ bound. The reservoir cubic is at most $P_n b\theta_0$. Thus no term without its required directional factor has been suppressed.

Use a simultaneous first-exit argument for weight comparability, $\mathcal{P}_{\mathrm{s}}'' \le M_{\mathrm{aux}0}\theta_0$, expanded light bounds, and a bounded potential sublevel. The coordinate displacement in Eq. (274) is at most $\delta_{\mathrm{end}}/4$, after fixing a_n, so endpoint distance remains at least $\delta_{\mathrm{end}}/2$. Relative variations of u_i, v_i, w_i, z_i have rate at most $P_n\sqrt{n\theta_0}$; scalar budget relative variations have rate at most $P_n b$. The same estimates used in Section 6.3, now integrated over Eq. (274), therefore close these comparisons and Eq. (276).

One additional point is needed because the Euclidean step can exceed one: boundedness of the modeled matrix must not be inferred merely from its Lipschitz bound times s. At the base the failed gradient test gives

$$|\mathcal{F}_{\mathrm{s}}'(0)| \le P_n^{-1}\sqrt{k\theta_0/n}.$$

While the temporary estimates hold, $|\mathcal{F}_{\mathrm{s}}''| \le P_n\theta_0$. For $|s| \le a_n/\sqrt{n\theta_0}$ this implies

$$|\mathcal{F}_{\mathrm{s}}(y+sr) - \mathcal{F}_{\mathrm{s}}(y)| \le P_n^{-1} a_n\sqrt{k}/n + P_n a_n^2/n < 1/2,$$

using $k \le n^2$ and a small enough prefactor. Hence the potential cannot exit a fixed larger bounded sublevel; $\|M\| \le \mathcal{P}_{\mathrm{s}} \le \mathcal{F}_{\mathrm{s}}$ then supplies the bounded-model root estimates. This closes the otherwise missing part of the first-exit argument.

In the symmetric average the linear terms cancel. The cubic remainder from Eq. (276) is at most a fixed small fraction of the negative quadratic term, because $P_n|s|(b+\sqrt{n\theta_0})$ is small relative to the certified curvature threshold. This proves Eq. (275). If $b^2 \le \mathsf{s}_n\theta_0$, its first entry is at least ${\delta_{\mathrm{end}}}^2/\mathsf{s}_n$, an inverse-polylogarithm, so the minimum is at least $1/(P_n n)$.

The joint root energy arc is at most $P_n\sqrt{\theta_0}|s|$ and its explicit-data transport integral at most $P_n\sqrt{n\theta_0}|s|$. Shrinking the same prefactor makes both meet Lemmas 6.1 and 6.2 over each whole candidate. These are the existing executable anchored corrections, not exact optimizer calls. $\mathcal{P}_{\mathrm{s}}''$ and θ are continuous at zeros; one-sided bounds in Eq. (276) integrate across any number of them. No third derivative at a zero or higher Taylor jet is used. Final common-grid rounding is taken below fixed fractions of the true decrease, energy radius, and curvature and spread margins. The longest possible step is polynomial since $\theta_0 \ge \Lambda/(4L_0k_0)$, so this remains a polylogarithmic precision requirement. □

Batched deferred attenuation. At a corrected state, collect all coordinates satisfying the near-endpoint test into one set J. Absorb their unchanged fractional signed contribution into F, preserving M, and attenuate their entire positive coupling

$$\mathcal{K}_s = \mathcal{K}_{\mathrm{rest}} + e^{-s}\mathcal{K}_J, \qquad \mathcal{K}_J(D) = \iota \sum_{i\in J} \beta_i V_i D V_i.$$

Use one path to the required tiny residual coefficient, then delete all its remaining slots exactly and correct once. Hold their reservoir contribution fixed until that deletion. The proof of Lemma 6.3 applies to this positive submap verbatim: with K_s the corresponding pair map and $v = \langle \mathsf{Q}, K_s\mathsf{P}\rangle$,

$$\mathcal{P}_s' = -v, \quad e(\mathsf{P}') + \mathcal{E}(\mathsf{Q}') \le P_n\sqrt{v}, \quad |v'| \le P_n v, \quad q(K_s D) \le C\sqrt{v}\, e(D).$$

The last inequality is the weighted Cauchy–Schwarz proof summed over J, whose second factor is bounded by the full coupling dual energy; no factor $|J|$ is lost. Likewise $K_s\mathsf{P}$ is a PSD subpair of the full coupling, so $\mathrm{tr}[\mathsf{Q}(K_s\mathsf{P})^2] \le Cv$.

If there are b_{def} nonempty batches and their total matrix potential drop is Δ_{def}, their complete correction cost in stages is

$$\widetilde{O}(b_{\mathrm{def}} + \sqrt{n b_{\mathrm{def}} \Delta_{\mathrm{def}}}) = \widetilde{O}(b_{\mathrm{def}} + \sqrt{n b_{\mathrm{def}}}). \tag{277}$$

The last equality uses the initial bound and the total phase injections in Eq. (269). All final tiny-slot deletions, grid and scalar-exponential evaluations, positivity and residual enclosures are included, as in the parent attenuation proof. Computing or applying $\mathcal{K}_J$ uses the same rank-one action bound as a full coupling, so a batch stage still costs $\widetilde{O}(n^\chi)$.

Attenuation changes no remaining coefficient, so it cannot create a new near-endpoint coordinate. A regular endpoint changes only the removed coefficient, and phase changes alter no coordinate. No charge proportional to k_0 separate optimizer corrections is inserted for simultaneous removals. Recording their coefficients and updating active index lists costs $O(k_0)$ over the whole run.

6.10 Slack-scaled grouped reduction and fast final rounding

We use the spread construction of Section 6.9 and replace the small reduction inside the preprocessing group tree. It now leaves at most twice the vector-space dimension fractional coordinates. Constant-factor zero padding restores the active-count hypothesis of the subsequent matrix algorithm. We also accelerate the last deferred-rounding step; otherwise its n^5 term would remain in the total.

A small reduction with dimensional slack. In the next lemmas D is the real vector-space dimension, not a matrix order. The columns need not be rank one or positive; the application to matrix inputs is given afterward. Fix the same admissible square multiplication exponent $\omega_0 = 2.371177$ as in Eq. (132). A fixed bilinear recursion at a strict exponent below ω_0 gives polynomial normwise rounding amplification. No new multiplication bound is used.

Lemma 6.17 (Slack-scaled approximate null directions)**.** *Let $A \in \mathbb{R}^{D\times r}$ have column norms at most two, $2D < r \le r_0 \le 4D$, and let $x \in (-1,1)^r$. Put*

$$s_i = 1 - |x_i|, \qquad \mathsf{S} = \mathrm{diag}(s_i), \qquad \mathsf{B} = A\mathsf{S},$$

and, for $\delta > 0$, define

$$J_\delta = \delta I_D + \mathsf{B}\mathsf{B}^\top, \qquad \Pi_\delta = I_r - \mathsf{B}^\top J_\delta^{-1}\mathsf{B}. \tag{278}$$

Then $0 \prec \Pi_\delta \preceq I$, $\operatorname{tr}[\Pi_\delta^2] \geq r - D$, and $\|\mathsf{B}\Pi_\delta\| \leq \sqrt{\delta}/2$. Let $\ell = \lceil \log_2(64 r_0) \rceil$, and let ξ have independent uniform real signs. If a dyadic vector $\widetilde{v}$ satisfies

$$\|\widetilde{v} - \Pi_\delta \xi\|_2 \leq \nu \leq 1/64,$$

then, with probability at least $1/16$, it passes both exact tests

$$\|\widetilde{v}\|_\infty \leq \ell, \qquad \|\widetilde{v}\|_2^2 \geq (r - D)/4. \tag{279}$$

For any vector passing the tests, $h_i = s_i \widetilde{v}_i / H$ with $H = 4\ell$ satisfies $|h_i| \leq s_i/4$.

Proof. The singular-value formula for Π_δ gives eigenvalues $\delta/(\delta + \sigma_j(\mathsf{B})^2)$ and at least $r - D$ eigenvalues equal to one. Its corresponding residual singular values are $\delta\sigma_j/(\delta + \sigma_j^2) \leq \sqrt{\delta}/2$. This calculation is only a proof; the implementation uses a ridge inverse, not a singular-value oracle.

Write $V = \Pi_\delta \xi$ and $a = \operatorname{tr}[\Pi_\delta^2]$. Independent signs give

$$\mathbb{E}\|V\|^2 = a, \qquad \mathbb{E}\|V\|^4 \leq 3a^2.$$

For the latter, expand the square of $\xi^\top \Pi_\delta^2 \xi$; its variance is at most $2\operatorname{tr}[\Pi_\delta^4] \leq 2a^2$. Cauchy–Schwarz on the event $\|V\|^2 \geq a/2$ gives its probability at least $1/12$. Each row of Π_δ has Euclidean norm at most one. The elementary estimate $\cosh t \leq e^{t^2/2}$ and exponential Markov inequality therefore imply

$$\Pr\{\|V\|_\infty > \ell/2\} \leq 2 r_0 e^{-\ell^2/8} \leq 2^{\ell-5} e^{-\ell^2/8} < 1/192.$$

Here $r_0 \geq 3$ and $\ell \geq 8$; the last expression decreases for $\ell \geq 8$, and its value at eight is $8e^{-8} < 1/192$. On the intersection of these two events, the perturbation by at most ν preserves the infinity test. Also $\|V\| \leq \sqrt{r}$, so

$$|\|\widetilde{v}\|^2 - \|V\|^2| \leq 2\sqrt{r}\nu + \nu^2 \leq (r - D)/4$$

for $r > 2D \geq 2$. This gives the second test. The intersection has probability at least $1/12 - 1/192 > 1/16$. The last assertion follows from $H = 4\ell$. □

Lemma 6.18 (A capped small cube reduction)**.** *Let $C \in \mathbb{R}^{D \times r_0}$ have $r_0 \leq 4D$ columns of norm at most two, and let $y \in [-1, 1]^{r_0}$ be a short dyadic vector. Assume the rational column entries have a common denominator and word length p_C, and y has word length p_y. Given $0 < \varepsilon < 1$ and $0 < \zeta_{\mathrm{f}} < 1/2$, assume $p_C + p_y = O(\log(2D/(\varepsilon\zeta_{\mathrm{f}})))$, with a fixed constant in this short-data bound. A finite randomized procedure returns $x \in [-1, 1]^{r_0}$ with at most $2D$ fractional coordinates and*

$$\|C(x - y)\|_2 \leq \varepsilon. \tag{280}$$

Every random-bit sequence terminates with this guarantee. Its fast attempt costs $\widetilde{O}(D^{\omega_0})$ operations, and it calls a $\widetilde{O}(D^3)$ deterministic fallback with probability at most ζ_{f}. Thus its expected cost is $\widetilde{O}(D^{\omega_0} + \zeta_{\mathrm{f}} D^3)$ in both models. The suppressed factors include logarithms of the inverse error and failure budgets. The short-data hypothesis holds uniformly in the group tree below. No exact rank decision is required.

Proof. If there are already at most $2D$ fractional coordinates, return y. Otherwise $r_0 \geq 3$. Choose a dyadic σ_{r} satisfying

$$\frac{\varepsilon}{2^{11} r_0} < \sigma_{\mathrm{r}} \leq \frac{\varepsilon}{2^{10} r_0}.$$

Set the following parameters once for the entire small call:

$$\begin{aligned} \ell &= \lceil \log_2(64 r_0) \rceil, \quad H = 4\ell, \quad L_{\delta_{\mathrm{end}}} = \lceil \log_2(2/\sigma_{\mathrm{r}}) \rceil, \\ T_* &= 128 H^2 L_{\delta_{\mathrm{end}}} + 1, \qquad \nu = \frac{\varepsilon}{2^{14} T_* r_0}, \qquad \nu^2/2 < \delta \le \nu^2. \end{aligned} \tag{281}$$

The ridge δ is dyadic. Use a dyadic coefficient grid of mesh u no larger than

$$\min\{ \frac{\sigma_{\mathrm{r}}}{2^{16} H^2}, \frac{\varepsilon}{2^{14} T_* r_0}, 2^{-p_y} \}, \tag{282}$$

within a factor two of this minimum. Enlarge p_y so every inherited coefficient is a multiple of 2^{-p_y}. The grid therefore contains all initial coefficients exactly.

Keep the original y for a possible fallback. Freeze every current coordinate at distance at most σ_{r} from an endpoint by setting it to that nearest endpoint. This includes the initial scan. Each coordinate is frozen at most once. If at most $2D$ remain, stop. Otherwise form Eq. (278) for the current active columns and current slacks. Draw signs and compute $\widetilde{v}$ to the accuracy in Lemma 6.17. Reject a direction failing Eq. (279). For a passing vector put $h = \mathsf{S}\widetilde{v}/H$. Both $x + h$ and $x - h$ are strictly inside the active cube. Define the barrier

$$\Phi(x) = \sum_{i \text{ active}} \phi(x_i), \qquad \phi(t) = -\log(1 - t^2). \tag{283}$$

Compute the short dyadic approximation $\widehat{g}$ to $g = \nabla\Phi(x)^\top h$ specified below. Choose $\tau = 1$ if $\widehat{g} \ge 0$, and $\tau = -1$ otherwise. Round $x + \tau h$ inward, toward zero, to the common grid, freeze all newly triggered coordinates, and repeat.

Progress with finite endpoints. For $s_i = 1 - |x_i|$ and $|h_i| \le s_i/4$, the exact scalar identity is

$$\begin{aligned} &\frac{\phi(x_i + h_i) + \phi(x_i - h_i)}{2} - \phi(x_i) \\ &\quad = -\frac{1}{2}\log(1 - \frac{h_i^2}{(1 - x_i)^2}) - \frac{1}{2}\log(1 - \frac{h_i^2}{(1 + x_i)^2}) \ge \frac{h_i^2}{2 s_i^2}. \end{aligned} \tag{284}$$

The logarithm bound follows from $-\log(1-z) \ge z$ on $[0, 1)$. For the implemented sign, Lemma 7.17 gives an increase of at least $(2/25 - 1/256)(r - D)/H^2$ before rounding. Before inward rounding, every active endpoint distance is at least $3\sigma_{\mathrm{r}}/4$. On the inward segment $|\phi'| \le 4/\sigma_{\mathrm{r}}$, so the grid loses at most $4ru/\sigma_{\mathrm{r}}$. Eq. (282) and $r \le 2(r - D)$ bound this loss by $(r - D)/(8192 H^2)$. Since $2/25 - 1/256 - 1/8192 > 1/16$, the net increase is at least

$$\frac{r - D}{16 H^2} > \frac{D}{16 H^2}. \tag{285}$$

For analysis, assign every frozen coordinate the finite value $Q_{\delta_{\mathrm{end}}} = \log(2/\sigma_{\mathrm{r}})$ and leave active values equal to $\phi(x_i)$. This defines a capped bookkeeping potential $\mathcal{V}$. After a move, a newly frozen coordinate was still at distance at least $3\sigma_{\mathrm{r}}/4$; replacing its barrier value by $Q_{\delta_{\mathrm{end}}}$ cannot decrease $\mathcal{V}$. Initial freezes cause no problem: start the bookkeeping after that scan. Always

$$0 \le \mathcal{V} \le r_0 Q_{\delta_{\mathrm{end}}} \le 4 D L_{\delta_{\mathrm{end}}}.$$

Consequently fewer than $64 H^2 L_{\delta_{\mathrm{end}}} + 1 < T_*$ successful steps are possible while $r > 2D$. This is a deterministic bound for every sequence of accepted directions, not a drift-only stopping argument. The active scalar barrier is never evaluated at an exact endpoint.

The residual is a separate budget. Lemma 6.17 gives at a successful step

$$\|Ah\|_2 \le \frac{\sqrt{r}}{H}\left(\frac{\sqrt{\delta}}{2} + 2\nu\right) \le \frac{5r_0\nu}{8},$$

where $H \ge 4$ and $\sqrt{r} \le r_0$. Coefficient rounding adds at most $2r_0u$. All initial and subsequent endpoint freezes together add at most $2r_0\sigma_{\mathrm{r}}$. Thus any fast output obeys

$$\begin{aligned}\|C(x-y)\|_2 &\le 2r_0\sigma_{\mathrm{r}} + T_*(5r_0\nu/8 + 2r_0u)\\ &\le \varepsilon\left(\frac{2}{2^{10}} + \frac{5/8+2}{2^{14}}\right) < \varepsilon/128.\end{aligned} \tag{286}$$

No error term is charged per original leaf at every group update. All columns in this small call have norm at most two, so this bound also covers repeated, zero, or linearly dependent columns.

A finite rational direction oracle. At a current state, $\delta I \preceq J_\delta \preceq (\delta + 4r_0)I$. Form B, J_δ, and a short inverse R by the residual-certified Newton–Schulz construction already used in the manuscript. Require, for example, a true residual enclosure

$$\|I - J_\delta R\|_{\mathrm{F}} \le \frac{\delta\nu}{64r_0^{3/2}}.$$

The known inverse bound $\|J_\delta^{-1}\| \le \delta^{-1}$ converts this to an inverse-error bound. Applying $\xi - \mathsf{B}^\top R\mathsf{B}\xi$ and rounding to a common dyadic vector grid, with sufficiently smaller fixed error allocations, gives $\|\widetilde{v} - \Pi_\delta\xi\|_2 \le \nu$. A square-root factor in an error bound may be replaced by r_0 to use rational thresholds. The product residual includes construction error in J_δ; it is not a residual for a succession of different matrices.

Section 7.2 proves that this residual target is attainable in logarithmic precision. The inverse costs $\widetilde{O}(D^{\omega_0})$ arithmetic operations, and each subsequent direction trial costs $\widetilde{O}(D^2)$.

Short-precision barrier sign selection. The directional derivative has the cancellation formula

$$g = \sum_{i \text{ active}} \frac{2x_ih_i}{1-x_i^2} = \sum_{i \text{ active}} \frac{2x_i\widetilde{v}_i}{H(1+|x_i|)}. \tag{287}$$

Choose $q = \lceil \log_2(512H^2)\rceil$. Round each summand in Eq. (287) to the dyadic grid of mesh 2^{-q} with absolute error at most 2^{-q}, and sum these rounded values exactly to obtain $\widehat{g}$. Then $|\widehat{g} - g| \le r/(512H^2) \le (r-D)/(256H^2)$. The sign rule above gives $\tau g \ge -(r-D)/(256H^2)$, including when $\widehat{g} = 0$. Lemma 7.17 proves the stated progress and bit bounds. Each summand uses one rational division, and the common-grid sum uses only logarithmic-length integers. At $x_i = 0$ the summand is zero, so the formula applies without a separate case. Inward rounding restores the prescribed coefficient grid.

Caps, verification, and fallback. At any unsuccessful search for a direction, permit at most

$$R_* = 64\lceil \log_2(2T_*/\zeta_{\mathrm{f}})\rceil$$

independent sign trials. If they all fail, or if any required accuracy or consistency check fails, abandon the fast attempt and return the deterministic fallback from the original inherited y. Also cap the number of successful steps at T_*. Upon a prospective fast return, check the cube, the fractional

count, and $\|C(x-y)\|_2^2 \le \varepsilon^2$ by exact rational arithmetic. The common column denominator and coefficient grid make this last check $\widetilde{O}(Dr_0)$ work on short entries. Equation Eq. (286) leaves ample slack.

For fallback split the at most $4D$ columns into two sets of at most $2D$. Apply the deterministic small reducer in the proof of Lemma 5.2 to each set, with its inherited coefficients and residual $\varepsilon/2$. It leaves at most D fractional coordinates per set. Its combined cost is $\widetilde{O}(D^3)$ and its residual is at most ε. This fallback is short-precision by the preceding whole-subsequence ridge-update estimate, not exact elimination with growing word length.

All matrix and scalar tolerances above guarantee their tests in the stated caps; only the random-direction event may fail to occur. Conditionally on every previous history, a sign trial succeeds with probability at least $1/16$. Thus the probability of a fallback is at most

$$T_*(15/16)^{R_*} \le \zeta_{\mathrm{f}}.$$

Every other fast-attempt cost is bounded deterministically by T_*R_* times the displayed polynomial work, and T_*, R_* are polylogarithmic in the parameters. The resulting expected bound is as stated. Every random-bit sequence terminates, and every returned vector has the required verified or deterministic guarantee. □

Lemma 6.19 (Grouped reduction with a factor-two slack)**.** *Under the short-data hypotheses of Lemma 5.2, a zero-error randomized algorithm finds $x \in [-1,1]^m$ with at most $2D$ fractional coordinates and $\|\sum_i x_i b_i\|_2 \le \eta$ in expected cost*

$$\widetilde{O}(mD + D^{\omega_0}) \tag{288}$$

in both operation models. Original long input length is additive when this routine follows the one-time matrix input conversion. Every random-bit sequence terminates with the same residual guarantee.

Proof. Retain the balanced tree, the precomputed sums W_J, and the power of two $M_{\mathrm{grp}} \in [m, 2m)$ from Lemma 5.2. Maintain at most $2D$ active groups, not D. Splitting their nonsingleton groups gives at most $4D$ groups. At each of at most $R_{\mathrm{tree}} = \max\{1, \lceil \log_2 m \rceil\}$ levels, apply Lemma 6.18 to their normalized sums W_J/M_{grp} and inherited coefficients, with

$$\varepsilon = \eta/(M_{\mathrm{grp}} R_{\mathrm{tree}}), \qquad \zeta_{\mathrm{f}} = (2mD)^{-20}/R_{\mathrm{tree}}.$$

All normalized columns have norm at most two. Coefficients are copied to children before rounding; this changes no signed sum. Each level's returned residual is at most η/R_{tree} in the unnormalized sum, on every random-bit sequence. Summing gives the same single allowance η. After the last level the remaining groups are singletons, so at most $2D$ original coordinates are fractional.

Precomputing all group sums and writing final leaf coefficients cost $\widetilde{O}(mD)$. No accepted small-reduction step scans a group's leaves. There are logarithmically many fast calls of cost $\widetilde{O}(D^{\omega_0})$. The total expected fallback contribution is at most $(2mD)^{-20}\widetilde{O}(D^3)$, hence is smaller. This remains valid conditionally on inherited coefficients and previous random choices. All group entries and coefficient grids have $O(\log(2mD/\eta))$ bits up to fixed multipliers. In particular, the new coefficient-grid exponent is the maximum of the inherited exponent and the two accuracy exponents in Eq. (282); it is not a multiple of the inherited exponent. These latter exponents have one uniform logarithmic upper bound across all tree levels. Thus the coefficient lengths do not grow multiplicatively through the tree. Lemma 7.17 implements each barrier sign choice on the same short data with $\widetilde{O}(D)$ bit operations. Thus the bit and arithmetic totals have the same exponent. □

Speeding up the unchanged final rounding. A second replacement is needed below n^5; this one changes only numerical linear algebra, not the rounding rule.

Lemma 6.20 (Fast evaluation in deferred rounding)**.** *Under the hypotheses of Lemma 5.3, its deterministic rounding rule and discrepancy allowance can be obtained using $\widetilde{O}(kn^{\omega_0})$ arithmetic and bit operations.*

Proof. Retain the heavy/small split, scalar coefficients $G_i^{\pm}$, and conditional trace-exponential estimator in that lemma. Its proof gives known bounds $O(\log(2n))$ for both Hermitian exponential arguments at every accepted prefix and either next candidate. For example, $\|\theta E\| \le \log(4n)+\ell/16$ follows from positivity of the remaining $G_i^{\pm}$ and the estimator bound. A candidate changes this by at most $1/2$. A conservative integer $M = 4\ell + 4$ therefore bounds all argument norms, including the stipulated evaluation errors.

For $\|A\| \le M$, evaluate e^A by the factorial Taylor recurrence

$$T_0 = I, \quad T_j = T_{j-1}A/j, \quad E_J = \sum_{j=0}^{J} T_j.$$

Its operator tail is at most $e^M M^{J+1}/(J+1)!$. Since $e < 3$ and $(J+1)! \ge ((J+1)/e)^{J+1}$, taking, for example,

$$J = \lceil 16M + 2\log_2(512\max\{1,k\}n)\rceil$$

makes the trace tail smaller than $1/(256\max\{1,k\})$. This uses $J = O(\log(2kn))$ order-n multiplications. A recurrence error inserted at order i is propagated by a product bounded by $M^{j-i}i!/j! \le e^M$; summing all errors adds only a polynomial factor in n, k. The exact term and partial-sum norms are at most e^M, also polynomial. Stable fixed-recursion multiplication and a sufficiently small inverse-polynomial roundoff therefore enclose each candidate estimator to the original additive tolerance. This requires only $O(\log(2kn))$ accuracy and magnitude bits, with fixed guard factors. No order-n eigendecomposition is inserted here.

The rank-one formula for $G_i^{\pm}$ uses only scalar exponentials and logarithms, evaluated once per coefficient. For small traces, use the equivalent divided series discussed in the proof of Lemma 5.3. Maintain the two remaining matrix sums and the signed prefix by rank-one additions, costing $O(kn^2)$ altogether; their cumulative errors over at most k updates are bounded by k times the prescribed unit error. Choosing that unit smaller by this known polynomial keeps the estimator enclosure valid. There is no full reconstruction for every candidate and no multiplying of conditioning factors over the sequence.

A constant number of candidate trace evaluations per coordinate now costs $\widetilde{O}(n^{\omega_0})$ instead of $\widetilde{O}(n^3)$. The scalar work and rank-one updates are smaller. Choose the smaller candidate using the same conservative error allowance as in Lemma 5.3. Its conditional expectation argument and all errors then remain unchanged, including

$$\|\sum_i (\sigma_i - t_i)C_i\| \le 21\varepsilon_{\mathrm{round}}/128 < \varepsilon_{\mathrm{round}}/2.$$

The matrix algebra uses complex Hermitian inputs directly; real bilinear products compute their real and imaginary parts at constant factor overhead. This proves both operation bounds. □

6.11 Amortized gradient sweeps and regular endpoint continuation

Frozen-root sweeps and charged endpoint continuation reduce the expected cost of corrected local and face stages to $\widetilde{O}(n^{\chi+1})$.

As in Section 6.10, first use the slack-scaled grouped reduction and pad to $N = 2^{\lceil \log_2(2n) \rceil} < 4n$. The initial retained count satisfies $k_0 \le 2n^2 \le N^2$. All lemmas in this subsection refer to the padded order N, with $\rho = \varepsilon_{\mathrm{scale}}/N$ and the same cutoff δ_{end}, smoothing temperature τ_{s}, and reservoir phases. Write P_N for a known fixed polynomial in $\log(2N)$, enlarged finitely many times. Its coefficients can be obtained from the coefficient sums, truncated budget bounds, and root estimates already proved; no numerical norm estimate or unknown spectrum is used to choose it. The base lower bounds on the nonzero input scale and $\gamma \ge \Lambda/(L_0 k_0)$ are unchanged.

Positive auxiliary penalties and offsets. We enlarge the class of auxiliary evaluation problems, not the signing potential. Let $\mathcal{K}$ be a positive self-adjoint coupling of the retained rank-one form, let M be Hermitian, and let the matrix pairs $\Omega \succeq \rho E$ and $D \succeq 0$ be fixed. Define

$$\mathcal{R}_{\Omega,D} = \min_{\mathsf{P}=(U,V)\succ 0} \{\langle \Omega, \mathsf{P}\rangle + f_{\tau_{\mathrm{s}}}(\mathcal{C}(\mathsf{P}) + D)\}, \qquad f_{\tau_{\mathrm{s}}}(C) = \tau_{\mathrm{s}} \log \mathrm{tr}[e^{C/\tau_{\mathrm{s}}}], \tag{289}$$

where C denotes the block diagonal matrix associated to its pair and $\mathcal{C}(\mathsf{P}) = (U^{-1} + M + \mathcal{K}(V), V^{-1} - M + \mathcal{K}(U))$. All instances below have bounded $\|M\|, \|D\|, \|\Omega\|$ and a uniform constant upper bound on their optimum.

Lemma 6.21 (Uniform root machinery with positive additions)**.** *The bounded-model root, joint-energy ball, and inexpensive anchored correction statements hold for Eq. (289), with the same logarithmic losses. A corrected stage costs $\widetilde{O}(N^{\chi})$. The residual is*

$$(\Omega - \mathscr{J}_{\mathsf{P}}\mathsf{Q},\ \Pi_{\mathcal{T}}(\tau_{\mathrm{s}} \log \mathsf{Q} - \mathcal{C}(\mathsf{P}) - D)), \qquad \mathrm{tr}[\mathsf{Q}] = 1.$$

The joint Jacobian and metric are precisely $\mathbb{A}$ and $\mathbb{D}$ in Eq. (207) and Eq. (208). If only Ω, D vary, the Jacobian has no additional explicit-data derivative.

Proof. Convexity and attainment follow by the same operator-convex inverse and positive trace-penalty argument as for $\mathcal{P}_{\mathrm{s}}$. At a minimizer the Gibbs multipliers satisfy

$$W = U(\Omega_U + \mathcal{K}(Z))U, \qquad Z = V(\Omega_V + \mathcal{K}(W))V, \qquad \mathrm{tr}[W + Z] = 1.$$

They are positive definite. In particular $W \succeq \rho U^2$, $Z \succeq \rho V^2$, and $\rho\,\mathrm{tr}[U^2 + V^2] \le 1$. The value upper bound, $D \succeq 0$, and bounded M bound the primal inverses and $\mathcal{K}(U), \mathcal{K}(V)$ above. Thus $U, V \succeq a_0 I$, $\mathsf{Q} \succeq a/N$, and all the weighted bounds used in Lemma 6.1 hold. Strict convexity of the weighted inverse terms proves uniqueness.

The additional trace penalty is linear in P and the offset is constant in P. They therefore leave the joint Jacobian unchanged. Its coercivity identity is algebraic on the trace-zero space and does not require $\Omega = \rho E$. The inverse-word and resolvent proofs of the joint-energy ball apply without change. For the inexpensive root inverse, the complete-positive Schwarz proof uses $\mathcal{K}(W) \preceq V^{-1} Z V^{-1}$ and its other-block counterpart; these still follow from the displayed dual identities. Consequently $\mathscr{E} \preceq J_s = \mathscr{E} + \mathscr{J} Q_{\mathsf{Q}} \mathscr{J} \preceq (1 + C/\tau_{\mathrm{s}})\mathscr{E}$ remains valid. The explicit Lyapunov inverse and rank-one actions implement the same anchored correction and its residual certification. Positive additions have only a fixed number of rank-one terms in the application below. □

Charging regular endpoints to their true decrease. We first treat an exact trigger to $+1$. At the old root put $d = 1 - y_i \ge \delta_{\mathrm{end}}$, $\beta = \beta(y_i)$, $a = \mathrm{tr}[UV_i]$, $b = \mathrm{tr}[VV_i]$, $p = \mathrm{tr}[WV_i]$, and

$q = \mathrm{tr}[ZV_i]$. Let $\mathcal{K}_i(M_{\mathrm{aux}}) = \iota\beta V_i M_{\mathrm{aux}} V_i$ and let the target face have $M_+ = M + dV_i$, $\mathcal{K}_+ = \mathcal{K} - \mathcal{K}_i$. Define pairs

$$\mathsf{A} = (\mathcal{K}_i(Z), \mathcal{K}_i(W)), \qquad E_i = ((\iota\beta b - d)V_i, (\iota\beta a + d)V_i), \qquad v_0 = \langle \mathsf{A}, \mathsf{P}\rangle = \iota\beta(bp + aq). \tag{290}$$

An exact trigger makes $E_i \succeq 0$. Both A and E_i have bounded operator norm. Also $v_0 \le C$, because $\mathcal{K}_i(\mathsf{P})$ is a positive subpair of the full coupling at a bounded-model root.

Lemma 6.22 (A charged positive homotopy for a regular endpoint)**.** *Let $\Delta = \mathcal{P}_{s,\mathrm{old}} - \mathcal{P}_{s,\mathrm{new}}$ for a regular endpoint, with the conservative score tolerance tightened as below. Then $\Delta \ge a\rho{\delta_{\mathrm{end}}}^3$ and there is an executable continuation from the old root to the new one with at most*

$$\widetilde{O}(1 + \sqrt{N\Delta}) \tag{291}$$

corrected stages. All regular endpoint continuations together use $\widetilde{O}(N)$ stages, including their initial and final corrections. No cold initializer is invoked after these endpoints.

Proof. We use a quantitative consequence of Golden–Thompson [Tro12]. If $0 \preceq E' \preceq M_{\mathrm{aux}} I$, $M_{\mathrm{aux}} \ge 1$, $0 < \tau \le 1$, and $\mathsf{Q}_C = e^{C/\tau}/\,\mathrm{tr}[e^{C/\tau}]$, then

$$f_\tau(C) - f_\tau(C - E') \ge \frac{\tau(1 - e^{-M_{\mathrm{aux}}/\tau})}{M_{\mathrm{aux}}}\,\mathrm{tr}[\mathsf{Q}_C E'] \ge \frac{\tau}{2M_{\mathrm{aux}}}\,\mathrm{tr}[\mathsf{Q}_C E']. \tag{292}$$

Indeed $e^{-E'/\tau} \preceq I - (1 - e^{-M_{\mathrm{aux}}/\tau})E'/M_{\mathrm{aux}}$ by scalar convexity on $[0, M_{\mathrm{aux}}]$. Golden–Thompson gives $\mathrm{tr}[e^{(C-E')/\tau}] \le \mathrm{tr}[e^{C/\tau}e^{-E'/\tau}]$; then use $-\log(1-r) \ge r$ for $0 \le r < 1$. This proof applies directly to the block diagonal pairs and to complex Hermitian matrices.

At the old P, the new constraint pair is $\mathcal{C}_{\mathrm{old}} - E_i$. Testing this P in the new minimization and using Eq. (292) yields

$$\Delta \ge \frac{\tau_{\mathrm{s}}}{2M_{\mathrm{aux}}}\langle \mathsf{Q}, E_i\rangle.$$

The self-slot in the old dual equation gives $q \ge \iota\beta b^2 p$. Since $\iota\beta b \ge d \ge \delta_{\mathrm{end}}$ and $\beta \le 1$, $b \ge \delta_{\mathrm{end}}/\iota$, whence

$$v_0 = \iota\beta bp + \iota\beta aq \le (1 + \iota/{\delta_{\mathrm{end}}}^2)\langle \mathsf{Q}, E_i\rangle \le P_N\Delta. \tag{293}$$

The separate Sherman–Morrison estimate in Lemma 6.3 gives $\Delta \ge a\rho{\delta_{\mathrm{end}}}^3$. We retain this estimate for counting endpoint events; Eq. (293) supplies the additional path charge.

For $r = e^{-s}$ consider the auxiliary values

$$F(s) = \min_{\mathsf{P}}\{\rho\,\mathrm{tr}[\mathsf{P}] + r\langle \mathsf{A}, \mathsf{P}\rangle + f_{\tau_{\mathrm{s}}}(\mathcal{C}_{\mathrm{new}}(\mathsf{P}) + rE_i)\}, \tag{294}$$

where $\mathcal{C}_{\mathrm{new}}$ denotes the full constraint pair for the target face. At $r = 1$ the old root is an exact root of this problem: its constraint pair equals the old one, and $\mathscr{J}_+\mathsf{Q} = \mathscr{J}_{\mathrm{old}}\mathsf{Q} + (\mathcal{K}_i(Z), \mathcal{K}_i(W)) = \rho E + \mathsf{A}$. The initial value is $\mathcal{P}_{s,\mathrm{old}} + v_0$; the limiting value at $r = 0$ is $\mathcal{P}_{s,\mathrm{new}}$. Both positive additions decrease with s, so

$$F'(s) = -v(s), \qquad v(s) = r\langle \mathsf{A}, \mathsf{P}(s)\rangle + r\langle \mathsf{Q}(s), E_i\rangle \ge 0, \qquad \int_0^\infty v(s)ds = \Delta + v_0 \le P_N\Delta.$$

All these problems have the uniform bounds in Lemma 6.21.

The derivative forcing is $F_1 = -rE_i$ and $F_2 = -r\mathsf{A}$ in the convention of Eq. (212). The entropy variance bound gives $|\langle F_1, \mathsf{B}\rangle| \le P_N\sqrt{v}\mathcal{E}(\mathsf{B})$, since $\|rE_i\| \le C$ and $\mathrm{tr}[\mathsf{Q}(rE_i)^2] \le Cr\langle \mathsf{Q}, E_i\rangle$. For

the other forcing use $r\mathsf{A} \preceq \Omega = \rho E + r\mathsf{A} \preceq (U^{-1}WU^{-1}, V^{-1}ZV^{-1})$. For example, after congruence by $U^{1/2}$, write $J_U = U^{1/2} r\mathsf{A}_U U^{1/2} \preceq T_U$. Then

$$|\operatorname{tr}[r\mathsf{A}_U D_U]| \le \sqrt{\operatorname{tr}[J_U]}\sqrt{\operatorname{tr}[J_U \widehat{D}_U^2]}.$$

Summing both blocks proves $|\langle r\mathsf{A}, D\rangle| \le \sqrt{v/2}\, e(D)$. The joint energy identity therefore gives

$$e(\mathsf{P}') + \mathcal{E}(\mathsf{Q}') \le P_N\sqrt{v}, \qquad |v'| \le P_N v. \tag{295}$$

For the second inequality differentiate v; its two nontrivial pairings are bounded by the same two estimates. The normalized metric and Jacobian variation rate is at most $P_N\sqrt{Nv}$: there is no explicit coupling derivative on this path.

As in the charged attenuation proof, compute an upper approximation $v \le \bar{v} \le C(v + 1/N)$ and take increments within fixed factors of $a/[P_N(1 + \sqrt{N\bar{v}})]$. Each increment stays in the joint Newton and anchor-transport neighborhood. Over a polylogarithmic horizon L_* its count is at most $P_N(L_* + \sqrt{NL_* \int v}) + 1$. Choose the final r_* below the root cache floor divided by a fixed polynomial in N, k_0; then set $r = 0$ and correct once. Direct sensitivity in r, including at zero, is polynomial because its two fixed forcing matrices are bounded and the joint inverse has a polynomial norm bound. This proves Eq. (291).

We detail conservative triggers and numerical coefficients. A trigger may have $\iota\beta b - d \ge -2\delta_e$. Replace this coefficient by its nonnegative part, and round the coefficients of A, E_i to nonnegative dyadics, at a smaller prescribed accuracy. The resulting positive pairs $\widehat{\mathsf{A}}, \widehat{E}_i$ differ from the exact pairs by $P_N\delta_e$ in norm. With $b \ge \delta_{\mathrm{end}}/(2\iota)$, the same self-slot argument gives $v_0 \le (1 + 2\iota/{\delta_{\mathrm{end}}}^2)\langle \mathsf{Q}, E_i^+\rangle$. The uniform repair and Eq. (292) yield $v_0 \le P_N(\Delta + \delta_e)$. Tighten δ_e below both the preceding $a\rho{\delta_{\mathrm{end}}}^3$ margin and a sufficiently small fixed inverse-polynomial fraction of every root radius. Then the original regular decrease remains positive and $v_0 \le P_N\Delta$ still holds.

At the old root, the residual for the rounded initial auxiliary problem is $P_N\delta_e$. The initial inverse can be built there using $\mathscr{J}_+$, the old $\mathscr{E}$, and the old entropy operator: the inequalities $\mathcal{K}_+(W) \preceq \mathcal{K}(W)$ and $\mathcal{K}_+(Z) \preceq \mathcal{K}(Z)$ prove the same root Schur preconditioning bound directly. Thus it is not assumed to follow from an unproved off-root rule. The joint ball, small initial residual and the usual inverse perturbation argument supply a constant number of initial corrections. Their value changes are $P_N\delta_e$. For the exact rounded auxiliary path its total drop is at most $P_N\Delta$, after further tightening that accuracy. This makes the path executable without knowing any exact old multiplier.

There are $\widetilde{O}(N)$ regular events. Their true matrix drops sum to at most a constant: their reservoir removals are nonnegative, all other signing moves decrease their phase potential, and the total positive phase injections are at most Λ. Cauchy–Schwarz in Eq. (291) gives $\widetilde{O}(I_{\mathrm{reg}} + \sqrt{NI_{\mathrm{reg}} \sum \Delta}) = \widetilde{O}(N)$ stages. The negative endpoint exchanges the two blocks. □

A weighted parameter neighborhood of one light root. The coordinate counts can be improved without changing the scalar budget. At a light base state y^0 with corrected root $z_0 = (\mathsf{P}_0, \mathsf{Q}_0)$, fix the current reservoir coefficient and put

$$\kappa_i^0 = \iota(v_i^0 w_i^0 + u_i^0 z_i^0) + \gamma, \qquad \varphi_i(t) = (w_i^0 - z_i^0)t + \kappa_i^0 \beta(t). \tag{296}$$

Each φ_i is concave on the truncated interval. Its derivative is the true gradient coordinate evaluated at the frozen root, with only its scalar coordinate argument changed.

Lemma 6.23 (Frozen-root weighted neighborhood)**.** *There is a known P_N with the following property. For y, y^0 in the cube truncated at endpoint distance $\delta_{\mathrm{end}}/2$, let*

$$d_0(y, y^0) = \sum_i \sqrt{\kappa_i^0}|y_i - y_i^0|.$$

If $d_0 \le a/(P_N^2\sqrt{N})$, then the optimizer at y is obtainable from z_0 in $\widetilde{O}(N^\chi)$ work, using the same base inverse. Its root distance and gradient error satisfy

$$\begin{aligned} \|z(y) - z_0\|_{\mathbb{D}_0} &\le P_N d_0, \\ |g_i(y) - \varphi_i'(y_i)| &\le P_N^2\sqrt{\kappa_i^0}\, d_0, \\ \tfrac{1}{2}\kappa_i^0 &\le \iota t_i(y) + \gamma \le 2\kappa_i^0. \end{aligned} \tag{297}$$

It is not required that the whole coordinate segment stay light. All constants are uniform in the number of coordinates moved and in real or complex Hermitian inputs.

Proof. At the light base, $u_i^0, v_i^0 \le P_N$. On the entire truncated scalar interval, β, β' and $1/\beta_i(y_i^0)$ have polylogarithmic bounds. These facts, rather than a unit bound on a coordinate displacement, suffice below.

For a fixed test scalar t the first forcing is $F_{1,i} = ((1 + \iota\beta'(t)v_i)V_i, (-1 + \iota\beta'(t)u_i)V_i)$. At the base, and throughout a sufficiently small joint-energy ball,

$$\mathrm{tr}[\mathsf{Q}F_{1,i}^2] \le P_N\nu_i(w_i + z_i) \le P_N(v_i w_i + u_i z_i) \le P_N\kappa_i^0.$$

Use $u_i, v_i \ge a_0\nu_i$ and constant-factor Loewner comparisons with the base matrices. The entropy variance inequality bounds its dual energy norm by $P_N\sqrt{\kappa_i^0}$. For $F_{2,i} = \iota\beta'(t)(z_i V_i, w_i V_i)$, the weighted coupling proof at the base gives dual energy norm at most $P_N\sqrt{\kappa_i^0}$: replace its coefficient ratio by $|\beta'(t)|^2/\beta(y_i^0) \le P_N$. At nearby positive P, Q the scalars w_i, z_i remain within fixed factors of their base values and the metric remains comparable. The same estimate follows by absolute values of the two block pairings. This argument does not assert an order comparison of noncommuting products from separate Loewner bounds.

Integration at the fixed z_0 therefore gives

$$\|\mathcal{E}_{\mathrm{KKT,y}}(z_0)\|_{\mathbb{D}_0^{-1}} \le P_N d_0.$$

The optimizer-dependent Jacobian variation is that of Lemma 6.1. Its explicit coupling change is bounded by $P_N\sqrt{N}d_0$ in the normalized joint norm: use the just stated single-coordinate weighted coupling estimate, pair its value with a trace-zero dual test through the relative entropy bound, and sum the absolute coordinate displacements. In a joint-energy ball this yields

$$\|\mathbb{D}_0^{-1/2}(\mathcal{E}'_{\mathrm{KKT,y}}(z) - \mathbb{A}_0)\mathbb{D}_0^{-1/2}\| \le P_N\sqrt{N}(\|z - z_0\|_{\mathbb{D}_0} + d_0).$$

The square of P_N in the allowed radius pays for both the residual constant and the Jacobian Lipschitz constant. Choose a sufficiently small. The chord map $z \mapsto z - \mathbb{A}_0^{-1}\mathcal{E}_{\mathrm{KKT,y}}(z)$ is a contraction on a ball of radius $2P_N d_0$ about z_0, maps the ball to itself, and keeps it in the positive cones and trace-one affine space. This proves root existence there and the first bound in Eq. (297). The convex minimization has a unique root, so this is its optimizer. The inexpensive approximate base inverse and additive-error chord iteration give the executable correction with the preceding

short precision. This proof also establishes boundedness of the model; no large signed-sum change is presumed harmless from its Euclidean coordinate length alone.

The derivative in root variables of $g_i(z,t)$ is $\langle F_{2,i}, \Delta\mathsf{P}\rangle + \langle F_{1,i}, \Delta\mathsf{Q}\rangle$. The two dual energy bounds just proved hold along the straight segment inside the base ball. Integrating gives the gradient error in Eq. (297). Positive scalar trace comparisons give the stated bounds on $\iota t_i + \gamma$. The proof uses only continuous first and second derivatives at coordinate zero. □

Two separated gradient thresholds. Write $\kappa_i = \iota t_i + \gamma$, so $\varpi_i = (-\beta_i'')\kappa_i \geq \kappa_i/2$. Choose a positive rational inverse-polylogarithmic h_N small enough that

$$|g_i| \leq h_N \sqrt{\kappa_i/N} \quad \text{for every } i \tag{298}$$

implies the failed-gradient hypothesis of Lemma 6.14. One may take $h_N = \upsilon_N/[64(\mathfrak{m}_N + 1)]$ in the notation of that proof. Put $l_N = h_N/64$. A separated numerical high test either certifies some $|g_i| \geq (h_N/2)\sqrt{\kappa_i/N}$ or certifies Eq. (298). For instance compare an approximation with error at most $(h_N/8)\sqrt{\kappa_i/N}$ against $(3h_N/4)\sqrt{\kappa_i/N}$, with constant slack for square-root evaluation.

Lemma 6.24 (A gradient sweep using one root)**.** *A successful high test at a light root can be handled by a deterministic gradient sweep. It ends either saturated, with true potential decrease at least $1/(P_N N)$, or unsaturated, with all its remaining nondeferred coordinates satisfying*

$$|g_i|/\sqrt{\iota t_i + \gamma} \leq h_N/(8\sqrt{N}). \tag{299}$$

The sweep uses one final optimizer correction and a single initial scan, of total cost $\widetilde{O}(N^\chi)$, plus $O(N^2)$ work for each coordinate update. Over the whole signing trajectory the number of these coordinate updates is at most k_0 plus the number of saturated sweeps. Every sweep has a positive inverse-polynomial decrease.

Proof. At its base store Eq. (296) and dyadic weights $\sqrt{\kappa_i^0} \leq r_i \leq 2\sqrt{\kappa_i^0}$. Use an arc budget A_*, a positive dyadic within fixed factors of

$$\frac{l_N}{C P_N^2 \sqrt{N}}, \tag{300}$$

where the fixed C is sufficiently large for Lemma 6.23 and the margins below. Track $Z = \sum_i r_i |y_i - y_i^0|$. The low test selects coordinates with frozen derivative magnitude at least $2l_N\sqrt{\kappa_i^0/N}$, evaluated with error at most $(l_N/4)\sqrt{\kappa_i^0/N}$. Absence of this test bounds every such magnitude by $(9l_N/4)\sqrt{\kappa_i^0/N}$.

On selecting i, move monotonically downhill for φ_i toward the endpoint target $\pm(1 - \delta_{\mathrm{end}}/2)$, stopping early only when the arc budget is exhausted. Concavity ensures that the frozen derivative keeps the same sign and its magnitude does not decrease along this move. The gradient error in Eq. (297) is less than $(l_N/8)\sqrt{\kappa_i^0/N}$, by Eq. (300). Thus the true potential decreases throughout, with

$$\mathcal{F}_{\mathrm{s}}(y^0) - \mathcal{F}_{\mathrm{s}}(y) \geq \frac{l_N}{2\sqrt{N}} \sum_i \sqrt{\kappa_i^0}|y_i - y_i^0| \geq \frac{l_N Z}{4\sqrt{N}}. \tag{301}$$

Coordinates that reach their target are marked, but their coupling slots and reservoir are kept until the end of the sweep. They remain inside the same open face at endpoint distance $\delta_{\mathrm{end}}/2$. No optimizer is computed during these coordinate updates. Frozen functions of untouched coordinates do not change, so one initial eligible list suffices.

All coordinates lie on one common sufficiently fine dyadic grid. Take the largest grid displacement in the chosen direction not exceeding the endpoint or arc limit. If a last grid step cannot use the entire arc, declare saturation only after $Z \geq A_*/2$; a mesh smaller than a fixed inverse-polynomial fraction of $A_*/(k_0 \max r_i)$ ensures this. Unchanged coordinates need no repeated rounding. Eq. (301) then gives the asserted saturated decrease. In the unsaturated case the eligible list is empty and the same error bound, followed by $\kappa_i(y) \geq \kappa_i^0/2$, gives Eq. (299), with ample slack.

A selected coordinate is processed just once unless it is the final partial coordinate of a saturated sweep. Every complete coordinate update reaches a near-endpoint target and is permanently removed at the next deferred batch. Consequently there are at most k_0 complete updates and one partial update per saturated sweep. Each has only a constant number of scalar budget evaluations and a rank-one modeled sum update; charging $\widetilde{O}(N^2)$ per update is sufficient. Alternatively keep only the exact coefficient list inside a sweep and reconstruct its modeled sum at the corrected endpoint by the retained rank-one action. This avoids accumulation in the numerical sum and is already included in the stated scan cost. The value and gradient comparisons at the end use the root correction from Lemma 6.23. A nonempty unsaturated sweep has a complete coordinate displacement of at least $\delta_{\mathrm{end}}/2$, since its base has no near-endpoint coordinates. Its decrease is therefore at least $l_N \delta_{\mathrm{end}} \sqrt{\gamma}/(4\sqrt{N})$ up to the fixed evaluation slack. This and the saturated bound are inverse polynomial. The preceding finer common cache floor certifies the final value decrease. □

Why unsaturated sweeps cannot occur too often. A short unsaturated sweep need not itself pay $1/(P_N N)$. It is important not to charge every such sweep to the larger saturated progress bound.

Lemma 6.25 (Sweep and deferred-batch counts)**.** *With the high/low tests above, the total number of gradient sweeps is $\widetilde{O}(N + I_{\mathrm{flat}} + I_{\mathrm{reg}} + I_{\mathrm{fb}})$. All deferred attenuation batches use expected $\widetilde{O}(N)$ corrected stages. Together with regular endpoint continuation and spread-curvature candidates, all corrected local and face stages number $\widetilde{O}(N)$ in expectation.*

Proof. Saturated sweeps number $\widetilde{O}(N)$ by Eq. (301) and the bounded total phase potential decrease. After an unsaturated sweep, collect all marked coordinates in one nonempty deferred set and use the retained positive-submap attenuation path. No remaining coefficient changes during this path. Write Δ_J for its matrix-potential decrease before the final tiny-slot correction and let $L_* \leq P_N$ be a uniform upper bound on its logarithmic horizon. The preceding forcing estimate gives total energy arc at most

$$P_N \int_0^{L_*} \sqrt{v(s)}\, ds \leq P_N \sqrt{L_* \Delta_J}. \tag{302}$$

If $\Delta_J \leq 1/(P'_N N)$ for a sufficiently large known fixed polylogarithm P'_N, this arc and its normalized transport are small relative to $h_N/\sqrt{N}$. The final tiny-slot deletion is below the same margin by tightening its prescribed cache floor.

Here is the needed gradient comparison along that small path. For a remaining coordinate put $\widetilde{g}_i(s) = g_i(s)/\sqrt{\iota t_i(s) + \gamma}$. In a fixed-factor neighborhood of its starting root, the forcing proof in Lemma 6.23 and the relative scalar bounds give

$$|\widetilde{g}_i'| \leq P_N a(s) + P_N \sqrt{N} |\widetilde{g}_i| a(s), \qquad a(s) = e(\mathsf{P}') + \mathcal{E}(\mathsf{Q}').$$

Until $|\widetilde{g}_i|$ reaches $h_N/(2\sqrt{N})$ the right side is at most $P_N a(s)$. Starting from Eq. (299), Eq. (302) prevents that exit and keeps the metric comparable, by the same first-exit transport argument. Thus

after small attenuation no high gradient test can pass. This comparison uses the fixed remaining coordinates; it does not assume that a normalized weight has zero derivative.

If a new reservoir phase is opened, charge this event separately; there are at most L_0 phases. Otherwise a new sweep following an unsaturated one must be preceded either by attenuation drop at least $1/(P'_N N)$, or by a regular endpoint or curvature event. The large attenuation drops sum to at most the bounded total potential decrease, so there are only $\widetilde{O}(N)$ of them. This proves the stated sweep count, with no independence assumption between adaptive states.

There is at most one deferred batch per corrected accepted state; attenuation changes no remaining coefficient. Therefore the batch count is bounded by the number of sweeps, curvature events and regular endpoints, plus one. Eq. (277) gives total batch stages $\widetilde{O}(b + \sqrt{Nb})$, because their nonnegative drops sum to a constant. Successful spread moves and regular endpoints number $\widetilde{O}(N)$. For the retained capped curvature fallback, $\mathbb{E}I_{\mathrm{fb}} \le (2N)^{-20}/2$. The excess stage count is bounded by $P_N(1+\sqrt{N})I_{\mathrm{fb}}$, using $\sqrt{j} \le j$ for nonnegative integers. Taking expectations gives $\widetilde{O}(N)$ batch and sweep stages. Lemma 6.22 gives $\widetilde{O}(N)$ regular stages, and the initial explicit homotopy uses only $\widetilde{O}(\sqrt{N})$ stages. Both signs of each spread-curvature candidate retain their constant number of corrections. This proves the total claim. □

6.12 A finite norm certificate

Lemma 6.26 (A short-precision norm certificate)**.** *For rational Hermitian E and positive rational a, there is a certified test that accepts only if $\|E\| \le a$ and certainly accepts if $\|E\| \le a/2$. If the entries have a common rational denominator and the entries and threshold have b bits, its working precision is $O(b+\log(2n))$ bits. For logarithmic input length its arithmetic and bit costs are $\widetilde{O}(n^{\omega_0})$.*

Proof. Define $Z := (E/a)^2$ and $Q := I - Z$. The certificate is a computed rational matrix X with a verified enclosure

$$\|X^*QX - I\|_F \le 1/2. \tag{303}$$

Then $X^*QX \succ 0$, so X is nonsingular and $Q \succ 0$ by congruence. This implies $\|E\| < a$, independently of how X was obtained.

To find a certificate, define

$$c_j := 4^{-j}\binom{2j}{j}, \qquad X_s := \sum_{j=0}^{s-1} c_j Z^j, \qquad s := \lceil 10\log_2(2n)\rceil.$$

Compute the powers and sum on a common dyadic grid, and reject if any intermediate exceeds a fixed polynomial magnitude cap. If $\|E\| \le a/2$, then $0 \preceq Z \preceq I/4$ and $Q^{-1/2} = \sum_{j\ge 0} c_j Z^j$, with $0 < c_j \le 1$. The tail has norm at most $4^{1-s}/3$. Consequently $\|X_s Q X_s - I\|_F < 1/8$ for the prescribed s. The powers are contractive and all coefficients have $O(s)$ bits. Stable fast products with a sufficiently fine inverse-polynomial grid therefore approximate X_s and the final congruence with Frobenius error below $1/16$ in each required enclosure. On the good case no magnitude cap is reached, and a verified upper bound below $1/2$ accepts. On every other case the final enclosure still refers to the true rational E and a, including all scaling and product errors; otherwise the test rejects. Thus unsuccessful tests also have fixed caps. All products and enclosures use $O(b+\log(2n))$ bits, and there are $O(\log(2n))$ matrix products. No large exact matrix power is formed. □

6.13 Masked drift and curved moves

All dimensions in this subsection refer to the padded matrix order. We keep the potential, reservoir phases, and notation of Section 6.1. A fixed background of coordinates will be held at its current

coefficients while the other coordinates move. At a corrected state let

$$M = F + \sum_i y_i V_i, \qquad \mathcal{K}(D) = \iota \sum_i \beta_i \operatorname{tr}[V_i D] V_i, \qquad V_i = a_i a_i^* \succeq 0, \qquad \sum_i V_i^2 \preceq I.$$

Zero inputs can be omitted. The true matrices define the mathematical model, as in Section 6.1. Let $\mathsf{P} = (U, V)$ and $\mathsf{Q} = (W, Z)$ be the unique smoothed optimizer and its positive Gibbs state, with $\operatorname{tr}[\mathsf{Q}] = 1$. On bounded-model sublevels,

$$W = U\mathcal{K}(Z)U + \rho U^2, \quad Z = V\mathcal{K}(W)V + \rho V^2, \tag{304}$$

$$U, V \succeq aI, \quad \|U^{-1}\|, \|V^{-1}\|, \|\mathcal{K}(U)\|, \|\mathcal{K}(V)\| \le C, \quad \rho \operatorname{tr}[U^2 + V^2] \le 1. \tag{305}$$

For coordinate i put

$$u_i = \operatorname{tr}[UV_i], \quad v_i = \operatorname{tr}[VV_i], \quad w_i = \operatorname{tr}[WV_i], \quad z_i = \operatorname{tr}[ZV_i], \quad t_i = v_i w_i + u_i z_i,$$
$$p_i = 1 + \iota\beta_i' v_i, \quad q_i = 1 - \iota\beta_i' u_i, \quad M_i = p_i u_i - q_i v_i,$$
$$z_i^{\mathrm{mov}} = p_i^2 u_i w_i + q_i^2 v_i z_i, \qquad \varpi_i = (-\beta_i'')(\iota t_i + \gamma).$$

Here z_i^{mov} is a scalar movement weight, not an optimizer coordinate. The scalar gradient $g_i = \partial_i \mathcal{F}_{\mathrm{s}}$ satisfies

$$M_i g_i = z_i^{\mathrm{mov}} - p_i q_i t_i + \gamma M_i \beta_i'. \tag{306}$$

On light coordinates, and on their fixed-factor expanded neighborhood,

$$p_i, q_i > 0, \quad p_i q_i \ge 1/2, \quad |M_i|, |\beta_i'|, p_i, q_i \le P_n, \quad |M_i| \le P_n(u_i + v_i), \quad \varpi_i \ge (\iota t_i + \gamma)/2.$$

All phases use the existing reservoir allocation $\gamma = \Lambda/(L_0 K)$, where $\Lambda = 10^{-6}$, K is the count at phase start and $L_0 = 1 + \lceil \log_2 \max(2, k_0) \rceil$. The initial reservoir plus all positive phase jumps is at most Λ. No new allocation is made at the switch to the finishing algorithm.

Let $\mathscr{E}$ be the primal Lagrangian Hessian from Eq. (198). Write

$$e(D)^2 = \langle D, \mathscr{E} D \rangle, \qquad \mathcal{E}(B)^2 = \tau_{\mathrm{s}} \langle B, D \log_{\mathsf{Q}}[B] \rangle, \quad \operatorname{tr}[B] = 0, \qquad \mathbb{D} = \operatorname{diag}(\mathscr{E}, B_{\mathsf{Q}}).$$

The joint Jacobian $\mathbb{A}$ has symmetric part $\mathbb{D}$ and $\|\mathbb{D}^{1/2} \mathbb{A}^{-1} \mathbb{D}^{1/2}\| \le 1$.

Masking non-light coordinates costs only a rank defect.

Lemma 6.27 (A small mask)**.** *After removing near-endpoint coordinates, a conservative classification can mark a set B of at most*

$$H_n = \lceil 16\iota^2 n/(\varepsilon_{\mathrm{scale}} \delta_{\mathrm{end}}^2) \rceil$$

coordinates such that every unmarked coordinate is light. The classification costs one full score scan and its true-data enclosures. Marked coordinates are not changed, and their coupling is not removed.

Proof. At this point $1 - |y_i| \ge \delta_{\mathrm{end}}$. Mark an index if a certified upper enclosure of either endpoint score is at least $-\delta_{\mathrm{end}}/8$, using width at most $\delta_{\mathrm{end}}/16$. An unmarked score is strictly negative. A marked index has either $\iota\beta_i u_i \ge \delta_{\mathrm{end}}/2$ or $\iota\beta_i v_i \ge \delta_{\mathrm{end}}/2$. Since $\beta_i \le 1$, it has $u_i \ge \delta_{\mathrm{end}}/(2\iota)$ or $v_i \ge \delta_{\mathrm{end}}/(2\iota)$. The rank-one variance bound gives $\sum_i \operatorname{tr}[MV_i]^2 \le \|M\|_{\mathrm{F}}^2$ for Hermitian M. To see this, the nonnegative Gram matrix $J_{ij} = \operatorname{tr}[V_i V_j]$ obeys $J\nu \le \nu$, $\nu_i = \operatorname{tr}[V_i] > 0$, so its spectral norm is at most one. Its Gram factor and adjoint are contractions. Therefore the number marked is at most $4\iota^2 \delta_{\mathrm{end}}^{-2}(\|U\|_{\mathrm{F}}^2 + \|V\|_{\mathrm{F}}^2) \le 4\iota^2 n/(\varepsilon_{\mathrm{scale}} \delta_{\mathrm{end}}^2)$, leaving the stated slack for enclosures. □

Write J for the unmarked set, $p = |J|$, $q = |B|$, and $\mathcal{K} = \mathcal{K}_J + \mathcal{K}_B$. Write $\mathcal{A}_J(D_U, D_V) := (\mathcal{K}_J(D_V), \mathcal{K}_J(D_U))$, and define $\mathcal{A}_B$ similarly. All subspaces in the next lemma concern the p unmarked physical coordinates.

Lemma 6.28 (Covariance with a fixed background)**.** *Let $\mathcal{H} \subseteq \mathbb{R}^J$ have codimension r. If $r + 2q \le p/4$, there is a finitely supported centered direction in $\mathcal{H}$, extended by zero on B, whose covariance satisfies*

$$\mathbb{E}[\mathcal{Q}(h)] \le \sum_{i \in J} \kappa_i z_i^{\mathrm{mov}} \Sigma_{ii} + r + 2q, \qquad \sum_{i \in J} z_i^{\mathrm{mov}} \Sigma_{ii} \ge p/175. \tag{307}$$

Here $\mathcal{Q}$ is the original full-coupling fixed-constraint inversion energy. This covariance is only used in the proof, not computed by the algorithm.

Proof. We verify why a background coupling is not being silently discarded. The operator $\mathscr{J}_J = \mathrm{diag}(U^{-1}(\cdot)U^{-1}, V^{-1}(\cdot)V^{-1}) - \mathcal{A}_J$ has a positive inverse: its dual inequalities are strict by Eq. (304) and $\mathcal{K}_J \preceq \mathcal{K}$ as positive maps. Let $D_J(h) = \mathscr{J}_J^{-1} F_1 h$ be its fixed-constraint response. Impose the additional real linear equations

$$\mathrm{tr}[V_b (D_J(h))_U] = \mathrm{tr}[V_b (D_J(h))_V] = 0 \quad (b \in B).$$

They add at most $2q$ to the codimension. On this subspace $\mathcal{A}_B D_J(h) = 0$, whence $\mathscr{J} D_J(h) = F_1 h$. The response is thus the response for the full problem.

For clarity the normalized covariance construction remains valid with strict dual inequalities, rather than equality for $\mathcal{K}_J$. Set $s_i = (\beta_i p_i / q_i)^{1/2}$ and $t_i^* = (\beta_i q_i / p_i)^{1/2}$, and define response matrices

$$A_{ij} = \iota s_i s_j |a_i^* U a_j|^2, \qquad B_{ij} = \iota t_i^* t_j^* |a_i^* V a_j|^2.$$

The normalized response vectors $\mathbf{u}, \mathbf{v}$ obey $\mathbf{u} - B\mathbf{v} = \tau h$ and $A\mathbf{u} - \mathbf{v} = \tau h$, where $\tau_i = (p_i q_i / \beta_i)^{1/2}$. The energy matrices have entries $s_i s_j \,\mathrm{Re}\,\mathrm{tr}[W V_i U V_j]$ and $t_i^* t_j^* \,\mathrm{Re}\,\mathrm{tr}[Z V_i V V_j]$. Normalize their positive diagonals $E_{U,i} = s_i^2 u_i w_i$ and $E_{V,i} = (t_i^*)^2 v_i z_i$. The resulting block energy matrix M has trace $2p$ and unit diagonals. For the normalized off-diagonal response map T, the dual inequalities imply

$$M \succeq T^* \Gamma T, \qquad \Gamma_i = (\iota \beta_i u_i v_i)^{-1} \ge \iota > 5.$$

For example, writing $D = \sum_i s_i \mathbf{u}_i V_i$, Cauchy–Schwarz gives

$$\sum_{j \in J} \frac{\iota \beta_j z_j}{u_j} (a_j^* U D U a_j)^2 \le \mathrm{tr}[\mathcal{K}_J(Z) U D U D U] \le \mathrm{tr}[W D U D].$$

This is the required block inequality; the other block is identical. The light inequalities and $\beta_i \ge 1 - y_i^2$ give $\Gamma_i \ge \iota$. In each pair use the unit vector proportional to $(\sqrt{E_{U,i}}, -\sqrt{E_{V,i}})$ and its orthogonal complement. Their energy diagonals remain one, the two response diagonals sum to zero, and the first is nonpositive. The response graph is mapped by $I - T$ onto the span of the first paired vectors. Its dimension is p, since the strict dual inequalities make the response system invertible.

Lemma 6.12 therefore applies to its subspace of dimension at least $p - r - 2q$. Its loss is its codimension, at most $r + 2q$, and its movement is at least $p/175$ when this codimension is at most $p/4$. The physical movement identity is $\tau_i^2 (E_{U,i} + E_{V,i}) = z_i^{\mathrm{mov}}$. On the constrained response space the energy is the full problem's energy, as proved above. This gives Eq. (307). The second fixed-constraint derivative gives the same energy expression even though the background can affect the second response: pairing its defining equation with the full Q uses $\mathscr{J}\mathsf{Q} = \rho E$ and leaves exactly the original energy and mixed terms. No second-response vanishing is assumed. □

A drift-adjusted Hessian at every masked light state. Let $C_{2,i} := (1+\tau/2)\kappa_i$ and $C_\tau := 1+\tau/2$, as in Eq. (265). Its exact scalar certificate and Eq. (307) give

$$\tfrac{1}{2}\operatorname{tr}[\nabla^2 \mathcal{F}_{\mathrm{s}}\Sigma] - \sum_{i\in J} C_{2,i} M_i g_i \Sigma_{ii} \le -\upsilon_n \sum_{i\in J} \varpi_i \Sigma_{ii} + C_\tau(r+2q), \tag{308}$$

where $\upsilon_n^{-1} \le P_n$. This is the retained generator calculation; there is no failed-gradient hypothesis in Eq. (308).

Choose known $C_w \le P_n$ so that on light coordinates

$$p_i q_i t_i + \gamma |M_i \beta_i'| \le C_w \varpi_i.$$

For example the displayed light bounds and $\varpi_i \ge (\iota t_i + \gamma)/2$ give such a coefficient by sums of fixed budget bounds. Set

$$\begin{gathered} a_* := \frac{\upsilon_n}{4(1+C_w)}, \qquad d_i := (C_{2,i}+a_*)M_i, \qquad \varpi_i^\sharp := \varpi_i + z_i^{\mathrm{mov}}, \\ G_\sharp := \operatorname{diag}_J(\varpi_i^\sharp), \qquad \mathcal{H}_\sharp := (\nabla^2 \mathcal{F}_{\mathrm{s}})_{JJ} - 2\operatorname{diag}_J(d_i g_i). \end{gathered} \tag{309}$$

Using Eq. (306) in Eq. (308) proves

$$\tfrac{1}{2}\operatorname{tr}[\mathcal{H}_\sharp \Sigma] \le -a_* \operatorname{tr}[G_\sharp \Sigma] + C_\tau(r+2q). \tag{310}$$

All coefficients in Eq. (309) are evaluated at the base root and frozen during the following path.

Lemma 6.29 (Negative multiplicity without a gradient test)**.** *Put*

$$K_*(n) = \lceil \max\{16H_n, 4000 C_\tau H_n/a_*\} \rceil, \qquad \delta_n = a_*/4, \qquad \pi_n = a_*/(700 C_\tau).$$

For $k > K_*(n)$ the matrix $A_\sharp = G_\sharp^{-1/2}\mathcal{H}_\sharp G_\sharp^{-1/2}$ has at least $\pi_n p$ eigenvalues below $-4\delta_n$. Its spectrum is contained in $[-P_n, P_n p]$, and its diagonal is bounded in absolute value by P_n. Moreover $K_*(n) = nP_n$.

Proof. We have $q \le H_n$, $p \ge k/2$, and $2q \le a_* p/(700 C_\tau)$. Let r count eigenvalues below $-a_*$. If $r \ge p/8$ the conclusion follows. Otherwise $r + 2q \le p/4$, so use Eq. (307) on the preimage of the complementary spectral space. Write $D_\Sigma = \operatorname{tr}[G_\sharp \Sigma] \ge p/175$. On that space the left side of Eq. (310) is at least $-a_* D_\Sigma/2$. Thus

$$r \ge \frac{a_* p}{350 C_\tau} - 2q \ge \pi_n p.$$

This also deals with nonintegral thresholds by the integer eigenvalue count.

Eq. (216) gives $\nabla^2 \mathcal{F}_{\mathrm{s}} \succeq -P_n G$ and $|\partial_i^2 \mathcal{F}_{\mathrm{s}}| \le P_n \varpi_i$ on the light subset. Eq. (306) gives $|M_i g_i| \le z_i^{\mathrm{mov}} + C_w \varpi_i$. The added diagonal in Eq. (309) is therefore bounded by $P_n G_\sharp$. The lower bound and diagonal bound follow, and the trace of $A_\sharp + P_n I$ gives the upper norm bound. Finally all quantities in K_* except its explicit n factor are fixed polylogarithms. □

Computing a direction and its true relative response. The Schur formula in Eq. (233) is restricted to J. Under the $G_\sharp$ normalization its diagonal term changes, by Eq. (309), to a diagonal of norm at most P_n. Set

$$D_\sharp = G_\sharp^{-1/2} F_1^* Q_{\mathbb{Q}} F_1 G_\sharp^{-1/2}.$$

The same completing-square argument as Lemma 6.4 gives

$$a_{\mathrm{dom}} D_\sharp - P_n I \preceq A_\sharp \preceq D_\sharp + P_n I, \quad \operatorname{tr}[D_\sharp] \le P_n p, \quad a_{\mathrm{dom}}^{-1} \le P_n.$$

Indeed the F_2 bound uses $\varpi_i \le \varpi_i^\sharp$, and $\operatorname{tr}[\mathsf{Q} F_{1,i}^2] \le P_n z_i^{\mathrm{mov}}$ since $u_i, v_i \ge a_0 \operatorname{tr}[V_i]$. The fixed-root Schur preconditioner for $\mathscr{J}$ is unchanged; it includes $\mathcal{K}_B$. A sufficiently large known positive shift is therefore comparable to $D_\sharp + RI$. The preceding Gibbs feature identity, finite sampler, alternating row/column sketches, and inverse recurrences apply with $G_\sharp$ and this diagonal shift. They require no zero-gradient condition, only these order bounds. The weights have lower bound $\Lambda/(2L_0K)$, and all surrogate perturbations have the same polynomial conditioning margins.

We record the additional response test and its proof to identify the curved-path gain. For $D = (D_U, D_V)$ put $\widehat{D}_A = A^{-1/2} D_A A^{-1/2}$ and $T_U = U^{-1/2} W U^{-1/2}$, $T_V = V^{-1/2} Z V^{-1/2}$. Define

$$\mathfrak{r}(D, B) = \max_{A=U,V} \{ \|\widehat{D}_A\|, \|T_A^{1/2} \widehat{D}_A T_A^{-1/2}\|, \|\mathsf{Q}_A^{-1/2} B_A \mathsf{Q}_A^{-1/2}\| \}. \tag{311}$$

The middle norm is a singular-value norm of a generally non-Hermitian matrix; it must not be replaced by its spectral radius or omitted. Each of the six maps in Eq. (311) has norm at most $P_n\sqrt{n}$ from joint energy to Frobenius norm, by Eq. (198).

Lemma 6.30 (A verified drift-response direction)**.** *At $k > K_*(n)$ a finite randomized trial, with an absolute conditional success probability, finds a real unit direction r supported on J satisfying, for $\theta = r^\top G_\sharp r$ and $z_r' = (D, B)$,*

$$r^\top \mathcal{H}_\sharp r \le -\delta_n \theta/2, \qquad \|r\|_\infty \le P_n \sqrt{\theta}, \tag{312}$$

$$\mathfrak{r}(z_r') \le P_n \min\{\sqrt{n}, n/\sqrt{p}\} \sqrt{\theta}. \tag{313}$$

All inequalities and the derivative solve are checked against the true model. Lemma 6.37 gives their arithmetic cost and Section 7.12 supplies the finite-precision implementation. No covariance or negative-eigenspace basis is constructed.

Proof. Use Lemma 6.29 and the bounded flat polynomial of the shifted inverse, equal to within a small leakage of one on eigenvalues below $-3\delta_n$, and of zero on those above $-\delta_n$. Call it F; $0 \preceq F \preceq I$. Choose leakage at most $1/(P_n\sqrt{p})$ as well as the quotient error requirement in Lemma 6.15. The Bernstein-polynomial construction and finite inverse evaluation are unchanged.

The response map $\mathcal{R}$ from $G_\sharp^{-1/2} y$ to z_r' obeys $\|\mathcal{R}F\|_{2\to\mathbb{D}} \le P_n$. To verify this point, negative curvature of $A_\sharp$ bounds the *true* Hessian above by $P_n\|y\|^2$ on that band, because the removed drift diagonal has norm at most P_n. In

$$r^\top \nabla^2 \mathcal{F}_{\mathrm{s}} r = -r^\top G r + e(D)^2 + \mathcal{E}(B)^2 + 2\langle F_2 r, D\rangle$$

complete the square, using $r^\top G r \le \theta$ and the retained F_2 bound. This bounds the response energy on the band. The whole-space response norm is at most $P_n\sqrt{p}$; the chosen leakage handles its complement.

For a real-linear map $\mathcal{M} : \mathbb{R}^p \to \mathbb{C}^{n\times n}$ with $\|\mathcal{M}\|_{2\to F} \le a$, its coefficient matrices obey $\sum_i M_i M_i^*, \sum_i M_i^* M_i \preceq 2na^2 I$. This follows by applying the positive coefficient-covariance map to $a^2 I$ in the real matrix-unit and imaginary-unit basis. The Hermitian case has n instead of $2n$. The matrix Rademacher bound hence controls all six relative maps composed with $\mathcal{R}F$ by $P_n n$, with a fixed small failure probability. Meanwhile the flat-filter sign moments give $\|F\xi\|^2 \ge \pi_n p/4$ and $\|F\xi\|_\infty^2 \le 2\log(384p)$ on an event of fixed probability. A union bound, not an independence

assumption between these tests, gives their simultaneous event. Divide by $\|F\xi\|$ and then normalize $G_\sharp^{-1/2}F\xi$ physically. Since $p \geq k/2 > K/4$ in its reservoir phase and $\varpi_i^\sharp \geq \Lambda/(2L_0K)$, this proves the spread bound. The relative response bound is $P_n n/\sqrt{p}$ times $\sqrt{\theta}$; the deterministic energy-to-relative comparison also gives its $\sqrt{n}$ alternative. The modified Rayleigh inequality follows from the same bounded filter.

Compute z_r' through the actual differentiated joint equations and a true residual enclosure. Order-n matrix-function and singular-norm enclosures check Eq. (313); no individual eigenvector gap is required. These costs are included in the invocation below. A fixed smaller choice of all thresholds leaves numerical slack. On any trial, only a direction passing the true enclosures is accepted. □

The curved trial and its predictor. Freeze d_i from Eq. (309) and set it to zero on B. For either sign of s use the quadratic coordinate path

$$y_i(s) = y_i + sr_i - s^2 d_i r_i^2 \quad (i \in J), \qquad y_i(s) = y_i \quad (i \in B). \tag{314}$$

Its second potential derivative at zero is exactly $r^\top \mathcal{H}_\sharp r$. In particular the drift is retained rather than estimated as a small gradient error.

Lemma 6.31 (Curved prediction and progress)**.** *A direction passing Eq. (312)–(313) permits*

$$s_* = \frac{a_n \Gamma(n,p)}{\sqrt{n\theta}}, \qquad \Gamma(n,p) = \max\{1, (p/n)^{1/4}\}, \tag{315}$$

where $a_n > 0$ is a fixed sufficiently small inverse-polylogarithm. Choose a sign with certified nearly nonpositive initial linear derivative. A first-order optimizer predictor followed by a true root correction then gives a decrease at least

$$P_n^{-1} n^{-1} \max\{1, \sqrt{p/n}\}. \tag{316}$$

The complete chosen candidate costs a constant number of root solves and true value/derivative evaluations, up to logarithms. Coordinates in the mask stay fixed. All coordinate-zero crossings are allowed.

Proof. Write $b = \|r\|_\infty$, $z_0' = z_r'$, and $\ell = \mathfrak{r}(z_0') + b$. The proof's reference norms are at the base. For the straight predictor $z_p(s) = z_0 + sz_0'$, we establish

$$\|\mathcal{E}_{\mathrm{KKT},y(s)}(z_p(s))\|_{\mathbb{D}_0^{-1}} \leq P_n s^2 \ell \sqrt{\theta}, \tag{317}$$

$$\|z(s) - z_p(s)\|_{\mathbb{D}_0} \leq P_n s^2 \ell \sqrt{\theta}, \tag{318}$$

$$|(\mathcal{F}_{\mathrm{s}} \circ y)''(s) - (\mathcal{F}_{\mathrm{s}} \circ y)''(0)| \leq P_n \theta (|s|\ell + \sqrt{n} s^2 \ell \sqrt{\theta}), \tag{319}$$

provided $P_n(|s|\ell + \sqrt{n}s^2\ell\sqrt{\theta})$ is a sufficiently small fraction of δ_n and $|s|b$ is a sufficiently small fraction of δ_{end}. We give the extra estimates for the curved path explicitly.

First $|d_i| \leq P_n(u_i + v_i)$ and $|d_i| \leq P_n$. A positive majorant for $-2\sum_i d_i r_i^2 V_i = M''$ has norm at most $P_n b^2$: use $\beta_i^{-1} \leq P_n$ and $\mathcal{K}(U) + \mathcal{K}(V) \preceq CI$. Its weighted trace is at most $P_n\theta$, since

$$|d_i|(w_i + z_i) \leq P_n(u_i + v_i)(w_i + z_i) \leq P_n(t_i + z_i^{\mathrm{mov}}) \leq P_n \varpi_i^\sharp.$$

The middle inequality uses the positive lower bounds for p_i and q_i on the expanded light set. The remaining terms of $\mathcal{C}_{ss}$ are weighted by $\beta_i'' r_i^2 + \beta_i'(-2d_i r_i^2)$ and have the same operator-majorant and

trace bounds. If $-B_0 \preceq B \preceq B_0$, then $B^2 \preceq \|B_0\| B_0$, by writing $B = B_0^{1/2} C B_0^{1/2}$ with $\|C\| \le 1$. The entropy variance inequality therefore bounds their dual energy norm by $P_n b\sqrt{\theta}$. The primal component $\mathcal{A}''\mathsf{Q}$ has the same bound by weighted Cauchy–Schwarz with $|r_i|^4 \le b^2 r_i^2$. Along a short segment $y_i'(s) = r_i(1 - 2s d_i r_i)$ remains within a fixed relative factor, and the same assertions hold with base norms and weights. There is no cubic signed-sum term, since $M''' = 0$. The direct budget/coupling cubic terms are bounded by $P_n b\theta$; terms containing $\beta_i'' y_i' y_i''$ obey this same bound.

For optimizer variation, let R be a normalized Hermitian test matrix. The inverse-Hessian words are $\mathrm{tr}[T(\widehat{D}R^2 + R\widehat{D}R + R^2\widehat{D})]$. The middle word is bounded by $\|\widehat{D}\|\,\mathrm{tr}[TR^2]$; the exterior words are bounded by $\|T^{1/2}\widehat{D}T^{-1/2}\|\,\mathrm{tr}[TR^2]$. Dual variation uses $\|\mathsf{Q}^{-1/2}\Delta\mathsf{Q}\mathsf{Q}^{-1/2}\|$. For the logarithm block, the resolvent formula gives the relative bound

$$|D\langle B, D\log_{\mathsf{Q}}[B]\rangle[\Delta\mathsf{Q}]| \le 2\|\mathsf{Q}^{-1/2}\Delta\mathsf{Q}\mathsf{Q}^{-1/2}\|\langle B, D\log_{\mathsf{Q}}[B]\rangle.$$

Signed coefficient variations bounded by t times the original positive coupling have weighted output norm at most $Cte(D)$, by splitting the two sign sets and applying complete-positive Schwarz. The explicit coefficient changes here have relative size $P_n|s|b$. These identities prove normalized Jacobian variation at the predictor at most $P_n|s|\ell$; they retain matrix order and apply to Hermitian data. The first residual derivative at zero vanishes. The pure second parameter forcing just bounded and the mixed and optimizer-only terms are at most $P_n\ell\sqrt{\theta}$. Integration proves Eq. (317).

Set E_s to twice its right side. On the energy ball of radius E_s around the predictor, additional relative variation is at most $P_n\sqrt{n}E_s$. The chord map using the exact base joint inverse is contractive and maps this ball to itself when the stated bound is small. All four matrix blocks remain positive, with $\mathrm{tr}[\mathsf{Q}] = 1$. This proves Eq. (318); primal convexity identifies the root with the original optimizer. The auxiliary normalized square roots are not computed by the algorithm.

For Eq. (319), write the parameter forcing in the joint residual as $f_s = \partial_s \mathcal{E}_{\mathrm{KKT},y(s)}$. The root response is $\dot{z}(s) = -\mathbb{A}_s^{-1} f_s$. The force variation is bounded by $P_n|s|b\sqrt{\theta}$ plus $P_n b\|z(s) - z_0\|_{\mathbb{D}_0}$; the acceleration terms satisfy the same bound by the displayed M'' majorant. Thus

$$\|\dot{z}(s) - z_0'\|_{\mathbb{D}_0} \le P_n(|s|\ell + \sqrt{n}E_s)\sqrt{\theta}.$$

The direct second-parameter term includes the reservoir acceleration. Its change is bounded by

$$P_n\theta(|s|\ell + \sqrt{n}E_s).$$

Indeed its root differential is paired with the positive majorants above, while its explicit differential has the stated $b\theta$ bound. The energy, metric, and mixed forcing terms obey the same estimate by Cauchy–Schwarz. Substituting in the second-order envelope identity proves Eq. (319). This argument includes the gradient times coordinate-acceleration term; it has not been discarded.

By Eq. (313), $\ell \le P_n\sqrt{\theta}\min\{\sqrt{n}, n/\sqrt{p}\}$. Substitution of Eq. (315) shows $|s_*|\ell \le P_n a_n$ and $\sqrt{n}s_*^2\ell\sqrt{\theta} \le P_n a_n^2$ in both ranges of p. Also $s_* b \le P_n a_n n^{-1/4}$ since $p \le n^2$. The quadratic displacement in Eq. (314) is at most $P_n(s_* b)^2$. Choosing a_n smaller guarantees endpoint distance at least $\delta_{\mathrm{end}}/2$ and all required expanded light bounds. These are comparisons for the originally unmasked set, not a claim that the whole state becomes light.

Compute $d_0 = \sum_i g_i r_i$ with an enclosure of width at most $\delta_n s_*\theta/100$. Choose the sign so its initial derivative is at most that width. On the chosen side, Eq. (319) keeps the second derivative at most $-\delta_n\theta/4$. Integration yields decrease $\delta_n s_*^2\theta/16$, after reserving fixed fractions for all errors. This proves Eq. (316). Only the chosen side is needed. Its potential is bounded by the old value plus the reserved arbitrarily small initial slack; boundedness is not inferred from its Euclidean length.

The base inverse is applied using the true Schur solve and fixed-anchor refinement. Predictor, derivative, matrix-function, coordinate-grid, and value errors are below fixed fractions of the preceding radii and gains. All reciprocal margins and step lengths are polynomial, with logarithmic numerical word lengths by Section 7.12. At zeros the budget is C^2 and its one-sided third derivatives are bounded. The force estimates integrate through those points; no third-order jet crosses a discontinuity. □

6.14 Resident coordinates and delayed deletion

A resident coordinate still contributes its signed coefficient, coupling slot, and scalar reservoir to the potential. Let k count these coordinates, and let D be the pending subset whose coefficients have reached the near-endpoint region. Keep each pending coefficient fixed at its recorded value. Entering D changes neither the potential nor its optimizer. The set $\mathcal{D}$ of coordinates already removed for final rounding is separate from D.

With the constants of Section 6.13, define

$$\zeta_n := \min\{1/100, a_*/(10000C_\tau)\}, \qquad K_c := \lceil \max\{32, n/\zeta_n, H_n/\zeta_n\} \rceil. \tag{320}$$

Thus $\zeta_n^{-1} \le P_n$ and $K_c = nP_n$, with a known fixed polylogarithmic factor. Reservoir phases count all resident coordinates.

Lemma 6.32 (A resident mask)**.** *Suppose $k > K_c$ and $|D| < \lceil \zeta_n k \rceil$. Place every newly near-endpoint coordinate into D. If this trigger is still false, let B contain D and the non-light coordinates marked by Lemma 6.27, and let J be its complement. Write $q := |B|$ and $p := |J|$. Then*

$$q \le 2\zeta_n k, \qquad p \ge (1 - 2\zeta_n)k, \qquad 2q \le a_* p/(700C_\tau). \tag{321}$$

The masked direction and curved-move lemmas apply with this background, giving decrease at least $P_n^{-1}\sqrt{p}/n^{3/2}$.

Proof. Integrality gives $|D| < \zeta_n k$ when the trigger fails. Every movable coordinate has endpoint distance greater than δ_{end}, so Lemma 6.27 marks at most $H_n \le \zeta_n k$ others. This proves the first two bounds. The third follows from $4\zeta_n k \le a_* k/(2500C_\tau)$ and $p \ge 0.98k$.

Lemma 6.28 permits any fixed background with rank loss $2q$. If the negative eigenvalue count is at least $p/8$, the required multiplicity already holds. Otherwise its proof and Eq. (310) give

$$r \ge a_* p/(350C_\tau) - 2q \ge a_* p/(700C_\tau).$$

The movement weight is at least $p/175$ and $r + 2q \le p/4$. Thus Lemmas 6.30 and 6.31 apply, with smaller fixed numerical margins if necessary. Their derivative bounds are used only on J. Pending coordinates have zero parameter derivatives, even if their endpoint distance is smaller than the cutoff. The full background coupling remains in every root and response solve. □

After the initial corrected root, process its near-endpoint coordinates once by positive-submap attenuation. Subsequently, collect pending coordinates until $|D| \ge \lceil \zeta_n k \rceil$, then attenuate their combined coupling and delete them. Keep the old reservoir phase fixed through final deletion and root correction. Update the count phase only afterward. When $k \le K_c$, clear the pending set once and switch permanently to the finishing algorithm of Section 6.11, without allocating a new reservoir. Every newly pending coordinate after a curved trial has endpoint distance at least $\delta_{\mathrm{end}}/2$; the initial scan separately handles exact endpoints.

Lemma 6.33 (Bulk continuation and move counts)**.** *Apart from failure-induced work, the large-count regime has at most*

$$b \le 3 + \zeta_n^{-1} \log \max\{2, k_0\}$$

attenuation calls, including initial and final cleanup. All these calls and the initial data homotopy use $\widetilde{O}(\sqrt{n})$ corrected stages. In a phase $K/2 < k \le K$, successful curved moves number $\widetilde{O}(n^{3/2}/\sqrt{K})$.

Proof. Each triggered deletion removes at least a ζ_n fraction of the resident count. Count these multiplicative drops and add the two untriggered cleanups. For a pending batch use $\mathcal{K}_s = \mathcal{K}_{\mathrm{rest}} + e^{-s}\mathcal{K}_D$. The proof of Lemma 6.3 and the batched argument in Section 6.9 use positivity of this submap and complete-positive Schwarz. They incur no factor $|D|$. A batch with matrix-potential decrease Δ_j uses $\widetilde{O}(1 + \sqrt{n\Delta_j})$ stages, including its final tiny coefficient, exact deletion, and root correction. All decreases sum to a constant by Eq. (269). Therefore

$$\sum_j (1 + \sqrt{n\Delta_j}) \le b + \sqrt{nb \sum_j \Delta_j} = \widetilde{O}(\sqrt{n}).$$

The initial homotopy has the same stage bound. In a count phase $p \ge 0.98k > K/3$; summing the decrease from Lemma 6.32 proves the move count. Pending status itself adds no potential increase. □

6.15 Matrix functions without diagonalization

We replace the cubic spectral evaluations in the preceding implementation. All matrices in this subsection have known polynomial norm and inverse norm bounds, and the requested absolute Frobenius error is inverse polynomial. The number of products below is polylogarithmic. Section 7.12 proves their logarithmic working precision.

Lemma 6.34 (Logarithm and derivative actions)**.** *For $\ell I \preceq A = A^* \preceq UI$, with $\ell > 0$, both $\log A$ and $D \log_A[E]$ can be approximated to the prescribed error in $\widetilde{O}(n^{\omega_0})$ arithmetic operations.*

Proof. Use the ordered resolvent identities

$$\log A = \int_0^\infty ((1+t)^{-1} I - (A + tI)^{-1}) dt,$$
$$D \log_A[E] = \int_0^\infty (A + tI)^{-1} E (A + tI)^{-1} dt.$$

The first follows from scalar functional calculus; differentiation under the integral gives the second. For $M \ge \max\{1, \|E\|_F\}$, the omitted tails outside $[a_0, b_0]$ are bounded respectively by

$$\sqrt{n}(a_0(1 + \ell^{-1}) + (U+1)/b_0), \qquad a_0 M/\ell^2 + M/b_0.$$

Choose positive dyadic endpoints to make these smaller than one eighth of the target. Split the interval into dyadic bands $[a, b]$, $b \le 2a$. Put $c := (a+b)/2$, $h := (b-a)/2$, $R := (A + cI)^{-1}$, and $W := hR$. Then $\|W\| \le 1/3$, and for $t = c + hv$, $|v| \le 1$,

$$(A + tI)^{-1} = \sum_{j=0}^{d} (-v)^j W^j R + \mathcal{R}_d(v), \qquad \|\mathcal{R}_d(v)\| \le \tfrac{3}{2}\|R\| 3^{-d-1}.$$

Integrate monomials using $\mu_j := \int_{-1}^{1} v^j, dv$, equal to zero for odd j and $2/(j+1)$ for even j. The logarithm uses $h\sum_j (-1)^j \mu_j W^j R$ and the analogous scalar series at $1+c$. The derivative uses

$$h \sum_{i,j=0}^{d} (-1)^{i+j} \mu_{i+j} W^i R E R W^j.$$

No factor crosses E. The band errors are bounded by a constant times $h(\|R\| + (1+c)^{-1})3^{-d}$ and $hM\|R\|^2 3^{-d}$, respectively, with the Frobenius conversion for the first. The known bounds determine a logarithmic d and a logarithmic number of bands. Compute each shifted inverse by residual-certified Newton–Schulz iteration, and symmetrize the outputs. All operations are order-n products or inverses. □

Lemma 6.35 (Exponentials, powers, and Gibbs actions)**.** *For Hermitian L with $\|L\| = O(\log n)$, e^L has an $\widetilde{O}(n^{\omega_0})$ arithmetic implementation. This also implements A^t, $|t| \le 1$, and the operators Q_{Q} and B_{Q} at a root with polynomial spectral bounds. For s vectors and possibly different $t_j \in [0,1]$, all $A^{t_j} g_j$ can be computed by polylogarithmically many products of shape (n,n,s) after one matrix logarithm.*

Proof. The exponential series truncated at degree d has norm error at most $e^M M^{d+1}/(d+1)!$ for $M \ge \|L\|$. A sufficiently large constant times $M + \log(n/\varepsilon) + 1$ suffices for Frobenius error ε. Evaluate the normalized terms $L^j/j!$. For powers use $A^t = \exp(t \log A)$ and Lemma 6.34. For the Gibbs covariance put $L := \log \mathsf{Q}$ on the two blocks. The ordered expansion is

$$\int_0^1 e^{tL} E e^{(1-t)L}, dt \simeq \sum_{i,j=0}^{d} \frac{L^i E L^j}{(i+j+1)!}.$$

Indeed the beta integral cancels the two factorials. Evaluate with normalized powers and bounded beta-integral coefficients. The two exponential tails control the error. Subtract $\mathrm{tr}[\mathsf{Q}E]\mathsf{Q}$ and divide by τ_{s} for Q_{Q}; use $\tau_{\mathrm{s}} \Pi_{\mathcal{T}} D \log_{\mathsf{Q}}$ for B_{Q}.

For the batched assertion set $G := [g_1 \ \cdots \ g_s]$ and use

$$Y_0 := G, \qquad Y_j := L Y_{j-1} \operatorname{diag}(t_1/j, \ldots, t_s/j).$$

The sum of these arrays gives the different exponential actions. The columnwise tail bound, summed in Frobenius norm, fixes their common accuracy. For Gibbs features this prepares $\mathsf{Q}^{t_j/2} g_j$ and $\mathsf{Q}^{(1-t_j)/2} h_j$ without an eigenvector matrix. Scalar work and writing vectors cost $O(ns)$ per term. □

Lemma 6.36 (Dyadic-shift Lyapunov solves)**.** *If $\ell I \preceq T \preceq U I$, the equation $TX + XT = B$ can be solved to the prescribed Frobenius accuracy in $\widetilde{O}(n^{\omega_0})$ arithmetic operations, without an eigenvalue-separation assumption.*

Proof. For a positive dyadic shift p, put $R_p := (T + pI)^{-1}$ and $A_p := (T - pI)R_p$. Starting from zero, update

$$X^+ := A_p X A_p + 2p R_p B R_p. \tag{322}$$

The exact solution X_* satisfies

$$(T+pI)X_*(T+pI) - (T-pI)X_*(T-pI) = 2pB,$$

so $X^+ - X_* = A_p(X - X_*)A_p$. Cycle through shifts from $2^{\lfloor \log_2 \ell \rfloor}$ to $2^{\lceil \log_2 U \rceil}$. Every eigenvalue t has one shift $p \le t < 2p$, giving $|(t-p)/(t+p)| \le 1/3$; all other factors are at most one. The

A_p commute because they are functions of T. Thus one cycle contracts error by at most $1/9$, and $\|X_*\|_F \le \|B\|_F/(2\ell)$ supplies the initial error bound. A logarithmic number of cycles suffices. A final residual evaluation, together with the operator lower bound 2ℓ, certifies the solution. □

Lemma 6.37 (Count-sensitive root and feature actions)**.** *For $n \le k \le n^2$, put $u := \log_n k$ and use ϑ from Eq. (250). Every true residual, corrected root stage, response solve, full resident score scan, and relative-response norm certificate costs*

$$\widetilde{O}(n^{\vartheta(u)}). \tag{323}$$

For $k \le n$, use $u := 1$. Implicit feature actions with at most $\widetilde{O}(n^u)$ selected coordinates and features have the same bound, including initial sample preparation.

Proof. The coupling and its adjoint, signed sums, and score scans use products of shapes (n, n, k) and (n, k, n) and $O(nk)$ scalar work. Fix a common power-of-two block base, with even exponent, for the strict square, $(1, 3/2, 1)$ and $(1, 2, 1)$ bounds. At $O(\log n)$ recursive levels, mix the two relevant endpoint algorithms; choose the least number of wider levels covering k. Padding costs a constant factor. The product costs interpolate to $\vartheta(u)$, uniformly in the integer sizes. Tensor permutations give the adjoints.

The root Schur system in Eq. (203) is polylogarithmically conditioned relative to $\mathscr{E}$. Its inverse blocks reduce to

$$\widehat{D}_U T_U + T_U \widehat{D}_U = U^{1/2} F_U U^{1/2}, \qquad T_U = U^{-1/2} W U^{-1/2},$$

and the analogous V equation. Lemmas 6.34–6.36 implement these and the Gibbs actions in $\widetilde{O}(n^{\omega_0})$. The root bounds in Eq. (197) give all spectral floors and ceilings. Anchored correction in Lemma 6.2 uses the actual residual and the actual $D \log_{\mathsf{Q}}$ block off the root. It does not presume a current Gibbs identity at an approximate iterate. The same correction count therefore has the cost in Eq. (323).

Lemma 6.35 prepares the sampled power vectors. The product shapes are those of Lemma 6.9. Its rank-two feature and quadratic-form formulas then give restricted actions and adjoints, including repeated indices. A terminal factor of order $\widetilde{O}(q)$ and its Gram use square products when the last two sketch sizes equal $q \ge n$. The explicit ridge supplies its inverse floor. Every action refers to one frozen covariance, Hessian surrogate, and set of feature factors. Relative-response certificates, including the non-Hermitian similarities in Eq. (311), use the short norm test in Section 7.12. No separate cubic spectral step remains. □

6.16 Count-sensitive recursive inverse costs

We use the alternating ridge systems in Eqs. (242)–(247), restricted to the movable coordinates and normalized by $G_\sharp$. All coupling actions still include the full resident background. Write the padded order as $n = 2^r$, put $L := 2 + r$, and set

$$\kappa := 2^{\lceil \log_2 k \rceil} = 2^t, \qquad u := t/r, \qquad r \le t \le 2r.$$

The case $n = 1$ is handled directly. The large-count regime has $k > n$. For $n \le q \le \kappa$ define

$$\mathcal{M}(n, q; \kappa) := n^{\vartheta(u)} (q/\kappa)^{5/6}. \tag{324}$$

Both slopes of ϑ exceed $5/6$, so Lemma 6.37 bounds a restricted action by this quantity, up to a fixed logarithmic factor. Enlarged lists can be split into polylogarithmically many blocks.

Lemma 6.38 (Expanded recursive cost)**.** *Let $2 \le h \le L^2$ and let $n \le q_h \le \cdots \le q_1 \le \kappa$ be powers of two with $q_{h-1} = q_h$. There is a fixed integer $D \ge 1$ such that a complete capped drift-response trial costs at most*

$$\begin{aligned} &L^{4D}(n^{\omega_0} + \mathcal{M}(n, q_1; \kappa) + q_h^{\omega_0} + \sqrt{1 + \kappa/q_1}\, n^{\vartheta(u)}) \\ &+ \sum_{j=1}^{h-1} L^{D(j+4)} \sqrt{(\kappa + q_1)/q_{j+1}}\, \mathcal{M}(n, q_{j-1}; \kappa) \\ &+ L^{D(h+4)} \sqrt{(\kappa + q_1)/q_h}\, q_h^2, \qquad q_0 := q_1. \end{aligned} \tag{325}$$

The constant D is independent of all sizes and depth. The trial retains an absolute conditional success probability and verifies the true direction and response inequalities on every accepted sample.

Proof. The invariant in Eq. (248), now with the padded movable count in place of $\bar{k}$, remains at most $2q_j/d_0$. All ridges, factors, and inverse bounds are polynomial uniformly in depth. The sampling proofs in Lemmas 6.6 and 6.8 require polynomial sizes and positive ridges; they do not require the earlier terminal floor $n^{4/3}$. Take moment order $\lceil 128 \log_2(2kn\ell_{\mathsf{Q}}(h+1)) \rceil = O(L)$ and allocate each conditional index failure at most $1/(1000(h+1))$. The same union bound gives constant joint success probability. Lemma 6.30 adds its six norm tests with their reserved slack. Finite caps bound work on all other samples.

Use the normalized recurrence in Eq. (256), with the residual corrections in Lemma 7.12. Lemma 7.19 extends its common accuracy interface to these sizes. Choose D large enough to include the fixed per-level logarithmic overhead, including residual correction. One internal level costs at most $L^D \sqrt{q_j/q_{j+1}}$ times its restricted action plus one child solve. The initial K_0 solve has constant condition number; the shifted physical solve contributes $L^D \sqrt{1 + \kappa/q_1}$. Expanding gives the sum and terminal inverse-action term in Eq. (325). The last two equal sizes make terminal construction, Gram formation, and inversion cost $L^D q_h^{\omega_0}$; the inverse application costs $L^D q_h^2$.

Initial sample preparation costs $L^D(n^{\omega_0} + \mathcal{M}(n, q_1; \kappa))$ by Lemma 6.35. Full resident actions cost $\widetilde{O}(n^{\vartheta(u)})$. Index lists, random bits, all response and quotient checks, and the chosen curved candidate and correction fit the displayed terms. No child inverse or terminal setup is omitted. □

Fix this D, set $C_* := 2000(D+1)$, and define

$$b_{\max} := t - r, \qquad b_0 := \min\{b_{\max}, C_* \lceil \log_2 L \rceil\}.$$

If $b_0 = b_{\max}$, use the two deficits (b_0, b_0). Otherwise set $b_{i+1} := \min\{b_{\max}, \lceil 3b_i/2 \rceil\}$ until $b_T = b_{\max}$, and use

$$(\ell_1, \dots, \ell_h) := (b_0, b_1, b_1, \dots, b_T, b_T), \quad q_j := 2^{t-\ell_j}, \quad \ell_0 := \ell_1. \tag{326}$$

This integer schedule has $h = O(\log L)$, $h \le L^2$, $q_{h-1} = q_h = n$, and $2^{b_0} \le L^{2C_*}$.

Lemma 6.39 (Uniform count-sensitive trial bound)**.** *The schedule in Eq. (326) makes the complete cost in Eq. (325) equal to $\widetilde{O}(n^{\vartheta(u)})$, with a fixed logarithmic degree.*

Proof. The top term is at most $2L^{4D+C_*} n^{\vartheta(u)}$; setup obeys the same bound because $q_h = n$. An internal summand is at most

$$2n^{\vartheta(u)} L^{D(j+4)} 2^{-5\ell_{j-1}/6 + \ell_{j+1}/2} \le 4n^{\vartheta(u)} L^{D(j+4)} 2^{-\ell_{j-1}/12},$$

using $\ell_{j+1} \le 3\ell_{j-1}/2 + 1$. When $b_0 < b_{\max}$, put $v := \lfloor (j-1)/2 \rfloor$. Then $\ell_{j-1} \ge C_*(3/2)^v \log_2 L$ before the terminal truncation, while $j + 4 \le 8(3/2)^v$ and $j \le 4(3/2)^v$. Consequently

$$L^{D(j+4)} 2^{-\ell_{j-1}/12} \le L^{-2j}.$$

All internal terms sum to $O(n^{\vartheta(u)})$.

The leaf overhead needs a separate estimate. Each preterminal increase is at least $b_0/2$, so $h+4 \le 12b_{\max}/b_0$ and

$$D(h+4)\log_2 L \le (12D/C_*)b_{\max} < r/100.$$

The leaf inverse applications cost at most $2\sqrt{\kappa}n^{3/2+1/100}$. They fit because $\vartheta(u)-u/2 \ge \omega_0 - 1/2 = 1.871177 > 1.51$. If $b_0 = b_{\max}$, the depth is two, all logarithmic powers are fixed, and the sole internal term has the additional factor $2^{-b_0/3}$. The same comparison handles the leaf. Thus no depth-dependent logarithmic power is treated as a fixed polylogarithm without a bound. □

7 Finite-precision implementation

Sections 7.1–7.3 fix the precision requirements and acceptance margins. Sections 7.4–7.9 implement matrix operations, direction selection, sampling, and nested solves at these accuracies. Sections 7.10–7.12 account for endpoint continuation, short-precision sign selection, and resident batches.

The numerical routines used in the upper-bound proofs require certified residual, value, and curvature estimates. We collect their finite-precision implementations from Song–Yue [SY26a], retaining the hypotheses and notation of each routine. Lemma 3.8 specifies the trace-map interface used by the higher-rank algorithms.

7.1 Computational model and accuracy requirements

Let L_{in} be the total binary input length. Throughout this section, a machine word has $w = \lceil \log_2(2n) \rceil$ bits, up to an absolute constant factor for signs and carry registers. We charge bit operations for reading the original rational inputs and for all subsequent integer arithmetic. A numerical encoding of B bits is stored in $O(1 + B/w)$ machine words. This convention applies to numerators, denominators, dyadic mantissas and exponents, accumulators, random-bit strings, indices, counters, and addresses. In particular, indices into the original m inputs may occupy several words; no polynomial relation between m and n is assumed.

For completeness, an integer is stored as a signed sequence of base-2^w digits. Addition, comparison, shifts, multiplication, and division are implemented on these digits, with carries propagated between words. A product of two digits and its carry use $O(w)$ bits; longer intermediate results are stored as additional digits. The same representation applies to rational and interval endpoints. We charge all constituent operations, including those on multiword indices and addresses, according to their full bit lengths. Thus an operation on a multiword quantity is not a unit-cost machine operation. Any bit-cost bound below includes this representation and its arithmetic costs.

After the one-time input conversion, the preprocessing grids use $O(\log(2mn))$ bits per numerical entry, with fixed multipliers, hence $O(1 + \log(2mn)/\log(2n))$ machine words per entry. The nonlinear phase has $k = O(n^2)$ active coordinates and uses $O(\log(2kn)) = O(\log(2n))$ accuracy and magnitude bits per numerical entry, hence $O(1)$ machine words per entry. The constants cover the fixed discrepancy and residual allowances. In the lemmas below, "working bits" and precision bounds refer to the total bit length of a numerical encoding, not to the size of a single machine word. Dependence on inherited input lengths or on an explicitly supplied accuracy or failure parameter is retained in these bounds and in the charged bit costs.

The grouped reducers choose barrier signs using the short-precision directional derivative in Lemma 7.17. Each summand and the common-grid accumulator have the bit length proved there

and use the corresponding number of machine words. Original rational inputs are read and converted using the same multiword representation, with their full input length charged through L_{in} before conversion to the numerical grids.

The arithmetic analyses specify residual, quotient, endpoint, and value tolerances before an operation is accepted. We implement those same tests using enclosures smaller than their reserved margins. The potential, covariance inequalities, step sizes, and progress counts are unchanged. Conditioning bounds used in those analyses remain there; the arguments below convert them to working precision.

7.2 Stable partial signing and grouped reduction

The following lemmas implement the corresponding arithmetic reductions. Each uses the local notation of the reduction it implements.

The ridge reduction makes many updates between recomputations. The required stability estimate controls an entire consecutive sequence of updates.

Lemma 7.1 (Stability of a ridge-update epoch)**.** *Consider an epoch in Lemma 5.1, with its fixed column matrix C, ridge δ, dimension d_{vec}, and entry tolerance τ. Put $\mathfrak{b} := 1+\|C\|^2/\delta$. Suppose the initial matrix error is at most ϵ_0, and each update or product that writes pending updates introduces matrix error at most ρ. For a sufficiently small absolute constant $c > 0$, if $\epsilon_0 + d_{\mathrm{vec}}\rho \le c\mathfrak{b}^{-3}$, the error at every epoch prefix is*

$$O(\mathfrak{b}^2(\epsilon_0 + d_{\mathrm{vec}}\rho)).$$

The prescribed entry tolerance and residual allowance are attained with $O(\log(2md_{\mathrm{vec}}/\eta))$ working bits, with a fixed multiplier. The resulting bit cost is the arithmetic bound in Lemma 5.1, up to logarithmic factors.

Proof. Throughout an epoch, $\delta I \preceq G \preceq (\delta + 4(d_{\mathrm{vec}} + 1))I$ and $\|\mathsf{G}\| \le \|C\|^2/\delta = O(d_{\mathrm{vec}}/\delta)$. The insertion denominator in Eq. (133) is at least one. For deletion, write $G = G_- + C_iC_i^\top$, with $G_- \succeq \delta I$. Then

$$1 - \mathsf{G}_{ii} = (1 + C_i^\top G_-^{-1} C_i)^{-1} \ge \frac{\delta}{\delta + 4}.$$

All exact quantities and all divisions therefore have polynomial magnitude bounds in $m, d_{\mathrm{vec}}, 1/\eta$.

For a consecutive sequence of updates let D be the diagonal matrix of its net column changes. Every input is inserted and removed at most once, so $\|D\| \le 1$. The exact composition is

$$F_D(\mathsf{G}) := (I + \mathsf{G}D)^{-1}\mathsf{G}.$$

If G' is the Gram matrix at the end of this sequence, the identity

$$(I + \mathsf{G}D)^{-1} = I - C^\top (G')^{-1} CD$$

shows that the inverse on the left has norm at most $\mathfrak{b} := 1 + \|C\|^2/\delta$. This bound holds for every consecutive subsequence, including one containing both insertions and deletions. For a symmetric perturbation E with $\|E\| \le 1/(2\mathfrak{b})$,

$$\begin{aligned} F_D(\mathsf{G}+E) - F_D(\mathsf{G}) &= (I + \mathsf{G}D)^{-1} E (I + D(\mathsf{G}+E))^{-1}, \\ \|F_D(\mathsf{G}+E) - F_D(\mathsf{G})\| &\le 2\mathfrak{b}^2 \|E\|. \end{aligned} \tag{327}$$

The second inverse is bounded by $2\mathfrak{b}$ by the Neumann series.

Here is the resulting error accounting over a whole epoch. Suppose each update, evaluated at the currently represented matrix, introduces an additive matrix error of norm at most ρ, and each product that writes the pending updates explicitly has the same error bound. Such a product leaves the represented matrix unchanged in exact arithmetic, so it is an identity map with an additive error. Compare a computed prefix with the exact prefix by inserting one error at a time and composing all remaining exact updates. Eq. (327), applied to every such suffix, bounds the prefix error by $O(\mathfrak{b}^2(\epsilon_0 + d_{\mathrm{vec}}\rho))$, where ϵ_0 is the initial error. A bootstrap with this bound smaller than $1/(4\mathfrak{b})$ justifies the same bound for the perturbed suffixes. In particular no factor is raised to the epoch length.

The lazy representation satisfies this additive-error model. Summing at most b_{batch} stored outer products recovers a column with error at most a fixed polynomial in $d_{\mathrm{vec}}, b_{\mathrm{batch}}, \delta^{-1}$ times its scalar rounding unit. The lower bounds on the denominators then give the same type of bound for the error of an update. The separately stored diagonal has at most b_{batch} such rounding errors. Coefficient updates also use only polynomially bounded quantities. Taking $O(\log(2md_{\mathrm{vec}}/\eta))$ bits, with a sufficiently large absolute multiplier, makes all resulting matrix entry errors smaller than τ. The explicit matrix is symmetrized after each product, and its diagonal replaces the separately stored diagonal. The factors in each rectangular product have polynomially bounded entries. A fixed bilinear recursive algorithm implements the strict rectangular exponent bound; its depth is $O(\log d_{\mathrm{vec}})$ and all coefficients at a level are fixed. Hence its intermediate magnitudes and normwise error amplification are polynomial, and only a constant multiple of the same logarithmic precision is needed. The strict slack $2.054999 < 2.055$ permits this fixed algorithm. Each new epoch is recomputed from its retained input columns; errors are not passed through an unbounded sequence of epochs.

The epoch inverses are well conditioned up to a fixed polynomial, so the short-precision Newton–Schulz construction and stable fast products of Section 7.4 apply, with logarithmic factors in md_{vec}/η. They give the stated bit cost as well as the arithmetic cost. □

We next apply the epoch stability bound to the small reduction used at each level of the group tree.

Lemma 7.2 (Finite-precision grouped reduction)**.** *Under the hypotheses of Lemma 5.2, consider its small procedure on at most $2D$ columns with residual allowance ε and inherited coefficient encoding length p_y. The procedure preserves the cube, leaves at most D fractional coordinates, and changes the represented sum by norm at most ε, using $O(p_y+\log(2D/\varepsilon))$ working bits with a fixed multiplier. Its bit cost is $\widetilde{O}(D^{2.5275})$, with logarithmic factors including the input encoding length and $\log(1/\varepsilon)$.*

Proof. Throughout this procedure, $\delta I \preceq G \preceq (\delta+8D)I$, and every deletion denominator is at least $\delta/(\delta+4)$. For any consecutive sequence of deletions, Lemma 7.1 applies with $\mathfrak{b} \le 1+8D/\delta$. It bounds the amplification of an additive matrix error through the entire sequence by $O(\mathfrak{b}^2)$. Writing a pending sum to the base is an identity map with an additive product error. Recovering a column from at most b_{batch} terms and then computing its deletion update introduce errors bounded by a fixed polynomial in D, δ^{-1} times the scalar rounding unit. The separately stored diagonal has the same bound; refresh it from the base whenever a pending sum is written out. The suffix-error accounting in Lemma 7.1 therefore applies to all $O(D)$ deletions and products, including simultaneous hits.

Each rectangular product uses fixed bilinear recursion. Its factors have polynomially bounded entries, and its logarithmic recursion depth gives polynomial error amplification. The strict exponent slack permits a fixed algorithm at exponent $1+2a$. Together with the initial inverse, these bounds make all queried entries accurate to τ using a sufficiently large constant times $\log(2D/\varepsilon)$

guard bits. Boundary ratios use short dyadic coefficients and directions, and inward rounding restores the common grid after every move. The initial encoding length of y_{aux} is included in the logarithmic cost. Thus batching changes neither the residual allowance nor the cube and stopping guarantees of Lemma 5.2. □

The randomized small reducer also needs certified directions from a fixed ridge inverse. Its direction evaluations and the sign selection in Lemma 7.17 both use short numerical entries.

Lemma 7.3 (Finite ridge-direction evaluations)**.** *Use the data and parameters $D, \varepsilon, \nu, T_*, p_C, p_y$ of Lemma 6.18. Its ridge inverse and direction trials attain their prescribed residual, norm, and maximum-entry tolerances using*

$$O(p_C + p_y + \log(2D/\varepsilon) + \log(2T_*))$$

working bits, with a fixed multiplier. Forming the inverse costs $\widetilde{O}(D^{\omega_0})$ bit operations, and each direction trial after fixing the inverse costs $\widetilde{O}(D^2)$ bit operations.

Proof. All magnitudes and inverse bounds here are polynomial in D, ε^{-1} and the displayed logarithmic factors. Newton–Schulz reaches the stated accuracy in logarithmically many products: from $R_0 = J_\delta/U^2$ with $U \geq \|J_\delta\|$, its exact residual squares at each iteration, and its initial spectral gap is at least δ^2/U^2. Rounded residuals satisfy $e_{t+1} \leq e_t^2 + \epsilon_{\mathrm{unit}}$. Taking the product-error floor below a fixed fraction of the squared initial gap and final target proves the same count, with polynomial iterate norms. The fixed stable square recursion gives the required accuracy with $O(p_C + p_y + \log(2D/\varepsilon) + \log(2T_*))$ bits, up to a fixed constant multiplier. Forming or inverting an order-D matrix costs $\widetilde{O}(D^{\omega_0})$ in both models. A direction trial after this inverse is fixed costs $\widetilde{O}(D^2)$, including its exact short-dyadic norm and maximum-entry tests. No singular values, nullspace basis, or exact rational inverse is computed. □

For energy-controlled groups, the sampled preconditioner is held fixed throughout a solve. This permits one error estimate for the normalized recurrence.

Lemma 7.4 (Precision for the structured group solve)**.** *Fix a short rational group matrix $B \in \mathbb{R}^{D\times r}$ and represented factors $Y \in \mathbb{R}^{D\times s}$ and $K = K^{\mathsf{T}} \succeq (T/2)I$, where $T \geq 1$. Put $A := \delta I + BB^{\mathsf{T}}$ and $P := \delta I + YK^{-1}Y^{\mathsf{T}}$, with $\delta > 0$. Suppose a known interval $[a_{\mathrm{spec}}, \beta]$ contains the spectrum of $P^{-1}A$. Let $L \geq 2$ bound all dimensions, the input magnitudes, $T, \delta, \delta^{-1}, {a_{\mathrm{spec}}}^{-1}, \beta$, the right-hand-side norm, and the prescribed iteration count. Choose the iteration count so that the exact normalized Chebyshev recurrence meets half the requested residual and direction tolerances. If the represented entries have at most b_{data} bits, its finite-precision implementation attains the full tolerances $0 < \varepsilon < 1$ with $O(b_{\mathrm{data}} + \log(L/\varepsilon))$ working bits. The structured actions and the solve have bit cost equal to their arithmetic bound up to logarithmic factors. This applies on the good sampling event of Lemma 10.1; unsuccessful samples use the same finite work and magnitude caps.*

Proof. Freeze B, Y, K throughout the solve. They define one positive system and one positive preconditioner. The Woodbury formula is $P^{-1} = \delta^{-1}I - \delta^{-2}Y(K + \delta^{-1}Y^{\mathsf{T}}Y)^{-1}Y^{\mathsf{T}}$. Its small matrix has smallest eigenvalue at least $T/2$ and norm bounded by a fixed polynomial in L. All factors and inverse factors in this formula have polynomial bounds in L. Newton–Schulz and fixed recursive products therefore attain the required absolute accuracy using $O(b_{\mathrm{data}} + \log(L/\varepsilon))$ bits. Cancellation is covered by a finer absolute error in each term. Moreover $P^{-1}A$ is similar to the symmetric positive matrix $P^{-1/2}AP^{-1/2}$ with polynomial similarity condition number. The square roots in this observation are used only for analysis.

Use the normalized recurrence of Eq. (256). Enlarge β to $\max\{\beta, 2a_{\mathrm{spec}}\}$ if necessary, so the interval width also has an inverse-polynomial lower bound. In this recurrence, the scalar ratios lie in $[0,1]$ and their update is Lipschitz with constant at most one. An error inserted at step i propagates through

$$[T_i(\varrho)/T_J(\varrho)]U_{J-i}(D_0),$$

where D_0 is the scaled preconditioned operator. On the good interval its norm is at most a polynomial similarity factor times $J-i+1$. Vector and scalar errors therefore accumulate with a polynomial bound in the prescribed iteration count, not an exponential product of condition factors. All iterates and intermediate right-hand sides have polynomial magnitudes on that event. A sufficiently large constant times $b_{\mathrm{data}} + \log(L/\varepsilon)$ precision meets the true residual and direction targets. In Lemma 10.2, the ridge, sample sizes, iteration cap, and tolerances are polynomial in the displayed size and inverse-accuracy parameters, so this is the claimed logarithmic precision. Compute the final residual from the same fixed B and right-hand side, including its rounding enclosure. The structured actions and residual use the same rational vectorization and group sums. The separate input-cache error is included when transferring this residual to the original matrices. □

7.3 Numerical caches and acceptance margins

The exact retained instance and its numerical cache have different roles. The next lemma controls cache error against the true model at each accepted state.

Lemma 7.5 (Short caches and certified acceptance)**.** *Consider a retained instance with $k \le n^2$ active matrices, the symbolic fixed contribution, and the conditioning and acceptance margins specified in Section 5. A one-time PSD cache with Frobenius error at most $u = a(kn)^{-C_0}$ per retained matrix, for a sufficiently large fixed C_0, permits the KKT residuals, predictors, gradient and curvature tests, endpoint scores, and value comparisons to be evaluated within their reserved errors using $O(\log(2kn))$ accuracy and magnitude bits. Each accepted correction restores the prescribed root-error floor. The initial coordinate rounding uses only the reserved $\varepsilon_{\mathrm{round}}/2$ allowance, and subsequent solves use the short cache. The reference rank-one routines use rank-one caches. For normalized higher-rank inputs, the same precision conclusion holds for the trace-map interface of Lemma 3.8, using the PSD caches below and the dense action costs of the higher-rank algorithms.*

Proof. **One-time data approximation.**

After partial signing, the algorithm needs only its k active matrices. For the reference rank-one routines, store a short rational PSD rank-one approximation to each C_i, with Frobenius error at most u, where $u := a(kn)^{-C_0}$ and C_0 is a sufficiently large fixed constant.

Such approximations can be obtained without assuming a lower bound on a nonzero eigenvalue. Discard a matrix whose largest diagonal entry is below $u/(Cn)$; its Frobenius norm equals its trace and is $O(u)$. Otherwise choose a largest diagonal entry $a := (C_i)_{jj} \ge u/(Cn)$ and write $C_i = a\mathsf{c}\mathsf{c}^*$, where $\mathsf{c} := C_i e_j/a$. Here every $|\mathsf{c}_i| \le 1$ by the PSD minor inequality and maximality of a. Approximate a by a positive rational and c by dyadics to enough additional $O(\log(2kn))$ bits. The product remains PSD and rank at most one and has the required error. This is a numerical cache; the mathematical model continues to use the true C_i.

For arbitrary-rank retained inputs $A_i \succeq 0$ with $\sum_i \mathrm{tr}[A_i]A_i \preceq I$, we have $\|A_i\|_F \le 1$. Choose a dyadic mesh v within a factor two of $u/(8n^2)$. Round the diagonal entries and the real and imaginary parts above the diagonal to this grid and reflect conjugates, obtaining a Hermitian R_i. Then $\|R_i - A_i\|_F \le 2nv$. Define

$$\widehat{A}_i := R_i + 2nvI \succeq 0, \qquad \|\widehat{A}_i - A_i\|_F \le 2nv(1+\sqrt{n}) \le u. \tag{328}$$

The positivity follows from $A_i \succeq 0$ and $\|R_i - A_i\| \le 2nv$. This construction uses $O(n^2)$ rounded entries per matrix and $O(\log(2n/u))$ bits per entry. It requires no lower bound on a nonzero input eigenvalue and no rank decision. It leaves each original matrix associated with one signing coordinate.

Here is the perturbation bound for the higher-rank coupling. Write $\Delta_i = \widehat{A}_i - A_i$ and suppose $|\gamma_i| \le G$. Since $\|\widehat{A}_i\|_F \le 2$, the maps on Hermitian matrices obey

$$\|\widehat{\mathcal{K}}_\gamma - \mathcal{K}_\gamma\|_{F\to F} \le \sum_i |\gamma_i| \|\Delta_i\|_F (\|\widehat{A}_i\|_F + \|A_i\|_F) \le 3kGu. \tag{329}$$

Indeed subtract the two rank-one operators $X \mapsto \operatorname{tr}[A_i X]A_i$ and apply Frobenius Cauchy–Schwarz to their two difference terms. This estimate also applies to signed coefficient derivatives, with their established polynomial bounds in place of G. A signed sum with coefficients of absolute value at most two has cache error at most $2ku$. In addition, $\|\sum_i \operatorname{tr}[\widehat{A}_i]\widehat{A}_i - \sum_i \operatorname{tr}[A_i]A_i\| \le 3k\sqrt{n}u$. Thus a sufficiently small fixed inverse-polynomial u preserves all reserved coupling, derivative, and normalization margins. The true matrices still define the mathematical model; exact normalization of the numerical cache is not required.

Maintain F symbolically as a combination of the initial k columns. Each coefficient has absolute value at most two: it is $-\widetilde{y}_i^0$ plus the cached removal coefficient, if any. All coefficients use the same fine dyadic grid. Errors in F, D_{pow}, the coupling, and its first two derivatives can consequently be bounded afresh by a fixed polynomial in $k, n, \delta_{\mathrm{end}}^{-1}$ times u. They do not accumulate with the number of local steps. No long original rational coefficient is used in subsequent solves.

For initialization, the cap and feasibility-repair argument in [SY26a, Section 4.10] controls coefficient errors before recovering the true optimizer. For local tracking, short-cache evaluations are inexact Newton steps for the true KKT system. The forcing for a predictor is evaluated and solved to a sufficiently small inverse-polynomial error, so its product with the mesh length is below a fixed fraction of the Newton margin. The finitely many derivative formulas involved have polynomial perturbation bounds.

The initial coordinate rounding costs only the $\varepsilon_{\mathrm{round}}/2$ allowance already reserved in the initialization argument. It is what permits the nonlinear phase to forget the original partial-signing denominators.

Accuracy margins and tracking error. The weights satisfy $\varpi_i \ge \gamma/4$, and all derivative bounds and matrix condition numbers are polynomial in k, n for the fixed cubic predictor. The gradient and curvature tests in Lemma 5.4 and the Newton radius therefore have inverse-polynomial margins. The accepted decreases are bounded separately in each application. Choose the cache, derivative, jet, and coordinate-grid accuracies below a fixed fraction of these margins. The potential and KKT perturbation bounds validate comparisons for the exact model. Inexact Newton corrections contract errors to their specified floor at every stage, including initialization.

For endpoint scores use error at most δ, and trigger if an approximation to $\iota\beta_i v_i - (1 - y_i)$ or $\iota\beta_i u_i - (1 + y_i)$ is at least $-\delta$. If neither triggers, the exact state is light. A triggered move needs a feasibility repair at most $2\delta\|C_i\|$; choose δ smaller than a fixed fraction of $\min\{\gamma\delta_{\mathrm{end}}^{97/200}, \rho\delta_{\mathrm{end}}^3\}$. The reservoir pays for the repair, and the regular endpoint decrease in Lemma 5.7 retains a fixed fraction of its margin. Local trial coordinates are evaluated as $y + tr$ from the original dyadic base, without accumulating intermediate coordinate errors. At acceptance, round once and correct the optimizer again. Fixed-order powers and budget derivatives have polylogarithmic evaluation cost: the fixed algebraic power $(1 - y^2)^{97/200}$ is computed by rational bisection. Logarithms and exponentials needed for deferred rounding have convergent series with certified tails on the maintained ranges. □

7.4 Stable matrix products and inverse solves

The arithmetic constructions use products and inverses whose norms and condition numbers have already been bounded. These bounds determine the required guard precision.

Lemma 7.6 (Stable products and residual-certified inverses)**.** *Under the polynomial magnitude and inverse bounds in [SY26a, Section 4.10] and 5.9, the required matrix products and inverse solves attain their prescribed inverse-polynomial errors with $O(\log(2kn))$ working bits. An order-d_{sys} product or inverse costs $\widetilde{O}(d_{\mathrm{sys}}^{\omega_0})$ bit operations; in particular, $d_{\mathrm{sys}} = O(n^2)$ gives cost $\widetilde{O}(n^{2\omega_0})$. The structured rank-one actions of Lemma 5.6 attain their stated errors and bit costs with the same logarithmic guard precision.*

Proof. For the inverse, use A, Z_r, E_r, L and the exact residual recurrence from Section 5.9. Round each matrix product and iterate to the common grid. If its induced residual error is at most δ, then

$$\|E_{r_{\mathrm{aux}}+1}\| \le \|E_{r_{\mathrm{aux}}}\|^2 + \delta.$$

Take δ smaller than a fixed fraction of both the square of the initial spectral gap and the target residual. The same logarithmic iteration bound holds: while the residual is near one, its gap from one grows by a constant factor; below a fixed constant, the usual squaring convergence applies down to the error floor. Throughout, $\|\mathsf{Z}_{r_{\mathrm{aux}}}\| \le \|A^{-1}\|(1 + \|E_{r_{\mathrm{aux}}}\|)$ is polynomial. Hence $O(\log(2kn))$ bits per entry suffice for all products. A direct residual computation bounds the final inverse error by $\|A^{-1}\|\|E_{r_{\mathrm{aux}}}\|$.

For order $d_{\mathrm{sys}} = O(n^2)$, implement the products with the fast multiplication routine below. Inversion and multiplication by the right-hand sides then cost $\widetilde{O}(n^{2\omega_0})$ arithmetic and bit operations. The same procedure handles the order-n primal inverses. The residual and precision arguments above are unchanged. Approximate linear solves and KKT evaluations give the additive error asserted after Eq. (148).

Fast products.

For the exponent ω_0 in Eq. (132), stable recursive multiplication with exponent ω_0 is available by Demmel, Dumitriu, Holtz and Kleinberg [DDHK07]. The upper bound $\omega_0 = 2.371177$ follows from the strict bound $\omega < 2.371177$ of Dupont et al. [DEK+26]. The older bound of Coppersmith and Winograd [CW90] already suffices for a product cost of $\widetilde{O}({d_{\mathrm{sys}}}^{5/2})$. For polynomially bounded matrices and $p = O(\log(2kn))$ requested accuracy bits, their polynomial normwise error bounds require only $O(p+\log(2kn))$ working bits. Thus an order-d_{sys} product costs $\widetilde{O}({d_{\mathrm{sys}}}^{\omega_0})$ bit operations as well as arithmetic operations.

Rank-one actions. All factors have polynomial magnitude, including the fixed-order jets. Choose fixed bilinear recursive algorithms at the strict exponent bounds in Eq. (132). Each has logarithmic recursion depth and fixed coefficients at each level, so intermediate magnitudes and normwise error amplification are bounded by a fixed polynomial. Tensoring or permuting these fixed algorithms preserves that property. A fixed multiple of the logarithmic working precision therefore attains the prescribed error. This proves the bit bound. □

The initializer must preserve feasibility and the Newton decrement at every accepted stage. We apply the same residual-certified linear algebra to its capped barrier path.

Lemma 7.7 (Finite-precision capped initialization)**.** *Assume the bounds on inverse slacks, Hessian conditioning, and centers in [SY26a, Section 4.10]. Its initializer attains the requested objective accuracy with $O(\log(2nk/\varepsilon_{\mathrm{obj}}))$ working bits. It preserves the feasibility and decrement margins within the same stage count. Its bit cost is the corresponding arithmetic cost up to logarithmic factors.*

Proof. Use the inverse-slack bounds and barrier from [SY26a, Section 4.10], and Lemma 7.6 for the numerical solves. The Hessian is a sum of products of these inverses and coefficient maps, so its norm is polynomial. For completeness, the scalar and matrix cap terms give $\nabla^2 f \succeq (T_c + R_c)^{-2} I$, where T_c and R_c are the polynomial cap bounds of the initializer. The term $-\log b$ controls the scalar direction by T_c^{-2}, while $-\log\det(U - aI)$ and $-\log\det(V - aI)$ control the two matrix directions by R_c^{-2} in Frobenius norm. All other barrier Hessian terms are positive semidefinite. This is the lower bound in [SY26a, Eq. 138], and proves the required polynomial bound on the inverse Hessian. All coordinates, gradients, Newton directions, and objective parameters have polynomial magnitude. Consequently $O(\log(2nk/\varepsilon_{\mathrm{obj}}))$ accuracy and magnitude bits suffice to preserve feasibility and the decrement slack at every stage. Rounded Newton corrections need not accumulate errors: the decrement invariant is re-established after each step.

This also gives a constructive implementation: compute inverse slacks and the initial Newton-system inverse by short-precision Newton–Schulz iteration, with residual-based error bounds and safe polynomial norm/condition bounds. Recompute the Newton-system inverse at each subsequent stage. Stable fast multiplication supplies the required inverse-polynomial product accuracy with logarithmic guard precision. Determinants need not be evaluated by the algorithm; the barrier gradient and Hessian use only inverses and traces. Approximate b_{obj_0} and every oracle to a sufficiently small inverse-polynomial floor. The exact-center and Dikin arguments in [SY26a, Section 4.10] ensure that the prescribed floor works at every stage, without assuming a priori precision at unvisited centers. □

7.5 Smoothed root evaluations and matrix functions

At a corrected smoothed root, residual bounds suffice to evaluate the matrix functions used by the algorithm. The spectral implementation below uses their established spectral ranges.

Lemma 7.8 (Matrix functions and corrected smoothed states)**.** *Assume the geometric bounds of Section 6.1 and the state invariant of Section 6.7. The required order-n matrix logarithms, square roots and their inverses, Lyapunov solves, and entropy covariance actions attain the prescribed errors with $O(\log(2kn))$ working bits. An order-n spectral decomposition, certified by its residual, costs $\widetilde{O}(n^3)$ bit operations. If inverse and structured actions meet their error bounds, rounded Newton corrections preserve the state invariant and provide the required value, endpoint, and residual enclosures.*

Proof. The input conversion and grouped partial signing use their proved accuracy and additive L_{in} accounting. After that phase, the true model uses its retained matrices; numerical computations use the PSD approximations of Lemma 7.5. They are rank one for the reference rank-one routines; the higher-rank algorithms use Eq. (328) and the trace-map error bound in Eq. (329). The fixed contribution is stored symbolically as a combination of the initial k matrices with coefficients of absolute value at most two. Its matrix error, all coupling errors, and their derivative errors are bounded afresh by a fixed polynomial times the chosen cache error, independently of the number of steps or resets.

Use $p = C\log(2kn)$ accuracy bits in the nonlinear cache for a sufficiently large fixed constant C. Choose C to cover the prescribed inverse-polynomial action errors and the scalar and matrix-product guard bits. The input conversion is not rerun on the original L_{in}-bit data at each stage. Retained normalized rational inputs are converted once to the chosen cache accuracy. Their normalized magnitudes are bounded; no lower bound on their nonzero entries or eigenvalues is assumed. For higher-rank scale computation, division by the total input trace puts the trace-variance norm in $[1/(mn), 1]$, as shown in Section 3.11. This is a bound on the common normalization scale, not

on individual input eigenvalues. Its approximation is performed once in preprocessing and includes the original input-length charge.

Here is a constructive spectral subroutine for the order-n operations. A complex Hermitian Jacobi rotation annihilates a largest off-diagonal entry and reduces squared off-diagonal Frobenius norm by twice its squared modulus. Thus each rotation reduces this measure by a factor at most $1-2/(n(n-1))$. After $O(n^2\log(M_{\mathrm{aux}}/\eta))$ rotations it is at most η^2, for a known norm bound M_{aux} and requested residual η. A max-heap for pivot magnitudes and the two changed rows/columns give $\widetilde{O}(n)$ work per rotation; accumulating the rotations costs the same. Only a pivot above the stopping threshold is divided by when finding its complex phase. Rotations have norm one, so rounded similarity and orthogonality errors sum polynomially over the rotation count, while the off-diagonal recurrence has its usual additive-error floor. For $\eta=(2kn)^{-C}$ this remains $\widetilde{O}(n^3)$ work with $O(\log(2kn))$ working bits. Direct residual and orthogonality checks supply certified bounds. No small adjacent eigenvalue gap is used.

For a logarithm, the integral resolvent formula bounds perturbations of $D\log_{\mathsf{Q}}$ by $C\lambda_{\min}(\mathsf{Q})^{-2}$ times the matrix perturbation. For Lyapunov solves the residual is converted to solution error by $2\lambda_{\min}(T_U)$. Square-root and inverse-square-root perturbations have polynomial bounds on the same spectral ranges. These facts turn the residual-based decompositions into matrix-function actions without requiring individual eigenvectors to be accurately determined. Project approximate entropy actions to trace zero on both sides and symmetrize; on this space their exact smallest eigenvalue is inverse-polynomial, so the selected accuracy preserves positivity.

Scalar powers use the fixed rational exponent $97/200$, with certified bisection or series. Matrix logarithms, Lyapunov inverses, and Q_{Q} use gap-free order-n spectral decompositions with inverse-polynomial spectral lower bounds; the requested higher accuracy adds only polylogarithmic factors. Direct log-trace-exponential value evaluation uses a common spectral shift. On the bounded model its eigenvalue magnitudes divided by τ_{s} are $O(\log n)$ with fixed constants. No underflow assertion or unit cost for an unbounded exponential is used. There is also a direct residual-based value evaluation: put $h=\operatorname{tr}[\mathcal{C}-\tau_{\mathrm{s}}\log\mathsf{Q}]/(2n)$. Since $\operatorname{tr}[\mathsf{Q}]=1$, $f_{\tau_{\mathrm{s}}}(\tau_{\mathrm{s}}\log\mathsf{Q}+hE)=h$, so the absolute difference between $f_{\tau_{\mathrm{s}}}(\mathcal{C})$ and h is at most $\|\Pi_{\mathcal{T}}(\mathcal{C}-\tau_{\mathrm{s}}\log\mathsf{Q})\|$. This supplies an additional value error enclosure from the joint residual itself. The fixed rational real/imaginary Hermitian basis has polynomial condition and constant diagonal Gram factors for the full pair space; projection to trace zero adds only a polynomial factor. All analytical square roots of metrics used for error bounds need not be computed.

At a corrected Newton stage, approximate the root-anchor inverse and current actions below a fixed fraction of the contraction margin. The refinement error recurrence is a contraction plus additive error; Newton then restores the cache floor. At a curvature filter, instead use the explicit second-kind recurrence bound in its proof. These are two different error arguments; neither multiplies per-stage condition numbers over the whole signing trajectory. Root perturbation, coordinate-grid, value-comparison, and endpoint-score tolerances are chosen below fixed fractions of the proved energy radii and true potential decreases. All minima of weights, endpoint distances, and spectral values appearing in them have polynomial or inverse- polylogarithmic bounds. Final Rayleigh and solve residuals include true-versus-cached errors. □

7.6 Deterministic conditional choices and response blocks

The deterministic choices use conditional estimators with polynomially bounded values and derivatives. This supplies a common finite-precision implementation for the scalar and grouped choices.

Lemma 7.9 (Certified deterministic conditional choices)**.** *The scalar conditional choices in Section 5.10 take $\widetilde{O}(k^2)$ bit operations. Certified evaluation within their comparison margins uses*

$O(\log(2kn))$ *working bits. The grouped estimator in Lemma 5.9 uses* $O(\log(2D/\varepsilon))$ *working bits, with its inherited short-data factors, and preserves its arithmetic cost up to logarithmic factors. In both cases the accepted choices preserve the conditional-estimator guarantee, including when the two candidate values nearly tie.*

Proof. Use the filter F, coefficients λ_f, π_n, and conditional estimator from Section 5.10. For the scalar choices, if v is the sum of the fixed columns with their chosen signs and J indexes the remaining columns, then

$$\mathbb{E}[\|F\xi\|^2 \mid v] = \|v\|^2 + \sum_{j\in J} \|F_{:j}\|^2, \quad \mathbb{E}[\cosh((F\xi)_i) \mid v] = \cosh(v_i) \prod_{j\in J} \cosh(F_{ij}).$$

Maintain v and these products; each candidate decision requires $O(k)$ scalar operations. There is no large-exponent precision cost: the maintained conditional value of Φ is positive, while the conditional quadratic expectation is at most $4k$. Thus $\cosh(v_i) \le 4k/\lambda_f$. A trial adds at most two to $|v_i|$, and the remaining products are between one and eight. All numbers and derivatives used in a comparison are therefore polynomially bounded. Certified scalar Taylor approximations with logarithmic precision suffice, including when the two choices nearly tie.

Grouped deterministic directions. In Lemma 5.9, the conditional estimator uses the short matrix V. At every chosen prefix their value stays positive, whereas the conditional quadratic term is at most $\|V\|^2 r < 2r$. Hence each conditional hyperbolic cosine is at most $200r$. At either next candidate it increases by at most a fixed factor, since $|V_{ij}| \le 65/64$. The remaining products are bounded by $e^{(65/64)^2/2} < 2$. All arguments are $O(\log r)$ and all intermediate magnitudes are polynomial. Taylor evaluation and outward error bounds consequently need only $O(\log(2D/\varepsilon))$ bits, with the inherited short-data factors. The divisions removing a factor $\cosh(V_{ij}) \ge 1$ are stable. □

All response columns at one corrected root use the same spectral operators. Their simultaneous approximation can therefore be controlled in Frobenius norm.

Lemma 7.10 (Precision for block responses)**.** *Under the hypotheses of Lemma 5.10, all Hessian and response blocks can be evaluated within its prescribed inverse-polynomial operator tolerances using* $O(\log(2kn))$ *working bits, with a fixed multiplier. Their bit cost is the block-construction arithmetic bound in that lemma, up to logarithmic factors.*

Proof. Use the same approximate spectral operators for all columns. The energy preconditioning bounds give polynomial similarity condition numbers. The normalized Chebyshev recurrence has polynomial error amplification: the scalar propagation ratios are at most one, and a second-kind polynomial of degree j on the preconditioned spectral interval has norm at most a polynomial similarity factor times $j+1$. The same bound applies to a matrix of right-hand sides in Frobenius norm. Demand a polynomially smaller per-column error, so the sum of squared column errors is below the prescribed squared operator tolerance. This changes only logarithmic precision and iteration factors. No eigengap is needed, and no order-n^2 dense inverse is formed. □

7.7 Finite sampling

We realize the feature distribution using finitely many fair bits. A coupling to the ideal samples quantifies the effect on the feature covariance.

Lemma 7.11 (Finite feature sampling)**.** *Under the hypotheses and sample count s of Lemma 6.5, its feature sampler has a finite implementation using dyadic uniforms with* $b = O(\log(2kn))$ *bits, with*

a sufficiently large fixed multiplier. It can be coupled to the ideal sampler so that, except on an event of probability at most $1/100$*, the feature covariance changes by operator norm at most* $\mu/16$*. On the intersection with the ideal feature event, the represented features satisfy the two-sided bound with slack* $1/2, 2$*. Every finite sample has polynomially bounded magnitude, including samples outside that event. The sampler and its matrix-function evaluations use logarithmic working precision.*

Proof. For the finite sampler use midpoints of dyadic uniform intervals for t and the Box–Muller uniforms. Pair independent real normals into circular complex normals. Couple these finite samples to ideal samples using the same uniform intervals. Except on an event of probability $1/100$, every radial uniform is at least $(100ns)^{-4}$ and every normal is at most $C\sqrt{\log(2ns)}$. On this event all sampling errors are a fixed polynomial times the scalar grid error. The eigenvalues of Q lie in an inverse-polynomial interval, and $t \in [0, 1]$; the Frobenius divided-difference bounds for $\mathsf{Q}^{t/2}$, as well as its t derivative, are polynomial uniformly in t. A gap-free order-n diagonalization and certified scalar powers therefore approximate all r_b, v_b to any fixed inverse-polynomial accuracy, without identifying individual eigenvectors at repeated eigenvalues. The normalization constants can also be replaced by short dyadics at this accuracy.

Store these vectors and constants, and define the represented features by the exact rank-two and centering formulas using the same short Q, F_1, G. The feature perturbation and hence the perturbation of $VV^\top/s$ are as small as prescribed, since all factors have polynomial magnitude. Choose them to give operator error at most $\mu/16$. The finite version of the two-sided bound, with slack $1/2, 2$, follows. Every finite Box–Muller output is bounded by $C\sqrt{b}$ if b bits are used, even on the excluded event. Thus all represented features have known polynomial magnitude for every sample. Taking b a sufficiently large fixed multiple of $\log(2kn)$, with the polylogarithmic factors already displayed, suffices. There is no exact-real or unbounded rejection-sampling oracle. □

7.8 Uniform precision for nested solves

The recursion in Lemma 6.10 increases the number of calls. Residual correction allows every level to use the same accuracy target.

Lemma 7.12 (Depth-independent precision for nested solves)**.** *Use the parameters and frozen feature and index data of Lemma 6.10, with* $2 \le h \le \mathscr{J}^2$ *and* $\mathscr{J} = 1 + \log_2 N$*. There are fixed constants* B_0, D_0*, independent of the depth and sketch sizes, with the following properties. On the simultaneous good sampling event, each completed recursive solve, with matrix* K *and right-hand side* b*, returns* v *satisfying*

$$\|v - K^{-1}b\| \le \eta\|b\|, \qquad \eta := N^{-B_0},$$

and the shifted solves attain their prescribed errors. All numerical entries use at most $D_0\mathscr{J}$ *accuracy and magnitude bits. The finite conditional index events have joint probability at least* $99/100$ *given the feature event, and the original verified curvature trial retains conditional success probability at least* $1/24$*. All samples obey prescribed iteration and magnitude caps; the arithmetic call recurrence in Lemma 6.10 therefore also gives its bit bound with the stated logarithmic overhead.*

Proof. Freeze one short root, its consistently defined symmetric Hessian surrogate, all finite feature factors, and all index lists. No action changes these objects. The invariant in Eq. (248) is algebraic and holds for any depth. Store each coefficient from its closed formula $\iota_j = \bar{k}q_j/(\mu q_1 d_j e_j)$ rather than by accumulating rounded products of previous coefficients. All sizes are integers and μ is a fixed short rational, so these positive coefficients have uniformly $O(\log N)$ encoding lengths. On every finite sample the feature entries have magnitude at most N^{C_0}, for a fixed C_0, since $h \le \mathscr{J}^2$

makes the moment order and the finite-sampling encoding lengths $O(\mathscr{J})$. Consequently, at every level,

$$1 \le \lambda_{\min}(K_j), \qquad \|K_j\| \le 1 + \iota_j d_j e_j \max_{a,b} |V_{ab}|^2 \le N^{C_1}.$$

Similarly $\lambda_{\min}(W_j) \ge \lambda_j \ge 1$ and $\|W_j\| \le N^{C_1}$. The factors in their Woodbury formulas, their positive scales and inverse scales, and all individual operator norms are bounded by N^{C_2}. These exponents do not grow with depth: $\iota_j d_j e_j \le 2q_j/d_0$ telescopes exactly, rather than accumulating a new condition factor at every level. Sample-list sizes are at most $N^2 \mathscr{J}^{C_3}$, with a fixed exponent; indices may be composed once into lists of original coordinates and sample columns.

The proofs of the simultaneous entry and covariance event and of uniform ridge sampling now use $p = \lceil 128 \log_2(2kN\ell_{\mathsf{Q}}(h+1)) \rceil$. For $h \le \mathscr{J}^2$ this is still $O(\mathscr{J})$. The same moment and finite-grid bounds dominate the polynomial factors in all dimensions. Each conditional index-sampling failure is at most $1/(1000(h+1))$; unioning them and the finite rejection-cap failure leaves probability at least $99/100$, conditionally on the feature event. Thus these probability bounds are uniform in the present range of depths.

On the simultaneous good event the preconditioned systems are similar to symmetric positive matrices in their prescribed slack intervals, through transformations whose condition numbers are at most N^{C_4}. In Eq. (257), $|T_j| \le 1$ on $[-1,1]$ and $T_j(\varrho) \ge 1$. Therefore each exact normalized iterate has norm at most $N^{C_5}\|b\|$, independently of the requested accuracy and the iteration count. Normalize every right-hand side by a dyadic scalar before a recursive call, then restore that scale. This avoids propagating large right-hand sides as unbounded intermediate integers.

Here is an explicit error bound for the normalized recurrence. A vector update error e_i contributes at a later index J the term

$$\frac{T_i(\varrho)}{T_J(\varrho)} \mathsf{A}_{J-i}(D) e_i$$

(up to the first-step convention). The scalar ratio is at most one and $\|\mathsf{A}_{J-i}(D)\| \le N^{C_4}(J-i+1)$. Thus the total vector error is at most $N^{C_4}(J+1)^2 \max_i \|e_i\|$. Errors in the clipped scalar ratios accumulate at most linearly, since their update is nonexpansive on $[0,1]$. Their contribution, and the contributions of action and coefficient errors, are bounded by a fixed polynomial in N, J times the scalar unit and the right-hand-side norm.

We use residual correction to keep the accuracy request independent of depth. Fix $\eta := N^{-B_0}$, with a sufficiently large constant B_0 chosen below. Every completed recursive solve is required to return v with $\|v - K^{-1}b\| \le \eta\|b\|$. Normalize each right-hand side by a dyadic scalar within its own call; restore that scalar when returning, without multiplying scales along the recursion tree. A zero right-hand side returns zero.

First construct a coarse solve $S_j(b)$ at each level. Run the normalized Chebyshev recurrence for $O(\sqrt{\beta/a_{\mathrm{spec}}} \log N)$ steps, using completed child solves to accuracy η relative to their right-hand sides. On the good event, all condition, similarity, Woodbury, and recurrence factors are bounded by N^{C_8} for a fixed C_8 independent of h and the level. The preceding error convolution therefore bounds the effect of child errors by $N^{C_9}\eta\|b\|$, for one fixed C_9. Local arithmetic errors have the same bound with the scalar unit u in place of η. The exact recurrence error can be made smaller than $N^{-C_1-3}\|b\|$ by its fixed logarithmic multiplier. Choose $B_0 > C_9 + C_1 + 10$ and u sufficiently smaller than η. Since $\|K^{-1}b\| \ge N^{-C_1}\|b\|$, these choices give

$$\|S_j(b) - K^{-1}b\| \le \tfrac{1}{4}\|K^{-1}b\|.$$

This assertion needs only the same completed-child accuracy η, not a more accurate child when the parent requests a smaller residual.

Starting at $v_0 := 0$, compute fresh residuals and corrections,

$$r_\ell := b - Kv_\ell, \qquad v_{\ell+1} := v_\ell + S_j(r_\ell),$$

rounding each operation. Exact residuals would contract solution error by a factor at most $1/4$. With residual and arithmetic errors, all iterates stay polynomially bounded and

$$\|v_{\ell+1} - K^{-1}b\| \le \tfrac{1}{4}\|v_\ell - K^{-1}b\| + N^{C_{10}}u\|b\|,$$

where C_{10} is uniform in depth. The same estimates follow inductively for rounded iterates. Choose $u \le \eta N^{-C_{10}-C_1-10}$ and a fixed $O(\log N)$ correction cap. Then the solution error is below $\eta\|b\|/(4N^{C_1})$; a direct residual enclosure certifies the requested $\eta\|b\|$ solution error, using $K \succeq I$. Stopping and scaling tests use the same polynomial margins. The terminal inverse supplies the induction base. Induction from the leaf now proves the same output accuracy at every level, with $O(\log N)$ accuracy and magnitude bits throughout. The number of levels changes the number of calls but not their accuracy exponents. The shifted system is handled identically using $B \succeq RI$.

The terminal Gram and all materialized actions have inverse-polynomial error, and the explicit ridge and Gram forms supply the positivity bounds. No intermediate spectral decomposition is charged implicitly. Root and feature matrix functions approximate the same frozen data; their logarithmic precision is included in u. The bounded outer filter requires only inverse-polynomial action errors by its array contraction bound. Choose B_0 also large enough for these final norm and quotient margins. Structured rational real/imaginary coordinates and their fixed Gram weights add only polynomial condition factors.

Prescribe the common logarithmic precision, all iteration lengths, and a polynomial magnitude cap $N^{C_{11}}$ in advance. Normalize recursive right-hand sides as above and reject on any violated cap, required residual or positivity enclosure, or failed final quotient. Even off the sampling event, every operation and loop has these caps. On the good event the induction proves that no accuracy test removes the successful independent sign-overlap event. The conditional success probability remains at least $1/24$. □

7.9 Certified spectral filters

The normalized-squaring witness is evaluated on symmetric matrices at every step. Its error can be bounded by propagating each local perturbation through the remaining exact powers.

Lemma 7.13 (Stable normalized squaring)**.** *Use the normalized-squaring sequence of Section 5.9, with $q = 2^r$ and matrices of order k; write D_j and $\widehat{D}_j$ for the exact and computed iterates. If the initial error and every local normalized-squaring error are at most $\eta < 1/2$, with symmetrization after each product, then*

$$\|\widehat{D}_r - D_r\|_F \le 8q\sqrt{k}\eta.$$

For the prescribed polynomially bounded q, inverse-polynomial local errors preserve a computed Rayleigh quotient at most $-\delta/4$ and certify negative curvature for the true potential. The complete witness computation uses $O(\log(2kn))$ working bits and $\widetilde{O}(n^{2\omega_0})$ bit operations.

Proof. For a nonzero symmetric matrix X and an integer $s \ge 1$, define $\mathcal{N}_s(X) := X^s/\|X^s\|_F$. The product rule gives $\|\mathrm{D}(X^s)[E]\|_F \le s\|X\|^{s-1}\|E\|_F$ for symmetric E. Since $\|X^s\|_F \ge \|X\|^s$, differentiating the normalization gives

$$\|\mathrm{D}\mathcal{N}_s(X)[E]\|_F \le \frac{s}{\|X\|}\|E\|_F \le \frac{s\sqrt{k}}{\|X\|_F}\|E\|_F.$$

Consequently, on a segment within Frobenius distance $1/2$ of a symmetric matrix of unit Frobenius norm, $\mathcal{N}_s$ is $2s\sqrt{k}$-Lipschitz. Also, $\mathcal{N}_2^{\circ j} = \mathcal{N}_{2^j}$ for $j \geq 1$.

Symmetrize after each product. Let $\widehat{D}_j$ be the computed iterates, with initial error and local errors bounded by $\eta < 1/2$:

$$\|\widehat{D}_0 - D_{\mathrm{pow}_0}\|_F \leq \eta, \qquad \|\widehat{D}_{j+1} - \mathcal{N}_2(\widehat{D}_j)\|_F \leq \eta.$$

Each computed iterate has Frobenius norm between $1-\eta$ and $1+\eta$. Insert these errors one at a time and propagate each through the remaining exact normalized squarings. The remaining powers are $q, q/2, \ldots, 2$; the final local error is unchanged. The preceding Lipschitz bound and this geometric sum give

$$\|\widehat{D}_{r_{\mathrm{aux}}} - D_{\mathrm{pow}_{r_{\mathrm{aux}}}}\|_F \leq 8q\sqrt{k}\eta.$$

Choose an inverse-polynomial accuracy $0 < \epsilon \leq \delta/(100L)$ and define $\eta := \epsilon/(8q\sqrt{k})$. Since q is polynomially bounded in k, n, this local accuracy is inverse polynomial as well. Moreover, $\|\widehat{D}_j^2\|_F \geq (1-\eta)^2/\sqrt{k} \geq 1/(4\sqrt{k})$. Thus all normalizations are polynomially conditioned. The stable fast products above, Frobenius norms, and scalar square roots achieve the required local errors with $O(\log(2kn))$ working bits per entry and polylogarithmic overhead. The final trace changes by at most $L(2+\epsilon)\epsilon$, leaving a fixed fraction of its negative margin. Compute column norms and quadratic forms with conservative margins, discarding only columns below comparable inverse-polynomial thresholds. The trace and retained-column bounds then still supply a witness with Rayleigh quotient at most $-\delta/4$. Finally normalize that column and check its Rayleigh quotient against the computed Hessian directly. The Hessian approximation margin transfers this negative curvature to the true potential. The complete filter has arithmetic and bit cost $\widetilde{O}(n^{2\omega_0})$. □

The outer filters use a fixed self-adjoint contraction throughout their evaluation. The following array estimate controls additive action and rounding errors.

Lemma 7.14 (Stability of the de Casteljau array)**.** *Let $0 \preceq T \preceq I$ be fixed, and evaluate a degree-d_f de Casteljau array by the updates $(a, b) \mapsto (I-T)a+Tb$. Each level is a contraction in the Euclidean norm of the concatenated array. If every vector update has error at most η, the final output error is at most $d_f\sqrt{d_f+1}\eta$. Consequently the bounded polynomial arrays used in Sections 6.6 and 6.9 attain their prescribed inverse-polynomial errors with $O(\log(2kn))$ working bits.*

Proof. For any vectors a, b,

$$\|(I-T)a + Tb\|^2 \leq \langle a, (I-T)a\rangle + \langle b, Tb\rangle.$$

Summing over adjacent pairs proves that each level is a contraction in the Euclidean norm of the concatenated vector array. An error of norm at most η in each vector update therefore causes final error at most $d_f\sqrt{d_f+1}\eta$. Every exact intermediate vector has norm at most $\sqrt{k}$. All inverse actions approximate the same frozen T, so their errors enter as these additive update errors. Inverse-polynomial action and rounding errors give inverse-polynomial output error. The $O(d_f^2)$ actions change only logarithmic factors in the cost, and all vector entries use $O(\log(2kn))$ bits. □

We combine the finite sampler and nested solves with the array estimate to implement the spread-filter certificate.

Lemma 7.15 (Certified finite flat-filter trials)**.** *Under the hypotheses of Lemma 6.15, one capped finite trial uses $O(\log(2kn))$ working bits and $\widetilde{O}(n^\chi)$ bit operations, including unsuccessful trials. Its conditional success probability is at least $1/64$. Every accepted direction is normalized and satisfies both true curvature and spread inequalities in that lemma; all other samples obey the same finite work and magnitude caps.*

Proof. For Lemma 6.15, use the de Casteljau array in Eq. (273), with degree d_f and fixed operator T. Each exact intermediate is a polynomial between zero and one on $[0,1]$, applied to ξ, so its norm is at most $\sqrt{k}$. Lemma 7.14 applies: if each vector update has error at most η, the final error is at most $d_f\sqrt{d_f+1}\eta$. Choose inverse-polynomial action and scalar errors below the normalization and quotient margins divided by this factor. Thus $O(\log(2kn))$ working bits suffice, with a polynomial magnitude cap.

Freeze one symmetric surrogate, its exact matrix-function definitions, all feature factors and index lists during this computation. Every inexact application approximates the same T. Lemma 7.12 supplies these inverse-action errors with the same logarithmic precision. Its residual corrections add only fixed logarithmic factors at each level; choose its fixed overhead integer D before the integer sketch schedule. No recursion depth is hidden as a constant. The feature and conditional index events retain probability at least $(4/5)(99/100)$; the independent sign event has probability greater than $1/16$. After numerical slack their product is still greater than $1/64$.

Reject a candidate unless a lower enclosure for its norm is at least $\sqrt{\pi k/8}$ and both inequalities in Eq. (270) are certified for the true model. The norm threshold has an inverse-polylogarithmic lower bound, so normalization is conditioned. The true weight and Hessian enclosures absorb the surrogate discrepancy and the final grid errors. On the good events none of these tests rejects the ideal estimates with their slack. On every other sample, all lists, loops, precisions and magnitude checks retain their fixed caps. Thus neither failed sketches nor failed sign trials introduce unbounded work. No negative eigenspace or individual eigenvector is computed by this routine. □

7.10 Phases, sweeps, and endpoint continuation

The remaining implementation issue is compatibility across accepted moves and changes of face. The phase parameters and corrected-state margins provide a common precision for these transitions.

Lemma 7.16 (Certified sweeps and face continuation)**.** *Assume the corrected-state invariant and the phase, sweep, and endpoint hypotheses in Sections 6.9 and 6.11. The reservoir phases, frozen-root sweeps, deferred attenuation, and regular endpoint homotopies admit finite-precision implementations using $O(\log(2kn))$ working bits. Their accepted moves preserve the true feasibility, decrease, and discrepancy allowances, and each terminal correction represents the actual resulting face. Finite trial caps and certified filter success margins are retained. These routines have their arithmetic costs up to logarithmic factors in the bit model, with the original input length charged only in preprocessing.*

Proof. **Reservoir phases.** Use the phase weights and spread filter of Section 6.9. The exact numerical Schur surrogate, matrix-function definitions, root data, and weights are frozen within each filter. True-versus-cache errors include the new phase weight and flat-polynomial accuracy. Lemma 7.12 remains applicable because the new outer degree, bit target and list counts are fixed polylogarithms. By Lemma 7.15, every successful direction satisfies true, enclosed quotient and spread bounds; each chosen candidate has true certified decrease. Conservative root, endpoint, value, grid, phase-weight, and action accuracies are below fixed fractions of the same inverse-polynomial or inverse-polylogarithmic margins. A root correction restores its cache floor rather than multiplying past error factors.

All real/complex matrix operations use the same rational Hermitian coordinates and exact trace-zero projection. No phase choices of input vectors or accurate individual eigenvectors at a repeated eigenvalue are needed. The rank-defect covariance is solely a proof of spectral multiplicity; it creates no covariance or SDP oracle. The polynomial filter needs no negative-eigenspace basis.

The existing finite fallback certifies arbitrary curvature when necessary. Its slower progress has been included in both deterministic caps and expected costs, rather than presuming a spread fallback.

All frozen signed coefficients still refer to the original retained matrices with absolute symbolic coefficients at most two. Batched attenuation deletes only positive coupling slots and never changes the recorded fractional signs. The final joint rounding is identical. The one-time input conversion is not repeated, and original long rational inputs do not enter nonlinear arithmetic. Polylogarithmic working precision therefore retains additive L_{in} in the bit bound and the mn^2 input term in both models.

Frozen sweeps and positive continuation. We implement the sweeps and regular endpoint homotopies of Section 6.11. The implementation never evaluates an exact optimizer during a sweep. Its frozen scalar data come from a corrected short root, and their certified errors are below fixed fractions of $l_N\sqrt{\kappa_i^0/N}$. Positive $\kappa_i^0 \geq \gamma$ have known inverse-polynomial lower bounds. Square roots, arc limits, and all coordinate endpoints have dyadic upper/lower enclosures. The common grid is chosen once below the smallest of the arc, cutoff, value-decrease and root-radius margins. All these minima are inverse polynomial up to fixed logarithms. Coefficient and scalar errors are included in the chord contraction, not multiplied through all sweep updates. A final root and value correction certifies the true sweep decrease. No eigenvector or curvature oracle is inserted into the deterministic sweep.

For regular continuation, store the nonnegative scalar coefficients of $\widehat{\mathsf{A}}, \widehat{E}_i$ multiplying the true retained rank-one matrix; the numerical cache uses the corresponding PSD approximations. Their errors, including the conservative trigger repair, are chosen below the initial auxiliary residual and endpoint-drop margins. Coefficient rounding preserves the positive additions. Store s on a common dyadic grid and evaluate $r = e^{-s}$ from it; do not accumulate rounded multiplicative updates. A stage uses the joint residual and a certified anchored inverse, all using $O(\log(2kn))$ working bits. The extra linear penalty introduces no new determinant or SDP solve. The final $r = 0$ correction is an optimizer of the actual new face. □

7.11 Short-precision barrier sign selection

The grouped reducers need a sign with a fixed barrier increase. The directional derivative supplies such a sign with a prescribed additive error. This avoids comparing products of all coordinate denominators.

Lemma 7.17 (Short-precision barrier sign selection)**.** *Let $r > 2D$ be the number of active coordinates, let $H \geq 1$ be an integer, and let $\sigma_{\mathrm{r}} > 0$. Suppose $s_i = 1 - |x_i| > \sigma_{\mathrm{r}}$ and $h_i = s_i\widetilde{v}_i/H$, where $\|\widetilde{v}\|_\infty \leq H/4$ and $\|\widetilde{v}\|_2^2 \geq (r-D)/4$. The sign chosen from the dyadic approximation in Eq. (287) satisfies*

$$\Phi(x+\tau h) - \Phi(x) \geq (2/25 - 1/256)\frac{r-D}{H^2}. \tag{330}$$

After inward rounding on any dyadic grid of mesh $u \leq \sigma_{\mathrm{r}}/(2^{16}H^2)$, the increase is greater than $(r-D)/(16H^2)$. In particular this applies to Eq. (282) and to the grid in Lemma 10.2. Suppose each coordinate of x and $\widetilde{v}$ is represented as a signed integer numerator divided by a positive power of two, with both numerator and denominator having at most p bits. Then the sign can be computed with $O(p + \log(2rH))$ bits per intermediate number and $O(r)$ rational divisions. Under the short-data hypotheses of Lemma 6.18, its bit cost is $\widetilde{O}(r)$, with logarithmic factors in $2D/(\varepsilon\zeta_{\mathrm{f}})$.

Proof. Write $a_i = 2x_i\widetilde{v}_i/(H(1+|x_i|))$ and $g = \sum_i a_i$. Each $|a_i| \leq 1/4$. With $q = \lceil \log_2(512H^2)\rceil$, round each a_i to a multiple $\widehat{a}_i$ of 2^{-q} with error at most 2^{-q}. Sum these rounded numbers exactly

on that common grid. The resulting $\widehat{g}$ obeys

$$|\widehat{g} - g| \le r2^{-q} \le \frac{r}{512H^2} \le \frac{r-D}{256H^2}. \tag{331}$$

Choose $\tau = 1$ when $\widehat{g} \ge 0$, and $\tau = -1$ otherwise. Then $\tau\widehat{g} \ge 0$, so $\tau g \ge -(r-D)/(256H^2)$. This conclusion includes a tie and does not require deciding the exact sign of g.

For $0 \le t \le 1$, put $z_i = x_i + t\tau h_i$. The distance from z_i to an endpoint nearest to x_i is at most $s_i + |h_i| \le 5s_i/4$. Both endpoints remain outside the segment, since $|h_i| \le s_i/4$. Hence

$$\phi''(z_i) = \frac{1}{(1-z_i)^2} + \frac{1}{(1+z_i)^2} \ge \frac{16}{25s_i^2}. \tag{332}$$

Taylor's formula with integral remainder gives

$$\Phi(x + \tau h) - \Phi(x) \ge \tau g + \frac{8}{25H^2}\|\widetilde{v}\|_2^2. \tag{333}$$

Here $h_i^2/s_i^2 = \widetilde{v}_i^2/H^2$. Substituting the lower bounds on τg and $\|\widetilde{v}\|_2^2$ proves Eq. (330). Every candidate coordinate has endpoint distance at least $3\sigma_{\mathrm{r}}/4$. On its inward-rounding segment, $|\phi'| \le 4/\sigma_{\mathrm{r}}$. The total rounding loss is therefore at most $4ru/\sigma_{\mathrm{r}} \le (r-D)/(8192H^2)$. Since $2/25 - 1/256 - 1/8192 > 1/16$, the claimed increase survives rounding. For a newly frozen coordinate, the value $\log(2/\sigma_{\mathrm{r}})$ is at least its barrier value after the step and inward rounding. Indeed its endpoint distance before freezing is at least $3\sigma_{\mathrm{r}}/4$ and $1 - x_i^2 \ge 1 - |x_i|$. Thus the finite assignment used in Lemmas 6.18 and 10.2 cannot decrease the bookkeeping potential.

Write $x_i = N_i/2^{b_i}$ and $\widetilde{v}_i = M_i/2^{c_i}$, where N_i, M_i are signed integers with at most p bits and $0 \le b_i, c_i \le p$. Then

$$a_i = \frac{2N_iM_i}{H2^{c_i}(2^{b_i} + |N_i|)}. \tag{334}$$

Its numerator and denominator have $O(p + \log(2H))$ bits and are formed by a constant number of integer operations. Here p bounds the full numerator and denominator lengths; a short encoding of an arbitrarily large exponent alone would not imply this bound. Computing $\widehat{a}_i$ by integer division after scaling by 2^q uses $O(p+q+\log(2H))$ bits. The exact common-grid accumulator has magnitude at most $r(1/4 + 2^{-q})$, and consequently uses $O(q + \log(2r))$ bits. Thus every intermediate number has $O(p + \log(2rH))$ bits. Elementary integer arithmetic on these short operands gives the stated bit cost. No products across coordinates are formed.

The coefficient grid in Eq. (282) has a logarithmic grid exponent, and $|x_i| \le 1$, so its numerators and denominators have logarithmic length. The directional evaluations in Lemma 7.3 round to fixed dyadic grids at their prescribed absolute accuracy. Their grid exponents and the passing bound $\|\widetilde{v}\|_\infty \le H/4$ give the same bound on the full numerator and denominator lengths. The deterministic construction in Lemma 5.9 uses the same coefficient grid; summing its dyadic matrix columns adds only $O(\log r)$ numerator bits. Thus both implementations satisfy the stated hypothesis on p. The new sign changes neither the direction tests nor the bound on $\|Ch\|_2$. The progress, residual, step-cap, and fallback arguments of the small reducer therefore retain their stated bounds. The sign proof uses no column-norm or ridge assumption beyond the displayed bounds on h and $\widetilde{v}$. It consequently also applies to the energy-controlled dense groups in Lemma 10.2, with their own σ, mesh, and residual certificate. Their coefficient mesh exponent is the maximum of inherited and prescribed logarithmic exponents. Rounding the direction at its specified absolute tolerance also uses a logarithmic grid exponent, so the same numerator and denominator bounds apply. □

7.12 Resident batches and count-sensitive solves

The final randomized implementation changes the preprocessing sketch, the matrix-function evaluations, and the sketch sizes. We verify that these changes preserve logarithmic numerical precision and the arithmetic counts in Section 6.16. The factor construction below also implements the dense preprocessing sketch in Lemma 10.1.

Lemma 7.18 (Precision for streamed factors and rational matrix functions)**.** *For a fixed short group matrix B and a bounded finite sample Ω, form $U = B^{\mathsf{T}}\Omega$, $Y = BU$, and $K = U^{\mathsf{T}}U + TI$, where $T \geq 1$. Suppose the dimensions, magnitudes, inverse ridge, and requested inverse accuracies are bounded by fixed polynomials in $2mn$. These factors and their Woodbury actions can be computed with $O(\log(2mn))$ numerical bits. Their bit costs equal the arithmetic costs of the selected matrix products and actions up to logarithmic factors. In particular this applies to the preprocessing in Lemma 10.2. In the nonlinear phase, Lemmas 6.34–6.37 have their stated bit costs with $O(\log n)$ accuracy and magnitude bits per entry.*

Proof. Freeze the group matrix and finite sample, and take all target accuracies to be positive short rationals. Let $L_{\mathrm{aux}} \geq 2$ be a known polynomial bound for $1+D+r+s+T+\delta^{-1}+\nu^{-1}+\|\Omega\|+\|B\|$. If the streamed U error is at most e and the subsequent Y action error is at most e, then

$$\|\widehat{Y} - Y\| \leq 2L_{\mathrm{aux}}e, \qquad \|\widehat{K} - K\| \leq 4L_{\mathrm{aux}}^2 e.$$

These estimates include the Gram error and symmetrization. Since $K \succeq TI \succeq I$, for $e \leq (2L_{\mathrm{aux}})^{-16}$ the inverse perturbation identity gives

$$\widehat{K} \succeq (T/2)I, \qquad \|\widehat{Y}\widehat{K}^{-1}\widehat{Y}^{\top} - YK^{-1}Y^{\top}\| \leq 32L_{\mathrm{aux}}^8 e < \delta/64.$$

Allocate each column error below e/s and the small inverse error below its explicit polynomial amplification bound. This preserves the ridge-sketch interval. Lemma 7.4 applies to the same frozen factors with the prescribed polynomial iteration count. Its additive recurrence-error estimate requires only a larger fixed logarithmic multiplier. For a grouped reducer with the state and direction bounds of Lemma 7.17, use its short-precision sign rule. The lemma preserves the required barrier increase and residual allowance using logarithmic-length summands and a common-grid sum.

For nonlinear matrix functions, the known spectral endpoints and requested inverse-polynomial error have $O(\log n)$ bits. The shifted inverses in Lemma 6.34 use Lemma 7.6. Their normalized powers have norm at most $1/3$, so truncation and arithmetic errors add with polynomial amplification over logarithmically many bands and terms. In Lemma 6.35, normalized Taylor terms have norm at most $e^{\|L\|} = n^{O(1)}$. A term error is amplified by at most this factor in later terms. The ordered covariance expansion and batched different-power recurrence obey the same bound. In Eq. (322), exact updates are nonexpansive in the current error; errors in their shifted inverses and products add along the prescribed finite sequence. The iterate norms remain polynomial. Reserve half the target for truncation or contraction, and half for these explicit arithmetic errors. Stable recursive products therefore give the claimed bit cost at a fixed $O(\log n)$ precision.

For a possibly non-Hermitian response matrix M, the certificate in Lemma 6.26 extends by setting $Z := M^*M/a^2$ and $Q := I - Z$. If $\|M\| \leq a/2$, then $0 \preceq Z \preceq I/4$. Its logarithmic binomial series constructs X with a verified enclosure $\|X^*QX - I\|_F < 1/2$ using the same short products. Conversely this certificate implies $Q \succ 0$ by congruence, hence $\|M\| < a$. Reserve this factor-two gap in the six relative-response tests. Thus the non-Hermitian similarities require no long exact trace powers or spectral decomposition. Every finite evaluation has prescribed iteration and magnitude caps and rejects on failure of its required enclosure. □

Lemma 7.19 (Uniform precision for the resident trials)**.** *The trial in Lemma 6.38, the resident masks, and the curved and bulk continuations can be implemented with $O(\log n)$ numerical bits per entry. Their bit costs equal their arithmetic bounds up to the logarithmic factors explicitly allowed there.*

Proof. The proof of Lemma 7.12 extends from terminal size $n^{4/3}$ to n. Its norm, inverse, and coefficient estimates use the invariant in Eq. (248), positive ridges, and polynomial sizes. None uses the stronger terminal floor. With $h \le L^2$, the moment order and sample encoding lengths remain $O(\log n)$. The last two equal sizes allow the leaf factor, Gram, and inverse to be materialized at their charged square-product cost.

Freeze every feature factor and index list. At every level require the same completed-solve guarantee $\|v - K^{-1}b\| \le \eta\|b\|$, $\eta := n^{-B_0}$, with fixed B_0. Normalize each right-hand side within its own call. The normalized Chebyshev error convolution gives a coarse solve S with $\|S(b) - K^{-1}b\| \le \|K^{-1}b\|/4$ using this same child accuracy. Fresh residual correction $v^+ := v + S(b - Kv)$ then obeys

$$\|v^+ - K^{-1}b\| \le \tfrac{1}{4}\|v - K^{-1}b\| + n^C u\|b\|,$$

for a fixed C independent of depth and common rounding unit u. The residual is a direct action of the same frozen operator; it does not call a recursive inverse. Choose $u = n^{-D_0}$ with fixed sufficiently large D_0, and perform the fixed $O(\log n)$ correction cap from Lemma 7.12. Induction from the leaf proves the common η interface at all levels. Child accuracy does not grow with depth. The bounded filter array in Lemma 7.14 then requires only inverse-polynomial inverse-action errors.

The per-level correction factor is included in the fixed D of Eq. (325). Choose $C_* = 2000(D+1)$ after D. Lemma 6.39 explicitly absorbs the resulting depth overhead; it is not treated as one global logarithmic factor. Lemma 7.18 supplies the root primitives and norm certificates used in these calls.

Mask scores, drift coefficients, predictor radii, and certified gains have inverse-polynomial margins by Lemmas 6.27 and 6.31. Choose one finer common grid for these finitely many margins. Pending coefficients are stored without alteration and have zero derivatives. Their true coupling and reservoir remain until bulk attenuation; symbolic coefficients in the fixed contribution keep their original bounded lengths. Lemma 7.5 bounds cache errors afresh against this true model. The charged attenuation uses the same positive-submap estimates, root corrections, and final tiny-slot accuracy as Lemma 7.16. No derivative bound on a pending endpoint is used. Each correction restores the same root-error floor, and no long original input is read again. Bad trials obey the prescribed caps; successful sampling events have slack for all the stated enclosures. This proves the precision and cost claims. □

8 Shared tools for signing via Schatten norms

Let $A_1, \dots, A_N$ be nonzero Hermitian positive semidefinite matrices of order D with $\sum_i A_i = I_D$, $\|A_i\| \le \alpha$, and $\operatorname{rank}(A_i) \le r$. Let $s \ge 2$ be an integer.

We replace the trace functional in the rank-one coupling of Song and Yue [SY26a] by a Schatten-norm functional.

8.1 The Schatten-norm functional and its differential inequalities

For $X \succ 0$, define

$$f_i(X) := (\mathrm{tr}[(A_i^{1/2} X A_i^{1/2})^s])^{1/s},$$
$$C_i := A_i^{1/2} X A_i^{1/2}, \qquad B_i(X) := \frac{A_i^{1/2} C_i^{s-1} A_i^{1/2}}{f_i(X)^{s-1}}.$$

Powers and inverses of C_i are restricted to the support of A_i. The differential identities below express the derivatives of f_i through B_i and reduce the Hessian estimate to a scalar inequality.

Lemma 8.1 (Identities for the Schatten-norm functional)**.** *The function f_i is smooth on the positive definite cone, convex, and homogeneous of degree one. For every Hermitian H,*

$$Df_i(X)[H] = \mathrm{tr}[B_i(X)H], \quad 0 \preceq B_i(X) \preceq A_i, \quad \mathrm{tr}[XB_i(X)] = f_i(X), \quad f_i(X) \geq \|C_i\|. \tag{335}$$

Define $C := C_i$, $K := A_i^{1/2} H A_i^{1/2}$, and $f := (\mathrm{tr}[C^s])^{1/s}$. The Hessian of f as a function of C satisfies

$$D^2 f[K,K] = f^{1-s} \sum_{\ell=0}^{s-2} \mathrm{tr}[C^\ell K C^{s-2-\ell} K] - (s-1) f^{1-2s} (\mathrm{tr}[C^{s-1}K])^2. \tag{336}$$

For positive scalars a, b,

$$\sum_{\ell=0}^{s-2} a^\ell b^{s-2-\ell} \leq \frac{s-1}{2}(a^{s-2} + b^{s-2}) \leq \frac{s-1}{2}\left(\frac{a^{s-1}}{b} + \frac{b^{s-1}}{a}\right). \tag{337}$$

The support inverse satisfies

$$A_i^{1/2} C^{-1} A_i^{1/2} \preceq X^{-1}. \tag{338}$$

Proof. The functional f_i is the Schatten s-norm of C_i, which gives convexity and homogeneity. Since C_i is positive definite on its support, differentiation of $\mathrm{tr}[C_i^s]$ gives smoothness, the derivative formula in Eq. (335), and Eq. (336). The inequality $\|C_i\| \leq f_i(X)$ implies $0 \preceq C_i^{s-1} \preceq f_i(X)^{s-1} I$, hence $0 \preceq B_i(X) \preceq A_i$. Taking the trace in the definition of B_i gives $\mathrm{tr}[XB_i(X)] = f_i(X)$.

For $s = 2$, the first step in Eq. (337) is equality. For $s > 2$, it follows by applying weighted arithmetic–geometric mean to each summand and summing the weights. The second step follows from

$$\frac{a^{s-1}}{b} + \frac{b^{s-1}}{a} - a^{s-2} - b^{s-2} = (a-b)(a^{s-2}/b - b^{s-2}/a) \geq 0.$$

Finally, conjugating the left side of Eq. (338) by $X^{1/2}$ gives an orthogonal projection. Conjugating its upper bound by $X^{-1/2}$ proves Eq. (338). □

The preceding identities yield the differential inequalities used in the potential analysis. Keeping the squared derivative in the Hessian bound will sharpen the mixed-term estimate.

Lemma 8.2 (Differential inequalities for the Schatten-norm functional)**.** *For $X \succ 0$ and Hermitian H,*

$$(Df_i(X)[H])^2 \leq f_i(X)\,\mathrm{tr}[B_i(X) H X^{-1} H], \tag{339}$$
$$0 \leq D^2 f_i(X)[H,H] \leq (s-1)\,\mathrm{tr}[B_i(X) H X^{-1} H]. \tag{340}$$
$$D^2 f_i(X)[H,H] \leq (s-1)\left(\mathrm{tr}[B_i(X) H X^{-1} H] - \frac{(Df_i(X)[H])^2}{f_i(X)}\right). \tag{341}$$

For every $W \succeq 0$,

$$\mathrm{tr}[W A_i X A_i] \leq f_i(X)\,\mathrm{tr}[W A_i] \qquad (W \succeq 0). \tag{342}$$

Proof. Weighted Cauchy–Schwarz, with $B_i(X) \succeq 0$ and $\operatorname{tr}[XB_i(X)] = f_i(X)$, proves Eq. (339). For the Hessian, use C, K, f from Lemma 8.1. Define $R := (s-1)f^{1-2s}(\operatorname{tr}[C^{s-1}K])^2$. We have

$$\begin{aligned} D^2 f[K,K] + R &= f^{1-s} \sum_{\ell=0}^{s-2} \operatorname{tr}[C^\ell K C^{s-2-\ell} K] \\ &\le (s-1) f^{1-s} \operatorname{tr}[C^{s-1} K C^{-1} K] \\ &= (s-1) \operatorname{tr}[B_i(X) H A_i^{1/2} C^{-1} A_i^{1/2} H] \\ &\le (s-1) \operatorname{tr}[B_i(X) H X^{-1} H]. \end{aligned}$$

The first step is Eq. (336). The second step follows by diagonalizing C and applying Eq. (337) to each pair of eigenvalues. The third step substitutes $K = A_i^{1/2} H A_i^{1/2}$ and the definition of $B_i(X)$. The last step uses Eq. (338). Since $R = (s-1)(Df_i(X)[H])^2 / f_i(X)$, subtraction proves Eq. (341). Dropping $R \ge 0$ gives the upper bound in Eq. (340); convexity gives its lower bound.

To prove Eq. (342), pair $C_i \preceq f_i(X) I$ with $A_i^{1/2} W A_i^{1/2} \succeq 0$ and use cyclicity of the trace. □

8.2 Potential, initialization, and regularity

Set

$$\iota := 16(s+3), \qquad \beta(y) := \sqrt{1-y^2}, \qquad \mathcal{F}_y(X) := \iota \sum_{i=1}^N \beta(y_i) f_i(X) A_i, \qquad M(y) := \sum_{i=1}^N y_i A_i.$$

We call $\mathcal{F}_y$ the Schatten-norm coupling map. For $\rho > 0$, define

$$\begin{aligned} P_\rho(y) := \min_{b \in \mathbb{R},\ U,V \succ 0} \quad & b + \rho \operatorname{tr}[U+V] \\ \text{subject to} \quad & U^{-1} + M(y) + \mathcal{F}_y(V) \preceq bI, \\ & V^{-1} - M(y) + \mathcal{F}_y(U) \preceq bI. \end{aligned} \tag{343}$$

For primal optimizers U, V and dual multipliers $W, Z \succeq 0$, define

$$u_i := f_i(U), \qquad v_i := f_i(V), \qquad w_i := \operatorname{tr}[A_i W], \qquad z_i := \operatorname{tr}[A_i Z], \qquad \beta_i := \beta(y_i).$$

On Hermitian pairs, define the constraint derivative

$$J(H,K) := (U^{-1} H U^{-1} - D\mathcal{F}_y(V)[K],\ V^{-1} K V^{-1} - D\mathcal{F}_y(U)[H]).$$

The regularization gives an attained potential and a smooth response on each open face. Its dual equations also determine the weights in the covariance estimates.

Lemma 8.3 (Potential and regularity)**.** *For $\rho > 0$, the minimum in Eq. (343) is attained and $P_\rho(y) \ge \|M(y)\|$. The value is continuous on $[-1,1]^N$. The primal and dual optimizers are unique and smooth on every open face, as is P_ρ. They satisfy*

$$\begin{aligned} W &= U(\rho I + \iota \sum_{i=1}^N \beta_i z_i B_i(U)) U, \\ Z &= V(\rho I + \iota \sum_{i=1}^N \beta_i w_i B_i(V)) V, \qquad \operatorname{tr}[W+Z] = 1. \end{aligned} \tag{344}$$

Both primal constraints are tight, and $W, Z \succ 0$. At the origin,

$$P_\rho(0) \le 2\sqrt{\iota r^{1/s} \alpha + 2D\rho}. \tag{345}$$

Proof. For fixed y, this is a convex optimization problem: each scalar f_i is convex and is multiplied by a PSD matrix. Strict feasibility follows by fixing positive definite U, V and making b sufficiently large.

Always $P_\rho(y) \geq \|M(y)\|$, since the coupling and inverse terms are PSD. In particular $b \geq 0$. An objective sublevel bounds b, $\mathrm{tr}[U]$, and $\mathrm{tr}[V]$, while the constraints give $U^{-1}, V^{-1} \preceq (b + \|M\|)I$. Thus the minimum is attained. These bounds can be chosen uniformly for y in the cube. Continuity follows from compactness for the lower bound on the value, and, for the upper bound, by increasing b slightly to repair a fixed feasible solution as y changes.

Stationarity of the Lagrangian gives Eq. (344). These equations imply $W, Z \succ 0$, hence both primal constraints are equalities. All four scalars u_i, v_i, w_i, z_i are strictly positive for nonzero active inputs.

To prove smoothness, consider the constraint derivative J. The off-diagonal derivatives of the coupling map are positive by Eq. (335), and Eq. (344) says $J^*(W, Z) = \rho(I, I)$. Factor $J = S(I - L)$, with S the positive invertible diagonal inverse-congruence operator and L positive. Then $L^* S(W, Z) \prec S(W, Z)$. Order-unit comparison for this strictly positive pair implies spectral radius $\varrho(L) < 1$. Consequently J is invertible with positive inverse.

The primal-variable Hessian of the Lagrangian in (U, V) is positive definite: its inverse terms weighted by $W, Z \succ 0$ are strictly convex, and the coupling terms are convex. The primal optimizer is therefore unique. The dual optimizer is unique because J^* is invertible. The full KKT derivative is nonsingular. Indeed a homogeneous perturbation obeys $J\delta X = -\delta b(I, I)$, $\mathcal{H}\delta X - J^*\delta Q = 0$, and $\mathrm{tr}[\delta W + \delta Z] = 0$. Pairing the second equation with δX shows $\langle \delta X, \mathcal{H}\delta X \rangle = 0$, hence all perturbations vanish. The implicit function theorem proves smoothness of the optimizer and P_ρ on every open face.

At $y = 0$, take $U := V := \ell I$. Since

$$f_i(I) = \|A_i\|_{S_s} \leq r^{1/s}\alpha, \qquad \sum_{i=1}^{N} f_i(I) A_i \preceq r^{1/s}\alpha I,$$

optimizing over $\ell > 0$ proves Eq. (345).

□

8.3 Endpoints and light states

Let $I \subseteq [N]$ index the active coordinates of an open face. The vanishing endpoint budget permits a coordinate to be fixed whenever one coupling term absorbs the change in $M(y)$.

Lemma 8.4 (Endpoint updates)**.** *For $i \in I$, setting $y_i := 1$ does not increase P_ρ if*

$$\iota \beta_i v_i \geq 1 - y_i, \tag{346}$$

Setting $y_i := -1$ does not increase P_ρ if

$$\iota \beta_i u_i \geq 1 + y_i. \tag{347}$$

Proof. For the update $y_i := 1$, the first primal left side changes by $(1 - y_i - \iota\beta_i v_i)A_i \preceq 0$. The second changes by $(-(1 - y_i) - \iota\beta_i u_i)A_i \preceq 0$. Thus the current b, U, V remain feasible with the same objective. The update $y_i := -1$ follows by interchanging the two constraints. □

The remaining states satisfy strict inequalities in every active coordinate.

Definition 8.5 (Light state)**.** *A state on an open face is light if, for every $i \in I$,*

$$\iota\beta_i v_i < 1 - y_i, \qquad \iota\beta_i u_i < 1 + y_i. \tag{348}$$

At a light state, define

$$p_i := 1 + \iota\beta_i' v_i, \quad q_i := 1 - \iota\beta_i' u_i, \quad c_i := \sqrt{\beta_i p_i / q_i}, \quad d_i := \sqrt{\beta_i q_i / p_i}, \quad \tau_i := \sqrt{p_i q_i / \beta_i}.$$

Lightness ensures that the response normalization is positive. It also gives the coercivity bound used to construct a covariance.

Lemma 8.6 (Light-state weights)**.** *For $t := |y_i| < 1$, the square-root budget satisfies*

$$p_i, q_i > 0, \qquad p_i q_i \le \frac{1}{1-t}, \qquad \Gamma_i := \frac{1}{\iota\beta_i u_i v_i} > \iota. \tag{349}$$

The potential gradient is

$$\partial_i P_\rho = p_i w_i - q_i z_i. \tag{350}$$

Proof. For $y_i \ge 0$, substitute $\beta_i' = -y_i/\beta_i$ in the light-state inequalities to obtain $1/(1+y_i) < p_i \le 1$ and $1 \le q_i < 1/(1-y_i)$. For $y_i < 0$, interchange the roles of p_i and q_i. These bounds give positivity and the product bound in Eq. (349). Multiplying the two light-state inequalities gives $\iota^2\beta_i^2 u_i v_i < 1 - y_i^2$. Since $\beta_i \ge 1 - y_i^2$, this implies $\Gamma_i > \iota$. Finally, differentiating the value through the dual multipliers gives Eq. (350). □

8.4 A parameterized projection inequality

The following lemma adapts the weighted projection argument of Song and Yue [SY26a], whose original lemma and proof are reproduced as Lemma 4.2. For $\vartheta > 1$, define

$$\gamma := \vartheta(\vartheta + 1), \qquad \kappa := \frac{\vartheta^2(\vartheta+1)}{\vartheta^3 - 1}. \tag{351}$$

Define

$$w_e := \frac{\vartheta^4}{(\vartheta+1)(\vartheta^3+\vartheta+1)}, \qquad w_f := \frac{\vartheta}{(\vartheta+1)(2\vartheta+1)},$$
$$h := \frac{2\vartheta^3}{\vartheta^3+\vartheta+1}, \qquad j := \frac{2\vartheta}{2\vartheta+1}.$$

Taking $\vartheta = 2$ gives $\gamma = 6$ and $\kappa = 12/7$. The parameterized form below will also be used with other coercivity bounds.

Lemma 8.7 (Parameterized projection inequality)**.** *Let $e_1, \dots, e_k, f_1, \dots, f_k$ be an orthonormal basis of $\mathbb{R}^{2k}$. Let Π_0 project onto the e-span and Π be any rank-k orthogonal projection. Suppose a real matrix T and a PSD matrix D_0 satisfy*

$$(I - \Pi_0)(I - T)\Pi = 0, \qquad e_i^\top T e_i + f_i^\top T f_i = 0, \qquad e_i^\top T e_i \le 0,$$
$$e_i^\top D_0 e_i = f_i^\top D_0 f_i = 1, \qquad D_0 \succeq \gamma T^\top T.$$

There is a PSD covariance $\mathcal{A}$, positive definite on range Π *and zero on its orthogonal complement, such that*

$$\operatorname{tr}[D_0 \mathcal{A}] \le \kappa \sum_{i=1}^{k} e_i^\top (I - T)\mathcal{A}(I - T)^\top e_i, \tag{352}$$

and the sum on the right is positive. One such covariance is defined by

$$D := \sum_{i=1}^{k} (w_e e_i e_i^\top + w_f f_i f_i^\top), \quad \mathcal{A} := R(R^\top D^{-1} R)^{-1} R^\top, \quad \mathcal{B} := D - \mathcal{A},$$

where the columns of R are an orthonormal basis for range Π. *It satisfies*

$$\mathcal{A}, \mathcal{B} \succeq 0, \quad \mathcal{A} D^{-1} \mathcal{A} = \mathcal{A}, \quad \mathcal{A} D^{-1} \mathcal{B} = 0, \quad \operatorname{tr}[D^{-1}\mathcal{A}] = k. \tag{353}$$

Proof. The formula for $\mathcal{A}$ gives a covariance supported on range Π, positive definite there. Conjugating by $D^{-1/2}$ gives an orthogonal projection, which proves Eq. (353). Direct substitution of the weights gives

$$\frac{h}{w_e} - \frac{h^2}{4w_e^2}\left(\frac{1}{\kappa} - \frac{1}{\gamma}\right) = \frac{1}{w_e}, \qquad \frac{j}{w_f} + \frac{j^2}{4\gamma w_f^2} = \frac{1}{w_f},$$
$$\frac{h^2}{4\gamma w_e} + \frac{j^2}{4\gamma w_f} = 1 - w_e - w_f. \tag{354}$$

Also $h > j$, since the numerator of their difference has the sign of $(\vartheta - 1)(\vartheta + 1)^2$.

Put $C := \gamma\mathcal{B} + \kappa\mathcal{A} \succ 0$. The projection identities give

$$C^{-1} = \gamma^{-1} D^{-1} + (\kappa^{-1} - \gamma^{-1}) D^{-1} \mathcal{A} D^{-1}.$$

For each row $t := e_i^\top T$, completing a square with C and using Eq. (354) gives

$$\gamma t \mathcal{B} t^\top + \kappa (e_i^\top - t)\mathcal{A}(e_i - t^\top) + h\, t e_i \geq \frac{e_i^\top \mathcal{A} e_i}{w_e} - \frac{h^2}{4\gamma w_e}. \tag{355}$$

Indeed, the minimum of the left side over t is

$$\kappa e_i^\top \mathcal{A} e_i - (\kappa \mathcal{A} e_i - h e_i/2)^\top C^{-1} (\kappa \mathcal{A} e_i - h e_i/2),$$

which is the right side by the first identity in Eq. (354).

For the row $t := f_i^\top T$, the response condition implies $t\mathcal{A} = f_i^\top \mathcal{A}$, hence $w_f t f_i = f_i^\top \mathcal{A} f_i + t\mathcal{B} f_i$. Completing a square with the PSD matrix $\mathcal{B}$, without inverting it, gives

$$\begin{aligned} \gamma t \mathcal{B} t^\top + j\, t f_i &\geq \frac{j}{w_f} f_i^\top \mathcal{A} f_i - \frac{j^2}{4\gamma w_f^2} f_i^\top \mathcal{B} f_i \\ &= \frac{j}{w_f} f_i^\top \mathcal{A} f_i - \frac{j^2}{4\gamma w_f^2}(w_f - f_i^\top \mathcal{A} f_i) \\ &= \frac{f_i^\top \mathcal{A} f_i}{w_f} - \frac{j^2}{4\gamma w_f}. \end{aligned} \tag{356}$$

The first step completes the square in $\mathcal{B}$ using $w_f t f_i = f_i^\top \mathcal{A} f_i + t\mathcal{B} f_i$. The second step substitutes $\mathcal{B} = D - \mathcal{A}$. The last step uses the second identity in Eq. (354). Sum Eqs. (355) and (356) over i. The right sides add to $k(w_e + w_f) = \operatorname{tr}[D_0 D]$, using Eq. (354). The extra diagonal term on the left is $\sum_{i=1}^{k} (h - j) e_i^\top T e_i \leq 0$. It follows that

$$\gamma \operatorname{tr}[T\mathcal{B}T^\top] + \kappa \sum_{i=1}^{k} e_i^\top (I - T)\mathcal{A}(I - T)^\top e_i \geq \operatorname{tr}[D_0 D].$$

Since $D_0 \succeq \gamma T^\top T$, $\operatorname{tr}[D_0 \mathcal{A}] \leq \operatorname{tr}[D_0 D] - \gamma \operatorname{tr}[T\mathcal{B}T^\top]$, proving Eq. (352).

If the movement vanished, the response condition would imply $(I - T)\Pi = 0$. Then T would act as the identity on a k-dimensional subspace, so $\|T\|_F^2 \geq k$. But $\gamma \|T\|_F^2 \leq \operatorname{tr}[D_0] = 2k$, contradicting $\gamma > 2$. This proves positive movement. □

8.5 Response covariances

Let h be an active-coordinate direction and keep b fixed while solving the tight primal constraints locally. The inverse of J makes this response well defined. Define the response coefficients by

$$\dot{U} = UM_aU, \quad \dot{V} = VM_bV, \qquad M_a := \sum_{i\in I} c_i a_i A_i, \quad M_b := \sum_{i\in I} d_i b_i A_i.$$

One may choose the coefficients as the unique solution of

$$a - \mathsf{V}b = \mathsf{U}a - b = \tau h, \tag{357}$$

where

$$\mathsf{U}_{ij} := \iota c_i c_j \operatorname{tr}[B_i(U)UA_jU], \qquad \mathsf{V}_{ij} := \iota d_i d_j \operatorname{tr}[B_i(V)VA_jV]. \tag{358}$$

These matrices have nonnegative entries. The dual equations give the following transpose inequalities:

$$\mathsf{U}^\top (d_i z_i)_i < (c_i w_i)_i, \qquad \mathsf{V}^\top (c_i w_i)_i < (d_i z_i)_i. \tag{359}$$

The positive remainder in the first inequality is $\rho c_j \operatorname{tr}[A_jU^2]$ in coordinate j. The product has spectral radius less than one by positive-vector comparison; hence Eq. (357) has a unique solution for every h. Its graph has dimension equal to the active-coordinate count k, even if the matrices A_i are linearly dependent.

Set

$$Q := \operatorname{tr}[WM_aUM_a] + \operatorname{tr}[ZM_bVM_b],$$
$$E_a := \operatorname{diag}(c_i^2 u_i w_i), \quad E_b := \operatorname{diag}(d_i^2 v_i z_i), \quad E := \operatorname{diag}(E_a, E_b).$$

Let $M_{\mathrm{raw}} \succeq 0$ be the real symmetric matrix of this quadratic form in (a, b). It is block diagonal. Eq. (342) gives $\operatorname{diag}(M_{\mathrm{raw}}) \le \operatorname{diag}(E)$. Add the nonnegative diagonal difference Δ and define

$$\widehat{D} := E^{-1/2}(M_{\mathrm{raw}} + \Delta)E^{-1/2}, \qquad \widehat{T} := \begin{pmatrix} 0 & E_a^{1/2}\mathsf{V}E_b^{-1/2} \\ E_b^{1/2}\mathsf{U}E_a^{-1/2} & 0 \end{pmatrix}.$$

Thus $\widehat{D}$ has unit diagonal and no cross-block entries.

The derivative bound for f_i controls the response energy. Define $Q_U := \operatorname{tr}[WM_aUM_a]$. We have

$$\begin{aligned} Q_U &\ge \iota \sum_{i\in I} \beta_i z_i \operatorname{tr}[B_i(U)UM_aUM_aU] \\ &\ge \iota \sum_{i\in I} \frac{\beta_i z_i}{u_i} (\operatorname{tr}[B_i(U)UM_aU])^2 \\ &= \iota \sum_{i\in I} \frac{\beta_i z_i}{u_i} \Big(\frac{(\mathsf{U}a)_i}{\iota c_i}\Big)^2 \\ &= \sum_{i\in I} \Gamma_i (E_b)_{ii} (\mathsf{U}a)_i^2. \end{aligned} \tag{360}$$

The first step uses the dual equation for W in Eq. (344). The second step applies Eq. (339) with $H = UM_aU$. The third step substitutes the definition of U from Eq. (358). The last step uses $c_i d_i = \beta_i$ and the definitions of Γ_i and E_b. The corresponding inequality for V gives

$$\widehat{D} \succeq E^{-1/2} M_{\mathrm{raw}} E^{-1/2} \succeq \widehat{T}^\top \Gamma \widehat{T}, \tag{361}$$

Here Γ has value Γ_i on both coordinates of pair i. The first step uses the nonnegative padding Δ. The second step combines Eq. (360) with its V counterpart.

In each pair use the orthonormal vectors

$$e_i := \frac{\sqrt{(E_a)_{ii}}e_{a,i} - \sqrt{(E_b)_{ii}}e_{b,i}}{\sqrt{(E_a)_{ii} + (E_b)_{ii}}}, \qquad f_i := \frac{\sqrt{(E_b)_{ii}}e_{a,i} + \sqrt{(E_a)_{ii}}e_{b,i}}{\sqrt{(E_a)_{ii} + (E_b)_{ii}}}.$$

Let Π project onto the normalized response graph $E^{1/2}(a,b)$, and Π_0 onto the span of the e_i. Eq. (357) gives $(I-\Pi_0)(I-\widehat{T})\Pi = 0$. Also

$$\langle e_i, \widehat{D}e_i\rangle = \langle f_i, \widehat{D}f_i\rangle = 1,$$

$$\langle e_i, \widehat{T}e_i\rangle = -\frac{(E_a)_{ii}\mathsf{V}_{ii} + (E_b)_{ii}\mathsf{U}_{ii}}{(E_a)_{ii} + (E_b)_{ii}} \le 0, \quad \langle f_i, \widehat{T}f_i\rangle = -\langle e_i, \widehat{T}e_i\rangle. \tag{362}$$

Eqs. (349), (361), and (362) verify the hypotheses of Lemma 8.7 with $\gamma = 6$, since $\Gamma_i > \iota \ge 80$. Define $S_i := v_i w_i + u_i z_i > 0$. The projection covariance transfers to active-coordinate directions through the response normalization.

Lemma 8.8 (Response covariance)**.** *At a light state, there is a covariance of real active-coordinate directions h with $\sum_{i\in I}\mathbb{E}[h_i^2] > 0$ such that*

$$\mathbb{E}[Q] \le \frac{12}{7}\sum_{i\in I}(p_i^2 u_i w_i + q_i^2 v_i z_i)\mathbb{E}[h_i^2]. \tag{363}$$

At a stationary light state, the same covariance satisfies

$$\mathbb{E}[Q] \le 2\sum_{i\in I} p_i q_i S_i \mathbb{E}[h_i^2]. \tag{364}$$

Proof. Apply Lemma 8.7 with $\vartheta = 2$. The normalized e_i movement is $\tau_i\sqrt{(E_a)_{ii} + (E_b)_{ii}}h_i$. Its square satisfies

$$\tau_i^2((E_a)_{ii} + (E_b)_{ii})h_i^2 = (p_i^2 u_i w_i + q_i^2 v_i z_i)h_i^2.$$

The diagonal padding increases the comparison energy, so Eq. (352) implies Eq. (363). Positive movement gives $\sum_{i\in I}\mathbb{E}[h_i^2] > 0$.

At a stationary point, Eq. (350) gives $p_i w_i = q_i z_i$. Consequently $p_i^2 u_i w_i + q_i^2 v_i z_i = p_i q_i S_i$. Substitution in Eq. (363), followed by $12/7 < 2$, proves Eq. (364). □

8.6 The nonlinear Hessian cost and strict negative curvature

The fixed-b feasible response curve gives an upper bound on the second derivative of the optimized value, because its objective agrees with P_ρ at the base point and lies above it nearby. Differentiating its equalities twice and using Eq. (344) gives

$$\begin{aligned} D^2 P_\rho[h,h] \le 2Q &+ 2\iota\sum_{i\in I}\beta_i' h_i(w_i\dot{v}_i + z_i\dot{u}_i) + \iota\sum_{i\in I}\beta_i'' S_i h_i^2 \\ &+ \iota\sum_{i\in I}\beta_i[w_i D^2 f_i(V)[\dot{V},\dot{V}] + z_i D^2 f_i(U)[\dot{U},\dot{U}]]. \end{aligned} \tag{365}$$

The last line is the nonlinear functional-Hessian contribution. Eq. (340) and the dual equations bound it by $(s-1)Q$. Eq. (339) and the same dual equations bound the absolute value of the mixed term by $2\sqrt{Q\mathcal{M}}$, where

$$\mathcal{M} := \iota \sum_{i\in I} \frac{(\beta_i')^2}{\beta_i} S_i h_i^2.$$

Applying $2\sqrt{Q\mathcal{M}} \le 2Q + \mathcal{M}/2$ gives

$$D^2 P_\rho[h,h] \le (s+3)Q - \iota \sum_{i\in I} L(y_i) S_i h_i^2, \quad L(y) := -\beta''(y) - \frac{\beta'(y)^2}{2\beta(y)}. \tag{366}$$

For the square-root budget, the negative scalar term dominates the averaged response energy. This excludes stationary light states from the set of local minima.

Lemma 8.9 (Negative curvature at light states)**.** *At a light state with $\beta(y) = \sqrt{1-y^2}$,*

$$L(y_i) \ge \frac{1}{4} p_i q_i. \tag{367}$$

At a stationary light state, the covariance of Lemma 8.8 satisfies

$$\mathbb{E}[D^2 P_\rho[h,h]] \le (2(s+3) - \frac{\iota}{4}) \sum_{i\in I} p_i q_i S_i \mathbb{E}[h_i^2] = -2(s+3) \sum_{i\in I} p_i q_i S_i \mathbb{E}[h_i^2] < 0. \tag{368}$$

In particular, such a state is not a local minimum on its open face.

Proof. For the square-root budget,

$$L(y) = \frac{1 - y^2/2}{(1-y^2)^{3/2}}, \qquad L(y)(1-t) = \frac{1 - t^2/2}{(1+t)\sqrt{1-t^2}} \ge \frac{1}{4}, \qquad t := |y|.$$

Together with Eq. (349), this proves Eq. (367). Average Eq. (366) over the covariance in Eq. (364) and apply Eq. (367). Substituting $\iota = 16(s+3)$ gives Eq. (368). The strict inequality follows from $p_i q_i S_i > 0$ and $\sum_{i\in I} \mathbb{E}[h_i^2] > 0$. □

8.7 Hessian energy of the Schatten-norm functional and mixed terms

For any positive smooth budget on an open face, the fixed-b response in Eq. (365) remains valid. Define

$$\mathcal{R} := \iota \sum_{i\in I} \beta_i (w_i \frac{\dot{v}_i^2}{v_i} + z_i \frac{\dot{u}_i^2}{u_i}), \qquad \mathcal{M} := \iota \sum_{i\in I} \frac{(\beta_i')^2}{\beta_i} S_i h_i^2.$$

Retaining the squared derivative in Lemma 8.2 allows the mixed term to be absorbed into the functional-Hessian energy.

Lemma 8.10 (Refined curvature and covariance)**.** *For a positive smooth budget on an open face, the fixed-b response satisfies*

$$D^2 P_\rho[h,h] \le (s+1)Q - \iota \sum_{i\in I} L_s(y_i) S_i h_i^2, \qquad L_s := -\beta'' - \frac{(\beta')^2}{(s-1)\beta}. \tag{369}$$

At a stationary light state whose normalized response satisfies Lemma 8.7 with parameters γ, κ, there is a nonzero covariance satisfying

$$\mathbb{E}[Q] \le \kappa \sum_{i \in I} p_i q_i S_i \mathbb{E}[h_i^2]. \tag{370}$$

If $L_s(y_i) \ge \lambda p_i q_i$ for all active coordinates and $\iota\lambda > \kappa(s+1)$, this covariance gives strictly negative expected curvature.

Proof. Eq. (341) and the dual equations bound the functional-Hessian contribution by $(s-1)(Q - \mathcal{R})$. Weighted Cauchy–Schwarz bounds the absolute mixed term by $2\sqrt{\mathcal{R}\mathcal{M}}$. Applying

$$2\sqrt{\mathcal{R}\mathcal{M}} \le (s-1)\mathcal{R} + \frac{\mathcal{M}}{s-1}$$

in Eq. (365) proves Eq. (369). All denominators are positive on the open face.

At stationarity, Eq. (350) gives $p_i w_i = q_i z_i$. The normalization in Lemma 8.8 then converts Eq. (352) to Eq. (370). Averaging the curvature bound and using $L_s(y_i) \ge \lambda p_i q_i$ gives

$$\mathbb{E}[D^2 P_\rho[h,h]] \le (\kappa(s+1) - \iota\lambda) \sum_{i \in I} p_i q_i S_i \mathbb{E}[h_i^2] < 0.$$

□

8.8 Coordinate-dependent covariance and budget certificates

We retain both the coordinate-dependent coercivity and part of the negative functional-Hessian term. The retained fraction may depend on the coordinate. The resulting projection estimate gives the intermediate coefficient 4.335. Section 11.4 improves this to 4.225, and Section 12 supplies the quantitative curvature and reservoir margins for a deterministic construction with that coefficient.

Retaining response energy. Set $c := s+1$ and choose $0 \le \mu_i < (s-1)/(s+1)$, with $\eta_i := s - 1 - c\mu_i > 0$. Define

$$E_i := \iota\beta_i(w_i \frac{\dot{v}_i^2}{v_i} + z_i \frac{\dot{u}_i^2}{u_i}), \qquad L_i := -\beta_i'' - \frac{(\beta_i')^2}{\eta_i \beta_i}.$$

Retain the negative term in Eq. (341). Completing the two mixed squares using only $\eta_i E_i$ gives

$$D^2 P_\rho[h,h] \le c(Q - \sum_i \mu_i E_i) - \iota \sum_i L_i S_i h_i^2. \tag{371}$$

Indeed the mixed term in coordinate i is at most $\eta_i E_i + \iota(\beta_i')^2 S_i h_i^2/(\eta_i \beta_i)$. These are pointwise algebraic choices, so no derivative of μ_i or η_i enters. For the normalized response vector $v = E^{1/2}(a,b)$, Eq. (357) gives

$$E_i = \Gamma_i((e_i^\top \widehat{T} v)^2 + (f_i^\top \widehat{T} v)^2). \tag{372}$$

For example $\dot{u}_i = (\mathsf{U}a)_i/(\iota c_i)$; substitution of $(E_b)_{ii} = d_i^2 v_i z_i$ gives its contribution to this identity. The rotation within each pair preserves the sum of squares.

Projection weights. The rowwise square completion uses weights that depend on the coercivity and retained-energy parameters. We record their identities separately because they are also used in Section 11.

For each $\gamma_i > 2$ and $\mu_i \geq 0$, define $\vartheta_i > 1$ and κ_i by

$$\gamma_i = \frac{\vartheta_i(\vartheta_i + 1)}{1 + \mu_i \vartheta_i}, \qquad \kappa_i = \frac{\vartheta_i + 1}{\gamma_i + 1}\left(\frac{1}{\vartheta_i - 1} + \frac{\gamma_i}{\vartheta_i}\right). \tag{373}$$

Lemma 8.11 (Projection weight identities)**.** *Let $\mu \geq 0$ and $\gamma > 2$. Define $\vartheta > 1$ and κ by Eq. (373), suppressing the coordinate index. Define*

$$a := \frac{2(\vartheta + 1)}{\vartheta + \mu}, \qquad b := \frac{2\gamma}{\vartheta},$$
$$w_e := \frac{\vartheta}{(\vartheta + 1)(1 + a^2/(4\gamma))}, \qquad w_f := \frac{1}{(\vartheta + 1)(1 + b^2/(4\gamma))}.$$
$$h := w_e a, \qquad j := w_f b, \qquad d := \mu\gamma + \kappa.$$

These weights satisfy

$$\begin{aligned} \kappa - \frac{\kappa^2}{d} + \frac{\kappa h}{d w_e} - \frac{h^2}{4 w_e^2}\left(\frac{1}{d} - \frac{1}{\gamma}\right) &= \frac{1}{w_e}, \\ \mu\gamma + \frac{j}{w_f} + \frac{j^2}{4\gamma w_f^2} &= \frac{1}{w_f}, \\ \frac{h^2}{4\gamma w_e} + \frac{j^2}{4\gamma w_f} &= 1 - w_e - w_f. \end{aligned} \tag{374}$$

Moreover, $h > j$.

Proof. Direct substitution proves Eq. (374). After multiplication by positive factors, $h - j$ has the sign of

$$(\vartheta - 1)(\mu\vartheta^3 + (1 + \mu + \mu^2)\vartheta^2 + (2 + \mu^2)\vartheta + 1) > 0.$$

□

A projection inequality retaining response energy. The weights in Lemma 8.11 allow a different coercivity parameter in each coordinate pair.

Lemma 8.12 (Projection with retained response energy)**.** *Use the hypotheses of Lemma 8.7, replacing its coercivity inequality by*

$$D_0 \succeq T^\top \Gamma T, \qquad \Gamma := \sum_{i=1}^{k} \gamma_i (e_i e_i^\top + f_i f_i^\top), \qquad \gamma_i > 2.$$

Let $\mu_i \geq 0$ and define ϑ_i, κ_i by Eq. (373). There is a PSD covariance $\mathcal{A}$, positive definite on the response graph and zero on its orthogonal complement, such that

$$\begin{aligned} &\operatorname{tr}[D_0 \mathcal{A}] - \sum_i \mu_i \gamma_i (e_i^\top T \mathcal{A} T^\top e_i + f_i^\top T \mathcal{A} T^\top f_i) \\ &\qquad \leq \sum_i \kappa_i e_i^\top (I - T)\mathcal{A}(I - T)^\top e_i. \end{aligned} \tag{375}$$

The sum of the unweighted movement terms is positive.

Proof. Use the weights from Lemma 8.11 in each pair and define

$$D := \sum_i (w_{e,i} e_i e_i^\top + w_{f,i} f_i f_i^\top), \qquad \mathcal{A} := R(R^\top D^{-1} R)^{-1} R^\top, \qquad \mathcal{B} := D - \mathcal{A}.$$

Here the columns of R form an orthonormal basis of range Π. The construction satisfies Eq. (353). For the e_i row complete the square with

$$C_i := \gamma_i \mathcal{B} + d_i \mathcal{A}, \qquad C_i^{-1} = \gamma_i^{-1} D^{-1} + (d_i^{-1} - \gamma_i^{-1}) D^{-1} \mathcal{A} D^{-1}.$$

For $t = e_i^\top T$, Eq. (374) gives

$$\gamma_i t \mathcal{B} t^\top + \mu_i \gamma_i t \mathcal{A} t^\top + \kappa_i (e_i^\top - t) \mathcal{A} (e_i - t^\top) + h_i t e_i \geq \frac{e_i^\top \mathcal{A} e_i}{w_{e,i}} - \frac{h_i^2}{4 \gamma_i w_{e,i}}.$$

For $t = f_i^\top T$, the graph identity gives $t\mathcal{A} = f_i^\top \mathcal{A}$. The other square completion gives

$$\gamma_i t \mathcal{B} t^\top + \mu_i \gamma_i t \mathcal{A} t^\top + j_i t f_i \geq \frac{f_i^\top \mathcal{A} f_i}{w_{f,i}} - \frac{j_i^2}{4 \gamma_i w_{f,i}}.$$

On summation, the right side is $\operatorname{tr}[D^{-1}\mathcal{A}] - \sum_i (1 - w_{e,i} - w_{f,i}) = \operatorname{tr}[D_0 D]$. The extra diagonal term on the left is $\sum_i (h_i - j_i) e_i^\top T e_i \leq 0$. Using $\operatorname{tr}[D_0 \mathcal{B}] \geq \operatorname{tr}[\Gamma T \mathcal{B} T^\top]$ and subtracting proves Eq. (375). If all movement vanished, T would be the identity on the rank-k response graph. This would contradict $\|T\|_F^2 \geq k$ and $(\min_i \gamma_i)\|T\|_F^2 \leq \operatorname{tr}[D_0] = 2k$. □

Retaining the light-state correlation. Apply Eq. (375) with the actual $\gamma_i = \Gamma_i = 1/(\iota \beta_i u_i v_i)$. The movement normalization and stationarity $p_i w_i = q_i z_i$ give

$$\mathbb{E}[Q - \sum_i \mu_i E_i] \leq \sum_i \kappa(\Gamma_i, \mu_i) p_i q_i S_i \mathbb{E}[h_i^2]. \tag{376}$$

It therefore suffices to prove

$$\iota L_i > (s+1) \kappa(\Gamma_i, \mu_i) p_i q_i. \tag{377}$$

Define $t := |y|$, $z := 1 - t^2$, and $\beta(t) := \sqrt{z} f(t)$. Set

$$\begin{aligned} N &:= z f' - t f, \qquad A := f - z f', \qquad \eta := s - 1 - (s+1)\mu(t), \\ \widetilde{H}_s &:= -\eta f (z N' + t N) - N^2. \end{aligned} \tag{378}$$

Then $\beta' = N/\sqrt{z}$ and $L = \widetilde{H}_s / (\eta z^{3/2} f)$. We verify

$$f > 0, \quad N \leq 0, \quad f^2 \geq z, \quad f + z f' > 0, \quad A > 0, \quad \eta > 0, \quad \widetilde{H}_s > 0. \tag{379}$$

These imply the required budget and endpoint-slope conditions. Normalize lightness by

$$a := \frac{\iota \beta u}{1+t}, \qquad b := \frac{\iota \beta v}{1-t}, \qquad \ell := -\beta'/\beta, \qquad 0 < a, b < 1.$$

Then $p = 1 - \ell(1-t)b$, $q = 1 + \ell(1+t)a$, and $\Gamma = \iota f / (\sqrt{z}\, ab)$. For fixed t, b, both pq and $\kappa(\Gamma, \mu)$ increase with a. Define $x := 1/\vartheta$. Eq. (373) gives

$$\gamma = \frac{1+x}{x(x+\mu)}, \qquad \kappa(\gamma, \mu) = \frac{(1+x)(1+\mu x)}{(1-x)(1+(1+\mu)x + x^2)}. \tag{380}$$

The first expression decreases with x, while the derivative numerator of the second is $1+4\mu x+(3+\mu+2\mu^2)x^2+(2+2\mu)x^3+\mu x^4>0$. Set $a=1$ for an upper bound, retaining b. At this comparison point,

$$\gamma=\frac{\iota f}{\sqrt{z}\,b},\qquad q=\frac{A}{(1-t)f},\qquad p=1+\frac{\iota N}{(1+t)\sqrt{z}\,\gamma}.$$

Define $D(t,x):=(1-x)(1+(1+\mu)x+x^2)$ and

$$\begin{aligned}W_s(t,x)&:=\iota[\widetilde{H}_sD-(s+1)\eta AN x(x+\mu)(1+\mu x)],\\ R_s(t,x)&:=(s+1)\eta A(1+x)(1+\mu x)\sqrt{1-t^2}(1+t).\end{aligned}\tag{381}$$

The sufficient curvature condition is $W_s>R_s$. For the order-three parameters below, $\gamma\ge\iota>6$ and $\mu\ge0$ imply $x<1/2$. The further restriction $b\le1$ gives

$$G_s(t,x):=\iota^2f^2x^2(x+\mu)^2-(1+x)^2(1-t^2)\le0.\tag{382}$$

Both sides of the desired inequality are nonnegative and $W_s>0$. Thus it is enough to prove

$$P_s(t,x):=W_s(t,x)^2-R_s(t,x)^2>0\tag{383}$$

on $[0,1]\times[0,1/2]$ subject to $G_s\le0$.

Schatten order three. Take

$$\begin{aligned}s=3,\qquad \iota=\frac{123}{20},\qquad \mu(t):=\frac{499}{1000}(1-t)^2,\qquad \eta(t):=2-4\mu(t),\\ f(t):=1+0.0512t^2-0.1198t^4+0.4794t^6\\ -1.2091t^8+1.4060t^{10}-0.6785t^{12}.\end{aligned}\tag{384}$$

All decimals denote exact rational numbers. The even budget uses $t=|y|$. The auxiliary split $\mu(|y|)$ is only pointwise and need not be differentiated at zero. The minimum ordinary-degree Bernstein coefficients on $[0,1]$ are

polynomial	minimum coefficient
f	2323/2500
$-N/t$	561/625
$(f^2-(1-t^2))/t^2$	5396329/6250000
$f+(1-t^2)f'$	27061/32500
A	90247/97500
η	1/250
$\widetilde{H}_3$	561/156250

The polynomial quotients are extended continuously at zero. These positive coefficients prove Eq. (379). In particular $\beta(0)=1$, $0<\beta\le1$ on the open interval, $\beta(\pm1)=0$, and $\beta\ge1-y^2$.

The remaining condition is certified on the whole admissible domain. For $g(t)=\sum_{j=0}^d g_jt^j$, its degree-d Bernstein coefficient numbered i is $\sum_{j=0}^i g_j\binom{i}{j}/\binom{d}{j}$. On another interval first substitute its affine parametrization; apply this formula in each variable for a tensor polynomial. Positive Bernstein basis functions sum to one. The polynomials $P_3(t,u/2)$ and $G_3(t,u/2)$ have bidegrees $(64,6)$ and $(28,4)$. Appendix A gives the partition and its rational verification, which certify $P_3>1/50000$ throughout the admissible domain, including its boundary. The compact-minimum argument now gives

$$\|\sum_i\sigma_iA_i\|\le2\sqrt{\frac{123}{20}r^{1/3}\alpha}.\tag{385}$$

Schatten orders at least five. For every integer $s \geq 5$, use $f = 1$, $\mu = 0$, the square-root budget, and $\iota_s = \frac{3}{2}(s+1)$. First take $s = 5$, $\iota = 9$, and $\xi = 2/5$. Define $H_s^0 := \widetilde{H}_s/\eta = 1 - t^2/(s-1)$. Then $N = -t$, $A = 1$, $H_5^0 = 1 - t^2/4$, and the budget inequalities are immediate. After dividing both W_5 and R_5 by $\eta = 4$, the polynomial to certify is

$$[9((1 - t^2/4)(1 - x^3) + 6tx^2)]^2 - 36(1 + x)^2(1 - t^2)(1 + t)^2.$$

It has bidegree $(4, 6)$. With $x = (2/5)u$, its tensor Bernstein coefficients have the following lower bounds:

t-interval	u-interval	coefficient lower bound
$[3/4, 1]$	$[0, 1]$	$4891681/10^6$
$[1/2, 3/4]$	$[1/2, 1]$	$34749/20000$
$[5/8, 3/4]$	$[0, 1/2]$	$306909/10^6$
$[9/16, 5/8]$	$[1/4, 1/2]$	$208533/10^6$
$[1/2, 9/16]$	$[1/4, 1/2]$	$206021/10^6$
$[1/2, 5/8]$	$[0, 1/4]$	$359109/500000$
$[1/4, 1/2]$	$[1/2, 1]$	$729/400$
$[1/4, 1/2]$	$[0, 1/2]$	$409353/10^6$
$[0, 1/4]$	$[0, 1]$	$6309/15625$

The rectangles partition the unit square, and every lower bound exceeds $1/5$. This proves Eq. (381) for the base case.

For general $s \geq 5$, divide W_s by $\eta(s+1)$ to obtain

$$\frac{3}{2}[(1 - \frac{t^2}{s-1})(1 - x^3) + (s+1)tx^2].$$

It is at least its value at $s = 5$, while $R_s/(\eta(s+1))$ is independent of s. Also $\Gamma \geq \iota_s \geq 9 > 35/4$, so every relevant x still lies in $[0, 2/5]$. The base certificate thus proves strict curvature for every integer $s \geq 5$. Consequently

$$\|\sum_{i=1}^{N} \sigma_i A_i\| \leq 2\sqrt{\frac{3}{2}(s+1)r^{1/s}\alpha} \qquad (s \in \mathbb{Z},\ s \geq 5). \tag{386}$$

The signing may depend on the selected order.

The uniform coefficient. It suffices to treat integer r. For $1 \leq r \leq 3$, use the trace-variance bound of Theorem 3.7. For $4 \leq r \leq 32$, use Eq. (385); for $32 \leq r \leq 256$ and $256 \leq r \leq 4096$, use Eq. (386) at orders five and eight. For fixed order, $r^{1/s}/\log(2r)$ decreases and then increases, so endpoint checks suffice. The order-three endpoint values agree. For $r \geq 4096$, put $x = \log r$, $a = \log 2$, and $s = \lceil x \rceil + 1$. The squared coefficient is at most

$$F(x) := 6\frac{x+3}{x+a}\exp(\frac{x}{x+1}).$$

The numerator determining the sign of its logarithmic derivative is $(a-2)x^2+(3a-3)x+(4a-3) < 0$, since $0 < a < 3/4$. Thus $F(x) \leq F(12\log 2)$. Exact rational upper bounds give

rank range	upper bound on squared coefficient
$1 \leq r \leq 3$	18.786663
$4 \leq r \leq 32$	18.780402
$32 \leq r \leq 256$	17.494931
$256 \leq r \leq 4096$	17.313519
$r \geq 4096$	18.400913

(387)

Every entry is less than $4.335^2 = 18.792225$. The first row uses the rounded trace coefficient 3.34963. The comparisons follow from

$$0.6931 < \log 2 < 0.6932, \qquad \log 6 > 1.7917,$$
$$4^{1/3} < 1.587402, \qquad 256^{1/5} < 3.031434, \qquad 4096^{1/8} < 2.828428.$$

Integer powers verify the root bounds. The logarithmic series in Section 11.2, with 200 terms, and its degree-24 exponential bound verify the remaining comparisons. For the last row use the smaller logarithmic endpoint in the decreasing rational prefactor and the larger endpoint in the increasing exponential argument. Every signing has norm at most one. This proves the intermediate coefficient 4.335. The algorithmic analysis uses the refinement in Section 11.4.

9 Deterministic algorithm for trace-variance signing

The shared coordinate acceleration and deterministic direction selection give larger curvature moves through the potential comparison in Lemma 3.12. This yields the following arithmetic bound with the trace-variance budget and discrepancy allowance of Section 3.11.

Theorem 9.1 (Formal version of Theorem 1.2)**.** *Let $A_1, \dots, A_m$ be dense positive semidefinite Hermitian matrices of order n, of arbitrary ranks, and put $\sigma_{\mathrm{tr}} := \|\sum_i \mathrm{tr}[A_i]A_i\|^{1/2}$. For every admissible square matrix multiplication exponent $\omega \in [2,3]$, a deterministic algorithm returns one sign $\varepsilon_i \in \{-1,1\}$ per original input such that*

$$\|\sum_i \varepsilon_i A_i\| \le 3.34963\sigma_{\mathrm{tr}},$$

using

$$\widetilde{O}(mn^2 + n^{\max\{2\omega, 9/2\}}) \tag{388}$$

arithmetic operations and comparisons in the model of Section 3.11.

Using the current bound $\omega < 2.371177$ of Dupont et al. [DEK+26], the arithmetic bound is

$$\widetilde{O}(mn^2 + n^{4.742354}).$$

Proof. Normalize and apply the grouped reduction of Section 3.11, leaving $k \le n^2$ after constant-factor padding. Use the phase reservoir, small-input cutoff, and rounding allowances from that subsection. In the large-count regime use the shared resident rule of Section 3.19, so the movable light set J has $p \asymp k$ and the fixed background has size at most $\mu p/4096$. The acceleration and negative multiplicity in Section 3.17 then apply to every corrected state.

Deterministic selection with four response maps. Apply the deterministic spectral filter and matrix conditional estimator of Lemma 5.10 to $H^\sharp$, with the higher-rank block implementation in Eq. (68). On the filtered band, the true Hessian is bounded above by $P_n G_+$: it differs from $H^\sharp$ by the diagonal term in Eq. (74).

Let $A(r) := (\dot U, \dot V)$ and $B(r) := (\dot W, \dot Z)$ be the true first optimizer derivatives at the base with the full coupling, and define $\theta := r^T G_+ r$. Completing the response-energy square gives

$$e(A(r))^2 + \mathcal{E}(B(r))^2 \le P_n\theta. \tag{389}$$

Define the four relative Hermitian response maps by

$$\mathcal{T}_1 r := U^{-1/2}\dot U U^{-1/2}, \tag{390}$$

$$\mathcal{T}_2 r := V^{-1/2}\dot{V}V^{-1/2}, \tag{391}$$
$$\mathcal{T}_3 r := W^{-1/2}\dot{W}W^{-1/2}, \tag{392}$$
$$\mathcal{T}_4 r := Z^{-1/2}\dot{Z}Z^{-1/2}. \tag{393}$$

By Eq. (389), each map has scaled $2 \to F$ norm at most $P_n\sqrt{n}$ on the filtered band. The rank-$2n$ evaluation argument in Eq. (184) bounds each coefficient variance matrix by $P_n n^2$. The same conditional trace-exponential estimator, with the negative fraction absorbed in P_n, returns a unit physical direction r satisfying

$$r^T H^\sharp r \le -\theta/P_n,$$
$$R := \|r\|_\infty + \max_{1\le j\le 4} \|\mathcal{T}_j r\| \le P_n \frac{n}{\sqrt{p}}\sqrt{\theta}. \tag{394}$$

For the coordinate bound, the reservoir floor gives $\|r\|_\infty/\sqrt{\theta} \le P_n\sqrt{n}K^{-1/4}$, which is at most $P_n n/\sqrt{p}$ when $p \asymp K \le n^2$. The conditional-expectation procedure selects r deterministically. Its block cost is

$$B(n,k) = \widetilde{O}(n^3 + kn^\omega + n^2 k^{\omega-1} + k^\omega). \tag{395}$$

Changing the diagonal Hessian term adds only $O(k)$ work.

Larger deterministic moves. Orient r so that $g^T r \le 0$, with separated numerical slack, and put

$$t := \frac{a\sqrt{p}}{P_n n\sqrt{\theta}}. \tag{396}$$

Increase P_n as required by Lemma 3.12. The direction bounds in Eq. (394) meet that lemma's hypotheses at the step length in Eq. (396). Its finite-step estimate gives a true decrease of at least $ap/(P_n n^2)$. A dyadic step within fixed factors and inverse-polynomial comparison errors preserve this bound. Thus a phase $K/2 < k \le K$, $K \ge K_c = \widetilde{O}(n)$, has

$$I_K = \widetilde{O}(n^2/K). \tag{397}$$

After each accepted curvature move, recompute the root with the cold initializer. The arithmetic count includes this initialization at every move.

Bulk deletion and small-count finishing. The resident rule of Section 3.19 retains near-endpoint coordinates, including their positive coupling slots, until a fixed fraction can be removed together. There are only $\widetilde{O}(1)$ attenuation calls and $\widetilde{O}(\sqrt{n})$ corrected stages in total, as proved in Section 3.19. Classification costs one scan per signing invocation and changes no coefficient or potential. Once $k \le K_c = \widetilde{O}(n)$, use Lemma 3.11 from the current corrected state. This costs $\widetilde{O}(n^{9/2} + n^{\omega+2})$, with the same reservoir and rounding allowances.

Complete arithmetic count. One dense corrected stage has cost $F_{\mathrm{dense}}(n,k) = \widetilde{O}(n^3 + kn^2)$. For $n \le K \le n^2$, the block work in one phase satisfies

$$\begin{aligned} I_K B(n,K) &= \widetilde{O}(\frac{n^2}{K} B(n,K)) \\ &= \widetilde{O}(\frac{n^2}{K}(n^3 + Kn^\omega + n^2 K^{\omega-1} + K^\omega)) \\ &= \widetilde{O}(n^5/K + n^{\omega+2} + n^4 K^{\omega-2} + n^2 K^{\omega-1}) \\ &\le \widetilde{O}(n^{2\omega}). \end{aligned} \tag{398}$$

The first step uses Eq. (397). The second step uses Eq. (395). The third step follows from simple algebra. The last step uses $n \le K \le n^2$ and $\omega \ge 2$.

The cold root initializations after these moves cost

$$\begin{aligned} I_K \sqrt{n} F_{\mathrm{dense}}(n,K) &= \widetilde{O}(\frac{n^2}{K} \sqrt{n} F_{\mathrm{dense}}(n,K)) \\ &= \widetilde{O}(\frac{n^2}{K} \sqrt{n}(n^3 + Kn^2)) \\ &= \widetilde{O}(n^{11/2}/K + n^{9/2}) \\ &\le \widetilde{O}(n^{9/2}). \end{aligned} \tag{399}$$

The first step uses Eq. (397). The second step substitutes the definition of F_{dense}. The third step follows from simple algebra. The last step uses $K \ge n$.

There are logarithmically many phases, so summing Eqs. (398) and (399) preserves these powers. Initial and bulk-deletion roots cost at most $\widetilde{O}(\sqrt{n}\, n^4) = \widetilde{O}(n^{9/2})$. All scans outside corrected stages cost at most $\sum_K I_K O(Kn^2) = \widetilde{O}(n^4)$. Deferred rounding uses full matrix logarithmic moments and the shared fast matrix-function routine, at cost $\widetilde{O}(n^{\omega+2})$. The small-input Frobenius conditional rounding and all completed-coordinate sum updates cost $\widetilde{O}(n^4)$.

Including preprocessing and small-count finishing, the total work is

$$\begin{aligned} &\widetilde{O}(mn^2 + n^{2\omega} + n^{9/2} + n^{\omega+2}) \\ &\quad = \widetilde{O}(mn^2 + n^{\max\{2\omega, 9/2\}}). \end{aligned} \tag{400}$$

This simplification uses $\omega + 2 \le 2\omega$ for $\omega \ge 2$ and proves Eq. (388). When $m \le 2n^2$, grouped preprocessing can be omitted. The deterministic block work at $k \asymp n^2$ still contributes the $n^{2\omega}$ term.

Final discrepancy and finite comparisons. Use the potential, budget, scalar certificate, initial-value bound, phase reservoir, and rounding allowances of Section 3.11. Freezing preserves the signed coefficients and the potential. Every accepted accelerated step has a certified true decrease.

The number of steps and comparisons is polynomial. Choose inverse-polynomial comparison tolerances below the separated margins so that their total contribution remains within the 10^{-6} error allowance. The initial-value and rounding calculation in Section 3.11 then gives

$$\|\sum_i \varepsilon_i A_i\| \le 3.34963 \sigma_{\mathrm{tr}}.$$

Direction selection, root evaluation, and rounding are deterministic. Together with Eq. (388), this proves the theorem. □

10 A Las Vegas algorithm for trace-variance signing

The randomized construction uses the finite Gibbs-feature sampler of Section 6 with dense higher-rank inputs. The preprocessing lemmas below reduce arbitrary input counts to the range needed by the signing algorithm.

10.1 Randomized preprocessing

The following Nyström preconditioner uses a trace bound to control the ridge condition number, following the approach of [FTU23].

Lemma 10.1 (Ridge sketch from a trace bound)**.** *Let $C \succeq 0$ have order D, let $T \geq \max\{1, \mathrm{tr}[C]\}$, and $1 \leq q \leq D$. Set $s := 256q$, draw a D-by-s standard real Gaussian matrix Ω, and define*

$$Y := C\Omega, \quad K := \Omega^\top C\Omega + TI_s, \quad Z := YK^{-1}Y^\top. \tag{401}$$

Always $0 \preceq Z \preceq C$. With probability greater than 0.98,

$$0 \preceq C - Z \preceq 512Tq^{-1}I. \tag{402}$$

For a specified positive rational δ, a bounded finite-random-bit sampler and short matrix computations give fixed factors $\widehat{Y}, \widehat{K} = \widehat{K}^\top \succeq (T/2)I$ defining $\widehat{P} := \delta I + \widehat{Y}\widehat{K}^{-1}\widehat{Y}^\top$. With probability at least 9/10, the spectrum of $\widehat{P}^{-1}(\delta I + C)$ lies in

$$[a_{\mathrm{spec}}, \beta], \qquad a_{\mathrm{spec}} := 1/4, \quad \beta := 16(1 + 2^{12}T/(q\delta)). \tag{403}$$

For every finite sample the factors and inverse bounds are polynomial in D, T, δ^{-1} and the input magnitudes. All requested inverse-polynomial errors use logarithmic word lengths in these parameters.

Proof. Completing a positive quadratic gives the identity

$$x^\top(C - Z)x = \min_z \{\|C^{1/2}x - C^{1/2}\Omega z\|^2 + T\|z\|^2\}. \tag{404}$$

This proves both order bounds. Diagonalize C and split its first q eigenvalues from the tail Λ_2. The Gaussian blocks in this basis are denoted Ω_1, Ω_2. We verify the two events

$$\sigma_{\min}(\Omega_1) \geq \sqrt{s}/2, \qquad \|\Lambda_2^{1/2}\Omega_2\|^2 \leq 8T + 32s\|\Lambda_2\|. \tag{405}$$

For a fixed unit vector in $\mathbb{R}^q$, its squared image norm is a chi-square variable with s degrees of freedom. Its moment-generating function gives probability at most $2e^{-s/32}$ of leaving $[9s/16, 25s/16]$. For example use parameters $\pm 1/8$, and the elementary bounds $-\log(1-u) \leq u + u^2/(2(1-u))$ and $\log(1+u) \geq u - u^2/2$. A 1/8-net of at most 17^q points then gives operator norm at most $10\sqrt{s}/7$ and minimum singular value at least $4\sqrt{s}/7 > \sqrt{s}/2$. Its failure probability is at most $2\exp(q\log 17 - s/32) < 1/64$.

For the tail, set $L := \|\Lambda_2\|$. The second event holds when $L = 0$. Otherwise for a fixed unit vector in $\mathbb{R}^s$, the tail image norm squared is $X = \sum_j \lambda_j g_j^2$ and $\mathbb{E}e^{X/(4L)} \leq e^{T/(2L)}$. Thus $\Pr[X > 4T + 16sL] \leq e^{-4s}$. A 1/4-net of at most 9^s points and the norm extension factor 4/3 prove the second event, with failure at most $e^{-s} < 1/256$. The total failure is less than 0.02. The elementary net bound follows by comparing volumes of disjoint balls of radius half the net separation in a maximal separated set.

On these events insert $z = \Omega_1^\dagger x_1$ into Eq. (404). The top residual is zero. The tail residual operator is the concatenation $[-\Lambda_2^{1/2}\Omega_2\Omega_1^\dagger, \Lambda_2^{1/2}]$, so its squared norm is at most the sum of the squared norms of these blocks. The ridge term is at most $4T\|x_1\|^2/s$. Consequently,

$$\begin{aligned}\|C - Z\| &\le 129L + 36T/s \\ &\le 129T/(q+1) + 36T/s \\ &= 129T/(q+1) + 9T/(64q) \\ &\le 130T/q \\ &< 512T/q. \end{aligned} \tag{406}$$

The first step combines the residual and ridge bounds. The second uses $L \le T/(q+1)$. The third substitutes $s = 256q$. The fourth uses $129/(q+1) + 9/(64q) \le 130/q$, and the last compares the constants. For $q = D$, the tail is zero.

Use midpoint dyadic uniforms and certified Box–Muller evaluation for the actual sampler. Couple to ideal uniforms. Except on an event of probability less than $1/100$, all radial uniforms exceed $(100Ds)^{-4}$ and all normals have a fixed polynomial magnitude bound. On that event the derivatives of the scalar sampling maps are polynomial, so a sufficiently fine grid makes $\|\Delta\Omega\|$ as small as any prescribed inverse polynomial. Every finite output is bounded by $C\sqrt{b}$ when b uniform bits are used, even outside that event. The sampler uses a fixed finite number of random bits.

For $\|\Delta\Omega\| \le 1$, expanding Eq. (401), using $K \succeq TI$ for both samples, bounds its perturbation by $16T(1 + \|\Omega\|)^4\|\Delta\Omega\|$. Make this less than $\delta/64$. Compute the factors and their Gram to still finer normwise accuracy, symmetrizing the small Gram. Their polynomial bounds and $K \succeq TI$ give $\widehat{K} \succeq (T/2)I$ and permit

$$\|\widehat{P} - (\delta I + Z)\| \le \delta/16.$$

The same accuracy and positivity guarantees apply to every finite sample. On the ideal good event, $\delta I + Z \preceq \delta I + C \preceq (1 + 512T/(q\delta))(\delta I + Z)$. The perturbation preserves the bounds in Eq. (403). A union bound gives success probability at least $9/10$. All parameters, norm bounds and inverse lower bounds are explicit polynomials; logarithmic precision suffices for the construction. □

The ridge sketch gives a grouped cube walk with a certified operator-norm residual. The walk uses the grouping and slack scaling of [SY26a] with explicit dense columns.

Lemma 10.2 (Randomized preprocessing for dense higher-rank inputs)**.** *Suppose $A_1, \dots, A_m \succeq 0$ are Hermitian of order n and $\sum_i \operatorname{tr}[A_i]A_i \preceq I$. For a fixed $0 < \eta < 1$, there is a capped randomized algorithm returning $x \in [-1, 1]^m$ with at most $2n^2$ fractional coordinates and $\|\sum_i x_i A_i\| \le \eta$ for every random-bit sequence. Its expected arithmetic cost is*

$$\widetilde{O}(mn^2 + n^{9/2} + n^{\omega+2}). \tag{407}$$

The suppressed factors include logarithms of m, n and the fixed error allowance. The algorithm acts directly on the dense input matrices.

Proof. Set $D := n^2$. We measure the residual in operator norm.

A tree with bounded total energy. Let $\mathcal{V}$ list the diagonal entries and the real and imaginary parts above the diagonal of a Hermitian matrix. Then $\|\mathcal{V}(H)\|_2 \le \|H\|_F \le \sqrt{2}\|\mathcal{V}(H)\|_2$. For preprocessing only, approximate $\mathcal{V}(A_i)$ by a rational vector b_i on one common dyadic grid, so that $\|\mathcal{V}^{-1}(b_i) - A_i\|_F \le \eta/(16m)$. This costs $\widetilde{O}(mD)$ in the arithmetic model. The total energy satisfies

$$\begin{aligned}\sum_i \|\mathcal{V}(A_i)\|_2^2 &\le \sum_i \operatorname{tr}[A_i^2] \\ &\le \sum_i \operatorname{tr}[A_i]^2 \\ &= \operatorname{tr}[\sum_i \operatorname{tr}[A_i] A_i] \\ &\le n. \end{aligned} \tag{408}$$

The first inequality uses the vectorization bound. The second uses the nonnegative eigenvalues of each A_i. The equality uses linearity of trace, and the last inequality uses the normalization hypothesis.

Thus $\sum_i \|b_i\|_2^2 \le 3n$. The original matrices remain the inputs to the subsequent nonlinear algorithm.

Fix a balanced binary tree of height $R := \max\{1, \lceil \log_2 m \rceil\}$. At each internal node choose the relative sign of its two child sums so that their inner product is nonpositive. This can be done bottom up using $O(mD)$ operations: if the child sums are u, v, choose $s \in \{-1, 1\}$ with $s\langle u, v\rangle \le 0$ and store $u + sv$ and the edge sign s. A final traversal propagates all edge signs to leaf signs e_i. Changing the orientation of a whole subtree changes none of its squared norms. Consequently its globally oriented node sums $w_J := \sum_{i \in J} e_i b_i$ satisfy the following bounds, where $T := 3(R+1)n$:

$$\|w_J\|_2^2 \le \sum_{i \in J} \|b_i\|_2^2, \qquad \sum_{J \text{ a node}} \|w_J\|_2^2 \le T. \tag{409}$$

All comparisons are exact on the short rational data. In particular, any adaptively selected collection of distinct nodes has total squared norm at most $T = O(n \log(2m))$, deterministically.

A matrix residual for an approximate null direction. For a real linear map $\mathcal{L} : \mathbb{R}^r \to \mathbb{H}_n$ with $\|\mathcal{L}\|_{2\to F} \le a$, define $E_j := \mathcal{L}(e_j)$. Then

$$\sum_j E_j^2 \preceq na^2 I, \qquad \Pr[\|\mathcal{L}\xi\| \ge t] \le 2n \exp(-t^2/(2na^2)), \tag{410}$$

where ξ has independent uniform signs. To see the first assertion, choose a real orthonormal Hermitian basis $H_1, \dots, H_D$. It has $\sum_j H_j^2 = nI$. The coefficient matrix of $\mathcal{L}$ satisfies $LL^{\mathsf{T}} \preceq a^2 I$. The map $Q \mapsto \sum_{j,l} Q_{jl} H_j H_l$ is positive on real symmetric positive semidefinite matrices, since it maps vv^{T} to $(\sum_j v_j H_j)^2$. Apply this map to $LL^{\mathsf{T}} \preceq a^2 I$. The second assertion is the matrix Rademacher bound [Tro12], with this variance; it also follows from the matrix exponential argument used in Lemma 5.3.

Consider at most $r_0 \le 4D$ active group columns in a matrix A, with $\|A\|_F^2 \le T$ and inherited cube coefficients y. Let x be the current coefficient vector, initially y, and let r be the number of its fractional coordinates. Set $S := \operatorname{diag}(1 - |x_i|)$ and $B := AS$. The operator

$$\Pi_\delta := I - B^{\mathsf{T}}(\delta I + BB^{\mathsf{T}})^{-1} B$$

is a contraction with at least $r - D$ eigenvalues equal to one, and $\|B\Pi_\delta\| \le \sqrt{\delta}/2$. These facts follow from its singular-value formula. Since $\|\mathcal{V}^{-1}B\Pi_\delta\|_{2\to F} \le \sqrt{\delta/2}$, Eq. (410) gives

$$\Pr[\|\mathcal{V}^{-1}(B\Pi_\delta\xi)\| \ge t] \le 2n\exp(-t^2/(n\delta)). \tag{411}$$

Thus an inverse ridge of order n, up to logarithms and the fixed accuracy, suffices.

A finite small reduction. Fix a per-call error $0 < \epsilon < 1$, and a failure budget $0 < \zeta < 1/2$. Let M be a power of two with $2\sqrt{T} \le M < 4\sqrt{T}$. It bounds both $\|A\|$ and the operator norm of each corresponding Hermitian column. If at most $2D$ coordinates are fractional, no reduction is necessary. Otherwise set

$$\ell := \lceil \log_2(64 r_0) \rceil, \qquad H := 4\ell, \qquad \frac{\epsilon}{2^{12} r_0 M} < \sigma \le \frac{\epsilon}{2^{11} r_0 M},$$

where σ is dyadic, and set

$$\begin{aligned} L_\sigma &:= \lceil \log_2(2/\sigma) \rceil, & N_* &:= 128 H^2 L_\sigma + 1, \\ \nu &:= \frac{\epsilon}{2^{16} N_* M}, & L_n &:= \lceil \log_2(128 n) \rceil, \\ \delta_0 &:= \frac{\epsilon^2}{2^{24} n N_*^2 L_n}, & \delta_0/2 &< \delta \le \delta_0. \end{aligned}$$

Choose δ dyadically. Use a dyadic coefficient mesh u, within a factor two of

$$\min\{2^{-p_y},\ \sigma/(2^{16}H^2), \epsilon/(2^{16} N_* r_0 M)\},$$

where 2^{-p_y} divides all inherited coefficients. Mesh exponents are combined by taking their maximum at subsequent levels.

At each trial use Lemma 10.1 with $C = BB^{\mathsf{T}}$, $q = n$, and the same T, then draw a fresh independent sign vector ξ. Solve $(\delta I + BB^{\mathsf{T}})z = B\xi$ with true Euclidean residual at most $\delta\nu/(8M)$, and form a dyadic approximation $\widetilde{v}$ to $\xi - B^{\mathsf{T}}z$ satisfying $\|\widetilde{v} - \Pi_\delta\xi\|_2 \le \nu$. Allocate the remaining error to the right-hand side and matrix products. This implication uses only $\|(\delta I + BB^{\mathsf{T}})^{-1}\| \le \delta^{-1}$ and $\|B\| \le M$, and is certified by the true residual.

Require the exact tests $\|\widetilde{v}\|_\infty \le \ell$ and $\|\widetilde{v}\|_2^2 \ge (r - D)/4$. The moment and scalar-tail proof of Lemma 6.17 gives success probability at least $1/16$; that proof uses only the contraction and null-space dimension, so it does not require a unit bound on the group columns. For $h := S\widetilde{v}/H$, also apply the finite norm certificate of Lemma 6.26 to $E_h := \mathcal{V}^{-1}(Ah)$, with threshold $t_* := \epsilon/(2^8 N_*)$. By Eq. (411), the ideal unscaled matrix residual exceeds t_* with probability at most $2n\exp(-2^8 L_n) < 1/1024$. On the complementary event,

$$\|E_h\| \le (t_* + M\nu)/H < t_*/2,$$

so the norm certificate accepts. The sketch is independent of the sign vector. The complete trial therefore succeeds with conditional probability at least $(9/10)(1/16 - 1/1024) > 1/32$ at every state.

For an accepted direction, $|h_i| \le (1-|x_i|)/4$. For the barrier $\Phi(x) := -\sum_i \log(1-x_i^2)$, use the short-precision sign rule of Lemma 7.17. Approximate each summand of $g = \sum_i 2x_i\widetilde{v}_i/(H(1+|x_i|))$ on the grid of mesh 2^{-q}, where $q = \lceil \log_2(512H^2) \rceil$, and sum on that common grid. Choose $\tau = 1$ when the sum is nonnegative and $\tau = -1$ otherwise. Round $x + \tau h$ inward on the coefficient mesh and freeze coefficients within σ of an endpoint. The lemma gives gain at least $(2/25 - 1/256)(r - D)/H^2$ before rounding. Rounding loses at most $4ru/\sigma \le (r-D)/(8192H^2)$, so the net gain is greater than $(r-D)/(16H^2) > D/(16H^2)$ while $r > 2D$. Every sign-selection operand has logarithmic length; no product across all coordinates is formed. Changing this sign preserves $\|E_h\|$ and all residual allowances. Assign each frozen coordinate the bookkeeping value $\log(2/\sigma)$. Freezing after a step cannot lower this bookkeeping barrier, and its total never exceeds $4DL_\sigma$. Perform the initial near-endpoint scan before starting this accounting. Thus fewer than N_* accepted steps leave at most $2D$ fractional coefficients. Every accepted step, independent of its sampling outcome, changes the matrix sum by norm at most t_*. Mesh changes cost at most $r_0 Mu$ per step; all freezes together cost at most $r_0 M\sigma$. The output residual of the fast branch is consequently at most

$$\begin{aligned} N_* t_* + N_* r_0 M u + r_0 M \sigma &\le \epsilon(2^{-8} + 2^{-16} + 2^{-11}) \\ &< \epsilon/128. \end{aligned} \tag{412}$$

The first inequality substitutes the bounds on t_*, u, and σ; the second adds their error allowances. The reduction retains inherited coefficients and applies to arbitrary Hermitian group matrices.

Dense sketch setup and inverse work. The matrices B have $D = n^2$ rows and at most $4D$ columns. For the sketch with $s = 256n$, form

$$U := B^{\mathsf{T}}\Omega, \qquad Y := BU, \qquad K := U^{\mathsf{T}}U + TI_s.$$

A product of shape $(O(n^2), O(n^2), O(n))$ splits into $O(n^2)$ products of order $O(n)$ and costs $\widetilde{O}(n^{\omega+2})$. Thus both large products cost that amount. Forming the small Grams and inverting the order-s Woodbury matrix costs at most $\widetilde{O}(n^{\omega+1})$. Explicitly, for the represented factors in Lemma 10.1, use

$$\widehat{P}^{-1} = \delta^{-1}I - \delta^{-2}\widehat{Y}(\widehat{K} + \delta^{-1}\widehat{Y}^{\mathsf{T}}\widehat{Y})^{-1}\widehat{Y}^{\mathsf{T}}.$$

One application costs $\widetilde{O}(Ds + s^2) = \widetilde{O}(n^3)$. Apply BB^{T} through the stored factors B and B^{T} at cost $O(Dr_0) = O(n^4)$ per action.

On the good sketch event the preconditioned condition number is $O(1+T/(n\delta))$. The normalized Chebyshev iteration of Eq. (256), or its positive-system specialization, therefore attains the required residual using $\widetilde{O}(1 + \sqrt{T/(n\delta)})$ iterations. Its polynomial similarity and inverse bounds permit certified absolute errors with logarithmic precision, as in Section 7.2. The Woodbury inverse is approximated to the required absolute accuracy, including cancellation. The true residual and all matrix norm certificates cost $\widetilde{O}(n^4 + n^\omega)$ per trial.

In the tree application $T = 3(R+1)n$ and $\epsilon = \eta/(4R)$. Thus $\delta^{-1} = \widetilde{O}(n)$ and $1 + \sqrt{T/(n\delta)} = \widetilde{O}(\sqrt{n})$, with logarithms of m included. A full trial, including its setup, costs

$$\begin{aligned} &\widetilde{O}(n^{\omega+2} + n^4 + \sqrt{n}(n^4 + n^3)) \\ &\quad = \widetilde{O}(n^{\omega+2} + n^{9/2}). \end{aligned} \tag{413}$$

The three terms account for sketch setup, certification, and inverse iterations, respectively. The equality uses $n^4, n^{7/2} \le n^{9/2}$.

Every loop and intermediate magnitude is capped at the proved polynomial bound; a failed residual or enclosure rejects the trial. On the good sketch event the finite precision construction and the iteration bound guarantee these checks. On a bad event the cap still bounds the cost.

The full tree, certification, and fallback. Start with root coefficient zero. Keep at most $2D$ fractional groups, split each nonsingleton into its two children and copy its coefficient, and apply the small reduction to the resulting at most $4D$ columns. Eq. (409) bounds their energy at every state. Freeze a group by assigning its leaf signs $e_i x_J$; represent these assignments lazily until the final tree traversal. Splitting preserves the matrix sum exactly. After at most R levels the remaining fractional groups are leaves. Their total approximate-matrix residual is at most $R\epsilon \le \eta/4$. Transferring to the original matrices adds at most $\eta/16$, since all final leaf coefficients lie in $[-1, 1]$. In particular the true operator residual is strictly below $\eta/2$.

Permit $64\lceil \log_2(2N_*/\zeta)\rceil$ trials at a state and at most N_* successful steps in a small call. Choose $\zeta := (2mn)^{-30}/R$. Condition on the complete past at each trial. The success probability just proved and a union bound make the probability that any cap is exhausted at most $(2mn)^{-30}$. On exhaustion restart the deterministic grouped reduction of Lemma 5.9 from the saved original instance, with its Euclidean residual chosen below $\eta/2$ and the same input-approximation allowance. Its cost is $\widetilde{O}(mD + D^\omega)$, and its output also satisfies the required operator residual and fractional count. Before returning a fast output, check its cube, count, and true matrix residual using Lemma 6.26, with threshold η and enclosures for the original input sum. The strict $\eta/2$ margin guarantees acceptance of the good fast branch. The final accumulation costs $O(mD)$.

There are only polylogarithmically many steps, trials and levels. Building all tree sums and producing the final coefficients costs $\widetilde{O}(mD)$, once. Eq. (413) and the negligible expected fallback cost prove Eq. (407). All finite outcomes return a correct output. Only the expected work uses the sketch event. □

Increasing the sketch size reduces the number of inverse iterations. Rectangular matrix multiplication then gives the following preprocessing bound.

Lemma 10.3 (Rectangular implementation of randomized preprocessing)**.** *Under the hypotheses of Lemma 10.2, the same output guarantees hold in expected* $\widetilde{O}(mn^2 + n^{4.31})$ *arithmetic operations.*

Proof. In the proof of Lemma 10.2, set $q := \lceil n^{7/5}\rceil$ and $s := 256q$. Keep all other parameters, acceptance tests, and the fallback unchanged. For $n \ge 2$ we have $q \le D = n^2$; bounded orders take constant work per input. Lemma 10.1 applies at every adaptive state with the same success probability. Acceptance still uses the true ridge residual. All cube and residual bounds remain valid, and every finite outcome is correct.

The trace and inverse-ridge bounds are $T = \widetilde{O}(n)$ and $\delta^{-1} = \widetilde{O}(n)$. The good-event iteration count is now

$$\begin{aligned} J &= \widetilde{O}(1 + \sqrt{T/(q\delta)}) \\ &= \widetilde{O}(n^{3/10}). \end{aligned} \tag{414}$$

The first bound is the Chebyshev iteration estimate. The second substitutes $T, \delta^{-1} = \widetilde{O}(n)$ and $q = \lceil n^{7/5}\rceil$.

The two large setup products $B^\mathsf{T}\Omega$ and $B(B^\mathsf{T}\Omega)$ have dimensions $(O(D), O(D), O(D^{7/10}))$. The rectangular bound $\omega(1, 1, 7/10) \le 2.152770$ in [ADW+25], with any fixed positive exponent slack, computes both in $O(n^{4.305540+o(1)})$ operations. Permuting the three dimensions of the matrix multiplication tensor gives the bound for either orientation of these products.

The current bound $\omega < 2.371177$ of Dupont et al. [DEK+26] justifies the rounded upper bound $\omega_0 := 2.38$ for the smaller-matrix operation counts. Forming the order-s Grams costs $\widetilde{O}(Ds^{\omega_0-1})$, with exponent 3.932; inverting their Woodbury matrix costs $\widetilde{O}(s^{\omega_0})$, with exponent 3.332. A preconditioner application costs $\widetilde{O}(Ds + s^2) = \widetilde{O}(n^{3.4})$. A true BB^T action still uses the explicit factors and costs $O(n^4)$, so all inverse iterations together cost $\widetilde{O}(n^{4.3})$. True residual and matrix

norm checks cost $\widetilde{O}(n^4 + n^{\omega_0})$ per trial. The polynomial bounds on the represented factors and inverse remain valid for $n \le q \le D$, so the same absolute-error certification requires only logarithmic precision. Cap the iteration count at Eq. (414), with the original logarithmic accuracy factors, and reject failed certificates. On the good sketch event this cap suffices.

The strict slack $4.305540 < 4.31$ absorbs the exponent slack in the rectangular algorithm. Hence each capped trial costs $\widetilde{O}(n^{4.31})$. The conditional success probability, polylogarithmic number of trials and tree levels, and negligible expected deterministic fallback cost are exactly as in Lemma 10.2. Tree construction, input conversion, and final accumulation cost $\widetilde{O}(mn^2)$ once. These bounds prove the claim. □

10.2 The Las Vegas signing algorithm

The preprocessing bound combines with verified randomized curvature steps to give the Las Vegas signing algorithm. Its discrepancy guarantee holds for every random-bit sequence.

Theorem 10.4 (Formal version of Theorem 1.3)**.** *Let $A_1, \dots, A_m$ be dense positive semidefinite Hermitian matrices of order n, of arbitrary ranks, and put $\sigma_{\mathrm{tr}} := \|\sum_i \mathrm{tr}[A_i]A_i\|^{1/2}$. For every admissible square matrix multiplication exponent $\omega \in [2,3]$, a Las Vegas algorithm returns one sign $\varepsilon_i \in \{-1,1\}$ per original input satisfying*

$$\|\sum_i \varepsilon_i A_i\| \le 3.34963\sigma_{\mathrm{tr}}$$

in expected

$$\widetilde{O}(mn^2 + n^{\max\{9/2,\omega+2\}}) \tag{415}$$

arithmetic operations and comparisons in the model of Section 3.11. Every random-bit sequence terminates with the stated discrepancy guarantee. The current bound $\omega < 2.371177 < 5/2$ of Dupont et al. [DEK+26] gives $\widetilde{O}(mn^2 + n^{9/2})$ for arbitrary m.

Proof. Apply Lemma 10.3 with the same fixed partial-signing allowance reserved in Section 3.11. Let x be its output and $R_0 := \sum_i x_i A_i$, so $\|R_0\| \le \eta$. Keep each original positive semidefinite input matrix; the tree orientations affect only its coefficient. Initialize the signed-sum part of the modeled potential as $\sum_{i\ \mathrm{fixed}} x_i A_i + \sum_{i\ \mathrm{fractional}} y_i A_i - R_0$, which is zero at $y = x$. The trace coupling and its positivity are unchanged. The initial-value bound uses $\beta(x_i) \le 1$ and therefore is the same as for the previous grouped reduction. At the final transfer add back R_0 and charge only its operator norm to the existing partial-signing allowance. This transfer uses only the operator norm of the fixed residual. Constant-factor padding ensures $k \le n^2$ for the working order.

Using the shared acceleration estimates. Use the coefficients a_i, modified Hessian $H^\sharp$, and negative multiplicity proved in Section 3.17. These estimates hold at the adaptively chosen full root before every sampled trial and do not depend on the trial's random choices.

Verified directions and curved progress. The Schur completion in Lemma 6.4 applies to $A^\sharp$ as well: its added diagonal has norm at most P_n after G_+ scaling. Consequently a positive shift of $A^\sharp$ is comparable within polylogarithmic factors to $D + RI$, where

$$D := G_+^{-1/2} F_1^* \mathcal{Q} F_1 G_+^{-1/2} \succeq 0, \qquad D_{ii} \le P_n.$$

The finite Gibbs features and alternating ridge sampler of Sections 6.6–6.8 and 6.16 apply to this same frozen Gram matrix. Their distributional and sampling statements apply to arbitrary input ranks. We bound the dense implementation cost below.

Use the flat spectral filter and a fresh sign vector to find a real unit physical direction r, supported on J, with $r^{\mathsf{T}} H^{\sharp} r \le -\delta_n \theta/2$, where $\theta := r^{\mathsf{T}} G_+ r$. Verify this inequality and all six relative optimizer-response norms in Eq. (311), including the two generally non-Hermitian similarities. Let $\mathfrak{r}$ be the maximum of these six norms. The response energy on the filtered band is at most P_n after scaling: the true Hessian differs from $H^{\sharp}$ by a diagonal bounded by $P_n G_+$, so the same Schur energy completion applies. The coefficient-matrix variance bound used in Lemma 6.30 then gives

$$\ell := \|r\|_\infty + \mathfrak{r}(z'r) \le P_n \frac{n}{\sqrt{p}} \sqrt{\theta} \qquad (n \le p \asymp k \le n^2) \tag{416}$$

with absolute conditional success probability. For the coordinate part, $\min_i (G_+)_{ii} \ge a/(P_n n\sqrt{K})$ and the flat sign-vector bound give $\|r\|_\infty \le P_n \sqrt{n}\, K^{-1/4} \sqrt{\theta}$. This is at most the right side of Eq. (416) up to fixed factors when $p \asymp K \le n^2$.

Freeze a_i and take the quadratic path

$$y_i(s) := y_i + s r_i - s^2 a_i r_i^2 \quad (i \in J), \qquad y_i(s) := y_i \quad (i \notin J). \tag{417}$$

Its second potential derivative at zero is exactly $r^{\mathsf{T}} H^{\sharp} r$, including the gradient term. We use the direct comparison of Lemma 3.12. Its response-energy hypothesis follows from the Schur energy bound above, and its four Hermitian relative norms are among the six norms already verified. Thus its parameter R is at most ℓ in Eq. (416). Choose

$$s_* := \frac{a\sqrt{p}}{P_n n \sqrt{\theta}},$$

where $a > 0$ is sufficiently small and the fixed polylogarithm P_n is enlarged to absorb all comparison constants and the inverse curvature margin δ_n^{-1}. Then $|s_*| R$ is a sufficiently small inverse polylogarithm. The cubic and quartic errors in Eq. (76), divided by θs_*^2, are bounded by $P_n R |s_*|$ and $P_n R^2 s_*^2$, respectively. They are therefore smaller than a fixed fraction of δ_n. Also $|s_*| \|r\|_\infty$ is small enough to preserve the enlarged retained interval and the endpoint margin. Orient s_* so that the initial linear term is nonpositive, allowing the usual certified derivative error. The finite-step comparison gives true potential decrease at least $ap/(P_n n^2)$. Thus the number I_K of accepted moves in a phase satisfies

$$I_K = \widetilde{O}(n^2/K) \qquad (K/2 < k \le K). \tag{418}$$

All quantities at the base root and the comparison remainder are computed with the existing inverse-polynomial error slack. After accepting the move, initialize and certify the new full root by the shared root routine before drawing another direction. The comparison certifies the potential decrease. Each accepted move uses a new full-root initialization, whose cost is included below.

The cost of a high-rank sampled trial. For $n \le k \le n^2$ a full root action, score scan, true Hessian product, or response verification costs $\widetilde{O}(n^3 + n^2 k) = \widetilde{O}(n^2 k)$. For selected lists of q coordinates and s Gibbs features, store the dense input matrices and the sampled feature factors. Each feature is the sum of two block rank-two matrices and a multiple of the fixed dense Gibbs pair. A feature-matrix product and its adjoint can therefore be applied in

$$\widetilde{O}(n^2(q+s)) \tag{419}$$

operations: form one weighted sum of the selected input matrices and pair it with the feature matrices, or reverse these two steps. Scalar p_i, q_i factors, diagonal weights, and Gibbs centering do not change this bound. After one $\widetilde{O}(n^3)$ spectral preparation, sampled power vectors cost $O(n^2)$ each. Repeated indices can be combined or processed with their multiplicities.

Use the alternating ridge recursion of Eqs. (242)–(249), with terminal sketch size n up to polylogarithms. There is one extra terminal setup cost: its feature matrix is formed as a product of an $\widetilde{O}(n) \times O(n^2)$ coordinate matrix and an $O(n^2) \times \widetilde{O}(n)$ feature matrix. Splitting the long dimension into order-n blocks costs $\widetilde{O}(n^{\omega+1})$. Its Gram and inverse cost only $\widetilde{O}(n^{\omega})$. Terminal materialization therefore contributes $\widetilde{O}(n^{\omega+1})$ to the trial cost.

To sum the recursive work, pad $\kappa \asymp k$ and the terminal order to powers of two, put $L_0 = O(\log n)$, and let D absorb fixed per-level logarithmic costs. The expansion in Lemma 6.38, with the new action costs, is bounded by

$$L_0^{4D}(n^3 + n^2 q_1 + n^{\omega+1} + \sqrt{1+\kappa/q_1}\, n^2\kappa) + \sum_{j=1}^{h-1} L_0^{D(j+4)} \sqrt{(\kappa+q_1)/q_{j+1}}\, n^2 q_{j-1} + L_0^{D(h+4)} \sqrt{(\kappa+q_1)/n}\, n^2. \tag{420}$$

Here $q_0 = q_1$. Use the repeated-deficit schedule in Eq. (326), setting

$$b_{\max} := \log_2(\kappa/n), \qquad b_0 := \min\{b_{\max}, C_* \lceil \log_2 L_0 \rceil\},$$
$$b_{t+1} := \min\{b_{\max}, \lceil 3b_t/2 \rceil\}.$$

Use deficits $(b_0, b_1, b_1, \ldots, b_T, b_T)$; if already capped, use (b_0, b_0). Set $q_j := \kappa 2^{-\ell_j}$ and choose the fixed $C_* \geq 2000(D+1)$. Relative to $n^2\kappa$, an internal term in Eq. (420) is at most a fixed factor times

$$L_0^{D(j+4)} 2^{-\ell_{j-1}+\ell_{j+1}/2} \leq 2L_0^{D(j+4)} 2^{-\ell_{j-1}/4}.$$

The growing deficits dominate the level overhead, exactly as in Lemma 6.39, and make these terms summable with a fixed logarithmic degree. The top factor is polylogarithmic. For the leaf, $L_0^{D(h+4)} \leq n^{1/100}$ in the nontruncated schedule after increasing C_*; hence its cost is at most a fixed factor times $n^{3/2+1/100}\sqrt{\kappa}$, below $n^2\kappa$ for $\kappa \geq n$. In the immediately capped case the depth is two and all logarithmic powers are fixed. Thus a complete capped trial costs

$$\widetilde{O}(n^2 k + n^{\omega+1}). \tag{421}$$

This includes sample preparation, terminal materialization, recursive inverses, and the true response and Rayleigh checks. Root initialization after an accepted move is charged separately. As proved in Lemma 6.38, the sampler has absolute conditional success probability at every adaptively chosen root. The true response and Rayleigh checks certify acceptance.

Resident coordinates and bulk deletion. Use Section 3.19, with its switching threshold $K_c = \widetilde{O}(n)$ and fixed resident-count fraction. It ensures $p \asymp k$ and the background bound required by the shared negative-multiplicity argument. All large-count signing moves are curvature steps counted by Eq. (418). All bulk deletions use $\widetilde{O}(\sqrt{n})$ corrected stages. At $k \leq K_c$, invoke the shared finishing routine with the current reservoir and rounding ledger.

Shared final rounding and randomized total work. Use the common final-rounding routine of Section 3.15. Its full matrix-function calculation costs $\widetilde{O}(kn^\omega)$ for k deferred inputs.

For a phase with $n \le K \le n^2$, multiply the trial cost in Eq. (421) by the move count in Eq. (418). The expected direction construction and verification work is

$$\begin{aligned}\widetilde{O}(&\frac{n^2}{K}(n^2K + n^{\omega+1}))\\ &= \widetilde{O}(n^4 + \frac{n^{\omega+3}}{K})\\ &= \widetilde{O}(n^4 + n^{\omega+2}).\end{aligned} \tag{422}$$

The first expression multiplies the expected trial cost by the move count. The first equality expands the product, and the last step uses $K \ge n$.

Each accepted move additionally requires a certified full-root initialization. The shared bound is $\widetilde{O}(\sqrt{n})$ stages, each costing $\widetilde{O}(n^3 + n^2K)$. Its phase total is therefore

$$\begin{aligned}\widetilde{O}(&\frac{n^2}{K}\sqrt{n}(n^3 + n^2K))\\ &= \widetilde{O}(\frac{n^{11/2}}{K} + n^{9/2})\\ &= \widetilde{O}(n^{9/2}).\end{aligned} \tag{423}$$

The first expression multiplies the move count, the stages per initialization, and the cost per stage. The first equality expands the product, and the last step uses $K \ge n$.

Logarithmically many phases preserve these bounds. Initial and bulk continuations cost at most $\widetilde{O}(\sqrt{n}n^4) = \widetilde{O}(n^{9/2})$. By Lemma 3.11, small-count finishing costs $\widetilde{O}(n^{9/2} + n^{\omega+2})$, including the cold endpoint initializations, block curvature searches, sweeps, and final ledger. Final rounding costs $\widetilde{O}(n^{\omega+2})$. Lemma 10.3 bounds preprocessing by $\widetilde{O}(mn^2 + n^{4.31})$ in expectation, including input conversion, all group sums, and the capped deterministic fallback. Adding these costs proves Eq. (415). Although preprocessing and direction work have smaller exponents, the certified root initializations, bulk continuations, and small-count endpoint initializations still contribute $\widetilde{O}(n^{9/2})$ to the full bound.

Cap the number of attempts at each random invocation at $C \log n$, with C large enough for the proved polynomial event cap. On exhaustion, restart the deterministic algorithm of Theorem 9.1 from the saved initial instance. Its polynomial cost times the probability of any exhaustion is negligible in the displayed expectation. All random-bit sequences therefore terminate correctly; randomness affects work, not the signing guarantee. All accepted moves use certified true-model inequalities, and the original total 10^{-6} discrepancy allowance still suffices. The scalar certificate gives the same discrepancy bound.

If $m \le 2n^2$, preprocessing may be skipped. The current bound $\omega < 2.371177 < 5/2$ [DEK+26] reduces Eq. (415) to $\widetilde{O}(mn^2 + n^{9/2})$ for every m.

□

11 Signing via Schatten norms with logarithmic rank dependence

Throughout this section, $A_1, \dots, A_N \in \mathbb{C}^{D\times D}$ are nonzero Hermitian positive semidefinite matrices satisfying $\sum_{i=1}^N A_i = I_D$, $\|A_i\| \le \alpha$, and $\operatorname{rank}(A_i) \le r$, where $r \ge 1$ and $\alpha > 0$. Zero inputs may be discarded and assigned arbitrary signs afterward.

11.1 Compact minimum and removal of regularization

We first convert the endpoint and curvature estimates for the potential in Eq. (343) into an initial signing bound.

Lemma 11.1 (Compact-minimum signing bound)**.** *Under the standing hypotheses of this section, for every integer $s \geq 2$ there are signs $\sigma_1, \dots, \sigma_N \in \{-1, 1\}$ satisfying*

$$\|\sum_{i=1}^{N} \sigma_i A_i\| \leq \min\{1, 8\sqrt{(s+3) r^{1/s} \alpha}\}. \tag{424}$$

Proof. Fix $s \geq 2$ and set $\iota := 16(s+3)$. Choose a global minimizer of the continuous potential P_ρ on the cube with the fewest active coordinates. If an endpoint test succeeds, Eq. (346) or Eq. (347) gives a global minimizer with fewer active coordinates, a contradiction. Otherwise the minimizer is light. If an active coordinate remains, the minimizer is stationary on its open face and its active Hessian is positive semidefinite. This contradicts Eq. (368). Thus the minimizer is a vertex $\sigma \in \{-1, 1\}^N$.

We have

$$\|\sum_{i=1}^{N} \sigma_i A_i\| \leq P_\rho(\sigma) \leq P_\rho(0) \leq 2\sqrt{\iota r^{1/s} \alpha + 2D\rho}.$$

The first step follows from the potential constraints in Eq. (343). The second step follows from global minimality. The last step is the initialization bound in Eq. (345).

Take a sequence $\rho \downarrow 0$. There are finitely many signings, so one occurs along a subsequence. With $\iota = 16(s+3)$, this proves the square-root bound in Eq. (424). Every signing has norm at most one, since $-I \preceq \sum_{i=1}^{N} \sigma_i A_i \preceq I$, proving the cap by one in Eq. (424).

For the baseline uniform bound, define $L := \log(2r)$ and choose $s := \lceil 2L \rceil$. Then $s \geq 2$,

$$r^{1/s} \leq \sqrt{e} < 2, \qquad s + 3 \leq 2L + 4 \leq 8L,$$

where the last inequality uses $L \geq \log 2 > 2/3$. Therefore $8\sqrt{(s+3) r^{1/s} \alpha} \leq 32\sqrt{\alpha L}$.

□

11.2 Order-dependent budgets and scalar certificates

We first establish the intermediate bound

$$\|\sum_{i=1}^{N} \sigma_i A_i\| \leq \min\{1, 5.1\sqrt{\alpha \log(2r)}\}. \tag{425}$$

The coordinate-dependent estimates in Section 8.8 provide the additional budgets used in the rank-dependent refinement.

The negative functional-Hessian term and the coercivity in Eq. (361) permit smaller couplings than those in Lemma 11.1. The curvature and covariance estimates in Eqs. (369) and (370) reduce the choice of coupling to a scalar budget comparison.

Schatten order four. Choose

$$s := 4, \qquad \beta(y) := (1 - y^2)^{9/20}, \qquad \iota := \frac{165}{16}, \qquad \vartheta := \frac{11}{4}. \tag{426}$$

Then Eq. (351) gives $\gamma = 165/16$ and $\kappa = 1815/1267$. The budget is positive and smooth on $(-1,1)$, continuous on $[-1,1]$, vanishes at the endpoints, and satisfies $\beta(0) = 1$. Consequently the potential and endpoint estimates in Sections 8.2 and 8.3 apply.

At a light state, put $t := |y_i|$. Since $\beta'/\beta = -(9/10)y/(1-y^2)$, the coordinate coefficients satisfy

$$p_i, q_i > 0, \qquad p_i q_i \le \frac{1 - t/10}{1-t}, \qquad \Gamma_i > \iota \frac{\beta_i}{1-y_i^2} = \iota(1-y_i^2)^{-11/20} \ge \iota. \tag{427}$$

Indeed, for $y_i \ge 0$ the two endpoint tests give $(1+t/10)/(1+t) < p_i \le 1$ and $1 \le q_i < (1 - t/10)/(1-t)$; for negative y_i the roles interchange. The last inequality follows by multiplying the two lightness inequalities, giving $u_i v_i < (1-y_i^2)/(\iota^2 \beta_i^2)$. These properties are precisely what the response-graph construction requires. In particular Eq. (361) gives the coercivity with $\gamma = \iota$.

For the budget in Eq. (426),

$$L_4(y) = \frac{9}{10}(1-t^2)^{-31/20}(1 - t^2/5).$$

We claim that

$$L_4(y_i) \ge \frac{7}{10} p_i q_i. \tag{428}$$

By Eq. (427), it suffices to prove

$$\frac{\frac{9}{10}(1-t^2/5)}{(1-t/10)(1+t)(1-t^2)^{11/20}} \ge \frac{7}{10} \qquad (0 \le t < 1).$$

Convexity gives $(1-t^2)^{-1/20} \ge 1 + t^2/20$. Using this lower bound, squaring the two nonnegative sides, and clearing positive denominators reduces the claim to $P(t) \ge 0$, where

$$\begin{aligned} P(t) &:= 81(t^2+20)^2(5-t^2)^2 - 4900(10-t)^2(1+t)^2(1-t^2) \\ &= 320000 - 882000t - 51900t^2 + 970200t^3 + 296025t^4 \\ &\quad - 88200t^5 + 7330t^6 + 81t^8. \end{aligned} \tag{429}$$

The second step follows from simple algebra.

Define $a_0, \dots, a_8$ by $P(t) = \sum_{\ell=0}^{8} a_\ell t^\ell$. On an interval $[u, u+v]$, the coefficients in the degree-eight Bernstein basis are

$$c_j(u,v) := \sum_{k=0}^{j} \frac{\binom{j}{k}}{\binom{8}{k}} v^k \sum_{\ell=k}^{8} a_\ell \binom{\ell}{k} u^{\ell-k} \qquad (0 \le j \le 8).$$

Substitution gives

$[u, u+v]$	$\min_{0 \le j \le 8} c_j(u,v)$
$[0, 1/4]$	$7371989361/65536$
$[1/4, 1/2]$	$809001/256$
$[1/2, 3/4]$	$267273/512$
$[3/4, 1]$	$7383020561/65536$

Every coefficient is positive. Since the Bernstein basis functions are nonnegative and sum to one, this proves $P > 0$ throughout $[0,1]$ and hence Eq. (428).

The strict curvature margin is

$$\frac{165}{16}\frac{7}{10} - 5\frac{1815}{1267} = \frac{2277}{40544} > 0.$$

Eqs. (369) and (370) exclude every stationary light point as a local minimum. The argument of Lemma 11.1 applies with the new parameters. Initialization uses $\beta(0) = 1$, and removal of ρ gives signs satisfying

$$\|\sum_{i=1}^{N} \sigma_i A_i\| \le 2\sqrt{\frac{165}{16} r^{1/4} \alpha}. \tag{430}$$

Schatten orders at least five. For every integer $s \ge 5$, use the square-root budget and set

$$\beta(y) := \sqrt{1-y^2}, \qquad \iota := \frac{81}{40}(s+1), \qquad \vartheta := 3. \tag{431}$$

Then $\gamma = 12$, $\kappa = 18/13$, and lightness gives $\Gamma_i > \iota \ge 12.15 > 12$ and $p_i q_i \le 1/(1-t)$. Moreover,

$$\frac{L_s(y_i)}{p_i q_i} \ge \frac{1 - t^2/(s-1)}{(1+t)\sqrt{1-t^2}} \ge \frac{1 - t^2/3}{(1+t)\sqrt{1-t^2}}.$$

The last function is minimized at $t = \sqrt{7} - 2$: its derivative has the sign of $t^2 + 4t - 3$. Its squared minimum is

$$\frac{40 + 28\sqrt{7}}{243} > (\frac{137}{200})^2.$$

For an exact rational comparison, $\sqrt{7} > 529/200$, and substituting this lower bound leaves a difference of $511/3240000 > 0$. The strict curvature margin per $s+1$ is therefore

$$\frac{81}{40}\frac{137}{200} - \frac{18}{13} = \frac{261}{104000} > 0.$$

The same compact-minimum argument proves that for every integer $s \ge 5$ there are signs satisfying

$$\|\sum_{i=1}^{N} \sigma_i A_i\| \le 2\sqrt{\frac{81}{40}(s+1) r^{1/s} \alpha}. \tag{432}$$

The signing may depend on the chosen order.

The uniform coefficient. We prove Eq. (425). It is enough to consider integer $r \ge 1$; otherwise replace r by $\lfloor r \rfloor$. Every signing has norm at most one, so we only need the bound with coefficient 5.1.

For $1 \le r \le 5$, use Theorem 3.7. Since $\mathrm{tr}[A_i] \le r\alpha$ and $\sum_i A_i = I_D$,

$$\sigma_{\mathrm{tr}}^2 = \|\sum_i \mathrm{tr}[A_i] A_i\| \le r\alpha.$$

Thus the squared coefficient relative to $\alpha \log(2r)$ is at most $(561/50) r/\log(2r)$. For $6 \le r \le 128$, use Eq. (430); the squared coefficient is at most $(165/4) r^{1/4}/\log(2r)$. For $128 \le r \le 512$, use Eq. (432) with $s = 6$; the squared coefficient is at most $(567/10) r^{1/6}/\log(2r)$. Each of these functions decreases and then increases, so its maximum on the stated interval is at an endpoint.

For $r \ge 512$, put $x := \log r$, $a := \log 2$, and $s := \lceil x \rceil + 1$. Then $s \ge 5$, $s + 1 \le x + 3$, and $x/s \le x/(x+1)$. Eq. (432) bounds the squared coefficient by

$$F(x) := \frac{81}{10}\frac{x+3}{x+a} \exp(\frac{x}{x+1}).$$

This function is decreasing for $x \geq 0$, since

$$\frac{F'(x)}{F(x)} = \frac{1}{x+3} - \frac{1}{x+a} + \frac{1}{(x+1)^2} < 0.$$

Indeed, after reversing the sign and clearing positive denominators, the numerator is $(2-a)x^2 + (3-3a)x + 3 - 4a > 0$, using $0 < a < 3/4$. Consequently $F(x) \leq F(9\log 2)$. The resulting upper bounds, rounded upward, are

rank range	squared coefficient
$1 \leq r \leq 5$	24.365
$6 \leq r \leq 128$	25.981
$128 \leq r \leq 512$	23.139
$r \geq 512$	25.561

(433)

All are less than $5.1^2 = 26.01$.

The scalar comparisons follow from the rational bounds

$$0.6931 < \log 2 < 0.6932, \qquad \log 10 > 2.3025, \qquad \log 12 > 2.4849,$$

$$6^{1/4} < 1.56509, \quad 128^{1/4} < 3.3636, \quad 128^{1/6} < 2.24493, \quad 512^{1/6} < 2.82843.$$

The root bounds follow by taking fourth or sixth powers. To verify the logarithmic bounds, put $z = (v-1)/(v+1)$ and use

$$0 \leq \log v - 2\sum_{j=0}^{n-1} \frac{z^{2j+1}}{2j+1} \leq \frac{2z^{2n+1}}{(2n+1)(1-z^2)}$$

with $n = 200$ and $v \in \{2, 10, 12\}$. For $F(9\log 2)$, its rational prefactor decreases with $a = \log 2$ and its exponential argument increases with a. Hence it is at most

$$\frac{81}{10}\frac{9a_- + 3}{10a_-}\exp\left(\frac{9a_+}{9a_+ + 1}\right), \qquad a_- := 0.6931, \quad a_+ := 0.6932.$$

For $0 \leq z < 1$, the exponential remainder satisfies

$$0 \leq e^z - \sum_{j=0}^{n} \frac{z^j}{j!} \leq \frac{z^{n+1}}{(n+1)!}\frac{1}{1 - z/(n+2)}.$$

Taking $n = 24$ verifies the last row of Eq. (433); substitution of the displayed rational bounds verifies the other rows. This proves the intermediate bound in Eq. (425).

11.3 The rank-dependent diagonal estimate and the coefficient 4.307

At an outer stationary point, the mixed term uses one response direction. The negative diagonal response has a rank-dependent lower bound in magnitude. Retaining both properties sharpens the order-three estimate.

Directional energy and the negative diagonal. Take $s = 3$ and define $g := \iota$. Retain the notation u_i, v_i, w_i, z_i, p_i, q_i, c_i, d_i, S_i, and $\Gamma_i = (g\beta_i u_i v_i)^{-1}$ from the response argument. Let ξ be the normalized response vector and T its normalized operator. Put

$$E_{a,i} = c_i^2 u_i w_i, \qquad E_{b,i} = d_i^2 v_i z_i, \qquad e_i = \frac{(\sqrt{E_{a,i}}, -\sqrt{E_{b,i}})}{\sqrt{E_{a,i} + E_{b,i}}}, \quad f_i = \frac{(\sqrt{E_{b,i}}, \sqrt{E_{a,i}})}{\sqrt{E_{a,i} + E_{b,i}}}.$$

Outer stationarity gives $p_i w_i = q_i z_i$, hence $c_i w_i = d_i z_i$ and $E_{a,i} + E_{b,i} = \beta_i S_i$. In particular,

$$(\sqrt{w_i/u_i}, \sqrt{z_i/v_i}) \cdot e_i = 0, \qquad m_i := e_i^\top (I - T)\xi = \sqrt{p_i q_i S_i}\, h_i.$$

The mixed term in Eq. (341) is therefore $2\lambda_i m_i f_i^\top T\xi$, where $\lambda_i = \beta_i' \sqrt{g\Gamma_i/(\beta_i p_i q_i)}$. Choose $0 < \mu_i < 1/2$, put $\eta_i = 2 - 4\mu_i$, and apply Young's inequality only in this direction. With $L_i = -\beta_i'' - (\beta_i')^2/(\eta_i \beta_i)$, we obtain

$$D^2 P_\rho[h,h] \le 4[Q - \sum_i \Gamma_i (\frac{1}{2}(e_i^\top T\xi)^2 + \mu_i (f_i^\top T\xi)^2)] - g\sum_i L_i S_i h_i^2. \tag{434}$$

No derivative of μ_i enters this estimate. For $C = A_i^{1/2} U A_i^{1/2}$, the rank bound gives $\mathrm{tr}[C^4] \ge r^{-1/3} u_i^4$. Consequently

$$\mathsf{U}_{ii} = g c_i^2 \frac{\mathrm{tr}[C^4]}{u_i^2} \ge g c_i^2 u_i^2 r^{-1/3},$$

and the same inequality holds for V_{ii}. Substitution in the normalized pair, using $E_{a,i} + E_{b,i} = \beta_i S_i$, yields

$$e_i^\top T e_i \le -r^{-1/3}/\Gamma_i. \tag{435}$$

The projection estimate. Use the normalized geometry of Section 8.4, including $D_0 \succeq T^\top \Gamma T$, $e_i^\top D_0 e_i = f_i^\top D_0 f_i = 1$, $f_i^\top (I - T)\xi = 0$ on the response graph, and $e_i^\top T e_i + f_i^\top T f_i = 0$. Fix $0 < \chi \le r^{-1/3}$. Suppress the pair index and define $\gamma := \Gamma_i$, $\mu := \mu_i$, and $m := 1/2$. Define

$$\gamma = \frac{\theta(\theta+1)}{1+\mu\theta}, \qquad a = \frac{2(\theta+1)}{\theta+m}, \qquad b = \frac{2\gamma}{\theta},$$

$$w_e = \frac{\theta}{(\theta+1)(1+a^2/(4\gamma))}, \quad w_f = \frac{1}{(\theta+1)(1+b^2/(4\gamma))},$$

$$h = w_e a, \quad j = w_f b, \quad \delta = \frac{\chi(h-j)}{\gamma}, \quad F_e = \frac{1-\delta}{w_e}.$$

Set

$$K = \frac{a^2(1-m)/4 + m\gamma F_e}{m\gamma + a + a^2/(4\gamma) - F_e}. \tag{436}$$

We use the negative diagonal estimate in Eq. (435) to sharpen the covariance bound on the response graph.

Lemma 11.2 (Rank-dependent projection)**.** *Assume the normalized response geometry of Section 8.4,* $\min_i \Gamma_i > 2$*, and the scalar domain in Eq. (442), on which* $0 < \delta_i < 1$ *and* $K_i > 0$*. There is a covariance* $\mathcal{A}$*, positive definite on the response graph, such that*

$$\mathrm{tr}[D_0 \mathcal{A}] - \sum_i \Gamma_i (m e_i^\top T \mathcal{A} T^\top e_i + \mu_i f_i^\top T \mathcal{A} T^\top f_i)$$
$$\le \sum_i K_i e_i^\top (I - T)\mathcal{A}(I - T)^\top e_i. \tag{437}$$

The sum of the unweighted movement terms is positive.

Proof. Direct substitution gives

$$\mu\gamma + b + \frac{b^2}{4\gamma} = \frac{1}{w_f}, \qquad w_e + w_f + \frac{h^2}{4\gamma w_e} + \frac{j^2}{4\gamma w_f} = 1,$$

and Eq. (436) is equivalent to

$$K - \frac{(K - a/2)^2}{m\gamma + K} + \frac{a^2}{4\gamma} = F_e.$$

Divide w_e, w_f, h, j by $1 - \delta$, denoting the resulting quantities by $\widetilde{w}_e, \widetilde{w}_f, \widetilde{h}, \widetilde{j}$. Let D have these weights in the e_i, f_i basis, and, for an orthonormal basis R of the response graph, set

$$\mathcal{A} = R(R^\top D^{-1} R)^{-1} R^\top, \qquad \mathcal{B} = D - \mathcal{A}.$$

Then $\mathcal{A}, \mathcal{B} \succeq 0$, $\mathcal{A}D^{-1}\mathcal{B} = 0$, and $\operatorname{tr}[D^{-1}\mathcal{A}]$ equals the number of active coordinates. Complete the e_i row with $C_i = \gamma_i \mathcal{B} + (m\gamma_i + K_i)\mathcal{A}$, exactly as in Eq. (374). The coefficient of $e_i^\top \mathcal{A} e_i$ is $F_{e,i} = 1/\widetilde{w}_{e,i}$. On the f_i row use $f_i^\top T \mathcal{A} = f_i^\top \mathcal{A}$; the coefficient is $1/w_{f,i} \geq 1/\widetilde{w}_{f,i}$. The constant subtracted by the two completions is

$$S_i^* = \frac{\widetilde{h}_i^2}{4\gamma_i \widetilde{w}_{e,i}} + \frac{\widetilde{j}_i^2}{4\gamma_i \widetilde{w}_{f,i}}.$$

By Eq. (435), the added linear term satisfies

$$\widetilde{h}_i e_i^\top T e_i + \widetilde{j}_i f_i^\top T f_i \leq -\frac{\delta_i}{1 - \delta_i}.$$

The normalization identity is $\widetilde{w}_{e,i} + \widetilde{w}_{f,i} + S_i^* - \delta_i/(1 - \delta_i) = 1$. Sum the row completions, use $\operatorname{tr}[D_0 \mathcal{B}] \geq \operatorname{tr}[\Gamma T \mathcal{B} T^\top]$, and cancel these constants. This proves Eq. (437). Positivity of the movement sum follows from the response-graph invertibility argument in Section 8.4, since $\min_i \Gamma_i > 2$. □

The two order-three budgets. The budgets must satisfy the negative-curvature condition obtained from Eq. (434) and Lemma 11.2:

$$gL_i > 4p_i q_i K_i \tag{438}$$

for every active coordinate. Put $t = |y|$, $z = 1 - t^2$, and $\beta(y) = \sqrt{z} f(t)$, where

$$\begin{aligned} f(t) := 1 &+ 0.07906t^2 - 0.16929t^4 + 1.05091t^6 - 4.52847t^8 \\ &+ 11.09191t^{10} - 15.75523t^{12} + 11.97023t^{14} - 3.81974t^{16}. \end{aligned} \tag{439}$$

The directional energy weight is

$$\mu(t) := \frac{1}{1000} + \frac{498}{1000}(1 - t)^2. \tag{440}$$

All decimals denote exact rationals. Use

$$(g, \chi) = (5.76, 69/100) \quad (r = 3), \qquad (g, \chi) = (5.9, 5/8) \quad (r = 4). \tag{441}$$

The rank estimates hold because $3(69/100)^3 < 1$ and $4(5/8)^3 < 1$. Let $\eta = 2 - 4\mu$, $N = zf' - tf$, $A = f - zf'$, and $H = -\eta f(zN' + tN) - N^2$. Then $\beta' = N/\sqrt{z}$ and $L = H/(\eta z^{3/2} f)$. The ordinary-degree Bernstein coefficients on $[0,1]$ certify positivity of

$$f, \quad -N/t, \quad (f^2 - z)/t^2, \quad f + zf', \quad A, \quad \eta, \quad H.$$

The quotients have their polynomial continuations at zero. These inequalities give $1-y^2 \le \beta(y) \le 1$, positive response scales, and $\eta \ge 1/250$. Symmetry exchanging U, V reduces the analysis to $y \ge 0$.

Normalize lightness by $\alpha_{\mathrm{loc}} = g\beta u/(1+t)$ and $\nu = g\beta v/(1-t)$. At fixed product $d = \alpha_{\mathrm{loc}}\nu$, Γ is fixed, whereas

$$pq = (1 + B\alpha_{\mathrm{loc}})(1 + Cd/\alpha_{\mathrm{loc}}), \qquad B = -\frac{(1+t)\beta'}{\beta} \ge 0, \quad C = \frac{(1-t)\beta'}{\beta} \le 0.$$

Its derivative in α_{loc} is $B - Cd/\alpha_{\mathrm{loc}}^2 \ge 0$. Since K depends on the fixed product through Γ, the supremum is at $\alpha_{\mathrm{loc}} = 1$, $\nu = d$. Put $x = 1/\theta$, so $\Gamma = (1+x)/(x(x+\mu))$. The physical domain satisfies

$$G(t,x) := g^2 f^2 x^2 (x+\mu)^2 - z(1+x)^2 \le 0. \tag{442}$$

The Bernstein check $G(t, 1/2) > 0$, together with monotonicity of $x(x+\mu)/(1+x)$, implies $0 < x < 1/2$ on this domain. On the enclosing rectangle, $\Gamma \ge 3$, $2 \le a \le 12/5$, and $1/w_e \le 111/50$. Define

$$A_1 = 1 + x/2, \quad X = x(x+\mu), \quad E = A_1^2 + X(1+x), \quad F = \mu + 2x + x^2, \quad R = A_1 F - xE.$$

Here

$$R = \mu(1 + x/2 - x^2 - x^3) + x(1 + x - 3x^2/4 - x^3) > 0, \qquad h - j = 2R/(EF).$$

Also $h \le 2$, so $0 < \delta \le 2\chi/3 < 1$. The denominator in Eq. (436) is at least $3/2+2-111/50 = 32/25$; its numerator is positive.

The scalar certificate. To specify the exact polynomial, retain $m = 1/2$ and put

$$B_* = (1+x)(1-m)FX + m((1+x)EF - 2\chi XR),$$

$$D_* := m(1+x)A_1^2 F + (1+x)(2A_1 - A_1^2)FX - x(1+x)X^2 F + 2\chi X^2 R.$$

Set $K_* := (1+x)B_*$. Substitution gives $K = K_*/D_*$, with positive denominators. At the maximizing pair, $q = A/((1-t)f)$ and $p = 1 + gN/((1+t)\sqrt{z}\,\Gamma)$. Thus Eq. (438) follows from

$$W := g(HD_* - 4\eta ANB_* X) > 4\eta A(1+t)\sqrt{z}\,K_*.$$

Since $W > 0$, it suffices to certify

$$P(t,x) := W^2 - 16\eta^2 A^2 (1+t)^2 z K_*^2 > 0 \qquad \text{whenever } G(t,x) \le 0. \tag{443}$$

The exact rectangle certificates, in coordinates (t,u) with $x = u/2$, have the following sizes:

r	P leaves	G leaves	total
3	65	52	117
4	12	7	19

Each leaf has strictly positive tensor Bernstein coefficients for the indicated polynomial. The rectangles cover $[0,1]^2$ and have disjoint interiors. A G leaf contains no admissible point, so the certificate proves Eq. (443) on the full admissible domain. The polynomial P has bidegree $(84,16)$. Appendix B.2 specifies every rectangle by a recursive subdivision rule and gives the exact Bernstein coefficient formula. Affine substitution into the displayed polynomials verifies every leaf using rational arithmetic. The same appendix specifies the univariate budget checks; Appendix B.7 gives the rational bounds used in the rank comparisons below.

Strict curvature follows at every stationary light state. The compact-minimum argument and removal of regularization therefore give

$$\|\sum_i \sigma_i A_i\| \le \begin{cases} 2\sqrt{5.76\,3^{1/3}\alpha}, & r=3, \\ 2\sqrt{5.9\,4^{1/3}\alpha}, & r=4. \end{cases} \tag{444}$$

The uniform coefficient. For ranks one and two use Theorem 3.7; for ranks three and four use Eq. (444). Use Eq. (385) for $5 \le r \le 16$, Eq. (386) at orders five and eight for $16 \le r \le 256$ and $256 \le r \le 4096$, respectively, and $s = \lceil \log r \rceil + 1$ for $r \ge 4096$. The endpoint comparisons and the decreasing envelope from Section 11.2 give the following upper bounds on the squared coefficient normalized by $\alpha \log(2r)$:

rank range	squared coefficient upper bound
$1 \le r \le 2$	112200211369/6931000000
$r = 3$	1661472/89585
$r = 4$	538257/29875
$5 \le r \le 16$	17527254/959375
$16 \le r \le 256$	540342/29875
$256 \le r \le 4096$	120000/6931
$r \ge 4096$	18.401

(445)

Supplement the rational logarithm and root bounds in Section 11.2 with $\log 10 > 2.3025$, $3^{1/3} < 1.44225$, $5^{1/3} < 1.709976$, $16^{1/3} < 2.519843$, and $16^{1/5} < 1.741102$. Integer powers verify the root bounds and the same logarithmic and exponential series verify the other comparisons. The largest displayed upper bound is $1661472/89585$, which is less than $4.307^2 = 18.550249$. Nonintegral rank bounds reduce to their integer parts.

11.4 Normalizing the projection weights and the coefficient 4.225

We incorporate the magnitude of the negative diagonal response directly into the normalization of the projection weights. Two order-three budgets then improve the existence coefficient to 4.225. The next subsection refines this comparison by retaining the joint energy–moment bound.

The retained response inequalities. Work at a stationary light state with Schatten order three. Let g denote the coupling and retain the response notation

$$u_i = f_i(U), \quad v_i = f_i(V), \quad w_i = \operatorname{tr}[A_i W], \quad z_i = \operatorname{tr}[A_i Z],$$

$$S_i = v_i w_i + u_i z_i, \qquad p_i = 1 + g\beta_i' v_i, \quad q_i = 1 - g\beta_i' u_i, \qquad \Gamma_i = (g\beta_i u_i v_i)^{-1}.$$

Let ξ be the normalized response and T its operator. For each active coordinate there is an orthonormal pair e_i, f_i. Outer stationarity gives $p_i w_i = q_i z_i$, so the mixed functional–budget term

uses only the f_i-response. Consequently, with

$$m = \tfrac{1}{2}, \qquad 0 < \mu_i < \tfrac{1}{2}, \qquad \eta_i = 2 - 4\mu_i, \qquad L_i = -\beta_i'' - \frac{(\beta_i')^2}{\eta_i \beta_i},$$

the directional estimate from Eq. (434) is

$$D^2 P_\rho[h,h] \le 4[Q - \sum_i \Gamma_i (m(e_i^\top T\xi)^2 + \mu_i (f_i^\top T\xi)^2)] - g \sum_i L_i S_i h_i^2. \tag{446}$$

The movement identity is

$$e_i^\top (I - T)\xi = \sqrt{p_i q_i S_i}\, h_i. \tag{447}$$

Each μ_i is a pointwise parameter in the directional estimate.

The rank-dependent diagonal estimate is

$$e_i^\top T e_i \le -r^{-1/3}/\Gamma_i. \tag{448}$$

It follows from $\mathrm{tr}[C^4] \ge r^{-1/3}\|C\|_{S_3}^4$ for the positive semidefinite compression $C = A_i^{1/2} U A_i^{1/2}$, and the analogous inequality for V, substituted into the normalized response.

The remaining normalized geometry is

$$D_0 \succeq T^\top \Gamma T, \qquad \Gamma = \sum_i \Gamma_i (e_i e_i^\top + f_i f_i^\top),$$

$$e_i^\top D_0 e_i = f_i^\top D_0 f_i = 1, \qquad e_i^\top T e_i + f_i^\top T f_i = 0, \qquad f_i^\top (I - T)\xi = 0$$

on the response graph, whose dimension equals the number of active coordinates. These are the normalization and diagonal padding of Section 8.4.

Normalization of the projection weights. Fix $0 < \chi \le r^{-1/3}$. Suppress the pair index and put $\gamma := \Gamma_i$, $\mu := \mu_i$. Choose positive row parameters a, b, and define

$$N_e = 1 + \frac{a^2}{4\gamma} - \frac{\chi a}{\gamma}, \qquad N_f = 1 + \frac{b^2}{4\gamma} + \frac{\chi b}{\gamma},$$

$$F_f = \mu\gamma + b + \frac{b^2}{4\gamma}, \qquad F_e = \frac{N_e F_f}{F_f - N_f}, \qquad w_e = F_e^{-1}, \quad w_f = F_f^{-1}. \tag{449}$$

Finally put

$$K = \frac{a^2(1-m)/4 + m\gamma F_e}{m\gamma + a + a^2/(4\gamma) - F_e}, \qquad m = \tfrac{1}{2}. \tag{450}$$

The scalar tests in Appendix B.2 establish positive denominators and $w_e a > w_f b$ throughout the admissible domain.

The essential identity is

$$w_e N_e + w_f N_f = 1. \tag{451}$$

It incorporates the magnitude of the negative diagonal term directly into the normalization. Set $h := w_e a$, $j := w_f b$. Eq. (451) says exactly

$$w_e + w_f + \frac{h^2}{4\gamma w_e} + \frac{j^2}{4\gamma w_f} - \frac{\chi(h-j)}{\gamma} = 1. \tag{452}$$

The normalization in Eq. (452) absorbs the negative diagonal contribution in the row completions. The next lemma gives the resulting covariance estimate.

Lemma 11.3 (Normalized projection weights)**.** *Under the normalized response geometry and weight hypotheses just specified, there is a covariance $\mathcal{A}$, positive definite on the response graph and zero on its orthogonal complement, such that*

$$\begin{aligned}\operatorname{tr}[D_0\mathcal{A}] - \sum_i \Gamma_i (m e_i^\top T\mathcal{A}T^\top e_i + \mu_i f_i^\top T\mathcal{A}T^\top f_i) \\ \le \sum_i K_i e_i^\top (I-T)\mathcal{A}(I-T)^\top e_i.\end{aligned} \tag{453}$$

The sum of the unweighted movement covariances is positive.

Proof. Let D be diagonal in the e_i, f_i basis with weights $w_{e,i}, w_{f,i}$. If R is an orthonormal basis of the response graph, set

$$\mathcal{A} = R(R^\top D^{-1} R)^{-1} R^\top, \qquad \mathcal{B} = D - \mathcal{A}.$$

Then $\mathcal{A}, \mathcal{B} \succeq 0$, $\mathcal{A}D^{-1}\mathcal{B} = 0$, and $\operatorname{tr}[D^{-1}\mathcal{A}] = k$, where k is the number of active coordinates.

Eq. (450) is equivalent to

$$K - \frac{(K - a/2)^2}{m\gamma + K} + \frac{a^2}{4\gamma} = F_e = 1/w_e. \tag{454}$$

For the row $t = e_i^\top T$, complete the square with

$$C_i = \gamma_i \mathcal{B} + (m\gamma_i + K_i)\mathcal{A}.$$

The identity

$$C_i^{-1} = \gamma_i^{-1} D^{-1} + ((m\gamma_i + K_i)^{-1} - \gamma_i^{-1}) D^{-1}\mathcal{A}D^{-1}$$

and Eq. (454) give

$$\begin{aligned}\gamma_i t\mathcal{B}t^\top + m\gamma_i t\mathcal{A}t^\top + K_i(e_i^\top - t)\mathcal{A}(e_i - t^\top) + h_i t e_i \\ \ge \frac{e_i^\top \mathcal{A} e_i}{w_{e,i}} - \frac{h_i^2}{4\gamma_i w_{e,i}}.\end{aligned}$$

For the row $t = f_i^\top T$, the graph identity gives $t\mathcal{A} = f_i^\top \mathcal{A}$. The analogous square completion, using $F_f = 1/w_f$, yields

$$\gamma_i t\mathcal{B}t^\top + \mu_i\gamma_i t\mathcal{A}t^\top + j_i t f_i \ge \frac{f_i^\top \mathcal{A} f_i}{w_{f,i}} - \frac{j_i^2}{4\gamma_i w_{f,i}}.$$

Because $h_i > j_i$, Eq. (448) bounds the added linear term by

$$h_i e_i^\top T e_i + j_i f_i^\top T f_i \le -\chi(h_i - j_i)/\gamma_i.$$

Summing and using Eq. (452), the quadratic terms without the added linear term are at least $\sum_i (w_{e,i} + w_{f,i}) = \operatorname{tr}[D_0 D]$. Since

$$\operatorname{tr}[\Gamma T\mathcal{B}T^\top] \le \operatorname{tr}[D_0\mathcal{B}],$$

subtraction proves Eq. (453).

If every movement term vanished, T would act as the identity on the response graph, giving $\|T\|_F^2 \ge k$. But coercivity gives $(\min_i \Gamma_i)\|T\|_F^2 \le \operatorname{tr}[D_0] = 2k$, contradicting $\min_i \Gamma_i > 2$. This proves positivity of the movement sum. □

Explicit budgets and row parameters. To choose the budgets, combine the directional and movement estimates in Eqs. (446) and (447) with Lemma 11.3. Strict negative curvature follows if

$$gL_i > 4p_iq_iK_i. \tag{455}$$

We use two order-three budgets to satisfy this condition. In each case put $t := |y|$, $z := 1 - t^2$, and

$$\beta(y) = \sqrt{z}\, f(t), \qquad f(t) = 1 + \sum_{j=1}^{d} c_j t^{2j}.$$

All decimals below are exact terminating rationals.

Parameter	Rank 3	Ranks 4 through 18
Coupling g	5.5435	5.82
Rank contribution χ	0.6933	0.3815
Row shift a_c	0.806	0.841
Row shift b_c	0.8793	0.7453

The coefficient lists $(c_1, \dots, c_d)$ are

$$\begin{aligned} \text{rank 3: } & (0.099459, -0.139141, 0.690006, -4.105105, \\ & 16.787172, -45.582901, 79.520157, -85.643210, \\ & 51.717346, -13.435034), \end{aligned}$$

$$\begin{aligned} \text{ranks 4–18: } & (0.072386, -0.141118, 0.861534, -3.843893, \\ & 9.668814, -14.068250, 10.913567, -3.548775). \end{aligned}$$

Specify $\mu(t)$ in the ordinary Bernstein basis:

$$\mu(t) = \sum_{j=0}^{q} m_j \binom{q}{j} t^j (1-t)^{q-j}.$$

For rank three, $q = 6$ and

$$(m_0, \dots, m_6) = (0.4998, 0.411076, 0.228422, 0.0002, 0.0002, 0.0002, 0.0002).$$

For ranks four through eighteen, $q = 4$ and

$$(m_0, \dots, m_4) = (0.4998, 0.343273, 0.0002, 0.0002, 0.0002).$$

Thus $0 < \mu \le 0.4998$ and $\eta = 2 - 4\mu \ge 1/1250$. The rank estimates hold because $3(0.6933)^3 < 1$ and $18(0.3815)^3 < 1$.

Define $x > 0$ by

$$\gamma = \frac{1 + x}{x(x + \mu)}.$$

The row parameters in Eq. (449) are

$$a = \frac{2(1 + x)}{1 + x/2} + a_c\chi, \qquad b = 2\gamma x - b_c\chi. \tag{456}$$

The scalar test uses these row shifts together with the normalization in Eq. (449).

Exact scalar reduction. Let

$$N = zf' - tf, \quad A_0 = f - zf', \quad H = -\eta f(zN' + tN) - N^2.$$

Then $\beta' = N/\sqrt{z}$ and $L = H/(\eta z^{3/2} f)$. Positive Bernstein coefficients on $[0,1]$ certify

$$f > 0, \quad -N/t > 0, \quad (f^2 - z)/t^2 > 0, \quad f + zf' > 0, \quad A_0 > 0, \quad \eta > 0, \quad H > 0. \tag{457}$$

The quotients have their polynomial continuations at zero. These give the budget bounds $1 - y^2 \le \beta(y) \le 1$, the endpoint-slope conditions and positivity of the response scales.

As in Section 11.3, introduce the light-state parameters

$$\alpha_{\mathrm{loc}} = \frac{g\beta u}{1+t}, \qquad \nu = \frac{g\beta v}{1-t}, \qquad 0 < \alpha_{\mathrm{loc}}, \nu < 1.$$

Fix $d := \alpha_{\mathrm{loc}}\nu$. Then Γ and K are fixed. Define

$$B = -\frac{(1+t)\beta'}{\beta} \ge 0, \quad C = \frac{(1-t)\beta'}{\beta} \le 0,$$

Then $pq = (1 + B\alpha_{\mathrm{loc}})(1 + Cd/\alpha_{\mathrm{loc}})$ has derivative $B - Cd/\alpha_{\mathrm{loc}}^2 \ge 0$. Thus it suffices to check $\alpha_{\mathrm{loc}} = 1$, $\nu = d$.

Define $J := 1 + x$, $D := x + \mu$, and $X := xD$. The physical domain is

$$G(t,x) := g^2 f^2 X^2 - zJ^2 \le 0. \tag{458}$$

The exact inequality $G(t, 1/2) > 0$ and monotonicity of $x(x+\mu)/(1+x)$ imply $0 < x < 1/2$. Therefore it suffices to certify the rectangle $[0,1] \times [0,1/2]$, excluding points with $G > 0$.

The following polynomials specify the scalar tests. Define

$$B_0 = 1 + x/2, \quad A_1 = 2J + a_c\chi B_0, \quad C_1 = 2J - b_c\chi D,$$

$$E = 4JB_0^2 + XA_1^2 - 4\chi X A_1 B_0,$$

$$U_0 = 4\mu J^2 + 4JxC_1 + x^2 C_1^2,$$

$$V_0 = \mu J^2 + Jx(C_1 - D) - \chi x^2 C_1 D,$$

$$K_N = J(4XA_1^2 V_0 + EU_0),$$

$$K_D = 2(8J^2 B_0^2 V_0 + 16JXA_1 B_0 V_0 + 4X^2 A_1^2 V_0 - XEU_0).$$

Substitution into Eqs. (449)–(450) gives

$$F_f = \frac{U_0}{4JX}, \qquad F_e = \frac{EU_0}{16JB_0^2 V_0}, \qquad K = \frac{K_N}{K_D}. \tag{459}$$

The polynomial

$$H_J = 4A_1 B_0 V_0 - xC_1 E$$

controls the sign of $w_e a - w_f b$, since

$$w_e a - w_f b = \frac{4JH_J}{EU_0}.$$

The tensor Bernstein coefficients of V_0, K_D, H_J are strictly positive on the entire rectangle, as specified in Appendix B.2. Also $E > 0$, since $N_e \ge 1 - \chi^2/\gamma > 0$ and $\gamma \ge 3$. We have $C_1 > 0$,

and hence $U_0 > 0$. All the denominator signs and weight hypotheses used in the projection lemma are therefore established. The formulas at $x = 0$ are interpreted as their polynomial limits; actual light states have $x > 0$.

At the maximizing light-state pair,

$$q = \frac{A_0}{(1-t)f}, \qquad p = 1 + \frac{gNX}{(1+t)\sqrt{z}\,J}.$$

Thus Eq. (455) follows from

$$W := g(JHK_D - 4\eta A_0 NXK_N) > 4\eta A_0(1+t)\sqrt{z}\,JK_N.$$

The left-hand side is positive by Eq. (457), the denominator certificates and $N \le 0$. Consequently we may square and certify

$$P(t,x) := W^2 - 16\eta^2 A_0^2(1+t)^2 z J^2 K_N^2 > 0 \quad \text{where } G(t,x) \le 0. \tag{460}$$

The certificates use rectangle coordinates (t,u), with $x = u/2$:

Branch	Bidegree of $P(t,x)$	$P>0$ leaves	$G>0$ leaves	Total
Rank 3	(144,22)	27	18	45
Ranks 4–18	(108,22)	57	40	97

In each branch the rectangles cover the unit square with disjoint interiors. At every leaf all tensor Bernstein coefficients of the named polynomial are strictly positive. A G leaf contains no physical point; hence Eq. (460) holds throughout the physical domain. The verification is rational throughout.

Direct affine substitution into the global monomial polynomials independently verifies every leaf using integer arithmetic after clearing denominators. The same exact checks establish the univariate budget conditions, positivity of V_0, K_D, H_J in each branch, and the row identities.

The directional estimate in Eq. (446), Lemma 11.3, and the strict inequality in Eq. (455) give negative curvature at every stationary light state. The compact-minimum proof of Lemma 11.1 reaches a vertex. Removing regularization gives

$$\|\sum_i \sigma_i A_i\| \le \begin{cases} 2\sqrt{5.5435\,3^{1/3}\alpha}, & r = 3, \\ 2\sqrt{5.82\,r^{1/3}\alpha}, & 4 \le r \le 18. \end{cases} \tag{461}$$

All-rank comparison and the sharper large-rank envelope. Use Theorem 3.7 for ranks one and two, Eq. (461) for ranks three through eighteen, and Eq. (386)

$$\|\sum_i \sigma_i A_i\| \le \sqrt{6(s+1)r^{1/s}\alpha}, \qquad s \ge 5,$$

for the remaining ranges. Choose order five for $19 \le r \le 256$, order eight for $256 \le r \le 4096$, and $s = \lceil \log r \rceil + 1$ thereafter. For a fixed order, $r^{1/s}/\log(2r)$ decreases and then increases, so compact intervals require only endpoint comparisons.

For the infinite range define $x := \log r$ and $a := \log 2$. Since $x + 1 \le s \le x + 2$, and

$$\frac{d}{ds}\log((s+1)e^{x/s}) = \frac{1}{s+1} - \frac{x}{s^2} > 0 \quad (s \ge x+1),$$

we have the sharper joint estimate

$$\frac{6(s+1)e^{x/s}}{x+a} \le F_2(x) := 6\frac{x+3}{x+a}\exp(\frac{x}{x+2}). \tag{462}$$

The sign of $F_2'(x)$ is the sign of

$$(a-1)x^2 + (6a-6)x + (10a-12),$$

which is negative for $x \ge 0$, since $0 < a < 3/4$. Hence the infinite range is bounded by $F_2(12\log 2)$.

The rank comparisons use the following rational inequalities:

$$0.6931 < \log 2 < 0.6932, \quad \log 6 > 1.7917, \quad \log 38 > 3.6375,$$

$$3^{1/3} < 1.44225, \quad 4^{1/3} < 1.587402, \quad 18^{1/3} < 2.620742,$$

$$19^{1/5} < 1.801984, \quad 256^{1/5} < 3.031434, \quad 4096^{1/8} < 2.828428.$$

Integer powers verify the root bounds. The logarithmic and exponential series in Appendix B.7 verify the remaining comparisons. The resulting squared coefficient bounds, rounded upward, are as follows.

Rank range	Squared coefficient upper bound
1–2	16.189
3	17.850
4–18	17.773
19–256	17.835
256–4096	17.314
At least 4096	16.876

Each resulting squared coefficient is strictly below

$$4.225^2 = 17.850625.$$

For rank three, the coefficient before outward rounding is $2\sqrt{5.5435\,3^{1/3}/\log 6}$, approximately 4.225. Nonintegral rank upper bounds reduce to their integer parts. Since positive semidefiniteness and isotropy give $-I \preceq \sum_i \sigma_i A_i \preceq I$ for every signing, the cap by one gives discrepancy at most $\min\{1, 4.225\sqrt{\alpha\log(2r)}\}$.

The subdivision rules and exact coefficient tests are given in Appendix B.2. They verify the scalar hypotheses of the projection lemma from the polynomials displayed here. Appendix B.7 specifies the rational comparisons for the numerical constants.

11.5 Joint response energy and the coefficient 4.171

We refine the order-three comparison by retaining the dependence between the diagonal response energy and the fourth moment of the normalized compressed matrix. This gives coupling 5.403 for rank at most three and 5.65 for rank at most thirty-two. Together with the higher-order estimate in Eq. (386), these bounds give the existence coefficient 4.171. The argument uses exact outer stationarity. The deterministic construction in Section 12 uses coefficient 4.225 with finite-accuracy computation. Section 11.7 further improves the existence bound.

The retained response geometry. Fix Schatten order three and coupling $g > 2$. For a nonzero input define

$$\phi_i(X) = \big(\mathrm{tr}[(A_i^{1/2} X A_i^{1/2})^3]\big)^{1/3}, \qquad B_i(X) = \frac{A_i^{1/2}(A_i^{1/2} X A_i^{1/2})^2 A_i^{1/2}}{\phi_i(X)^2}.$$

For an even budget β with $\beta(0) = 1$, $\beta(\pm 1) = 0$, and $\beta > 0$ on $(-1, 1)$, use

$$F_y(X) = g \sum_i \beta(y_i)\phi_i(X) A_i, \qquad M(y) = \sum_i y_i A_i,$$

with already fixed contributions retained in M on a face. The regularized potential is

$$P_\rho(y) = \min_{b \in \mathbb{R},\ U,V \succ 0} \left\{ b + \rho\,\mathrm{tr}[U+V] : \begin{array}{l} U^{-1} + M(y) + F_y(V) \preceq bI, \\ V^{-1} - M(y) + F_y(U) \preceq bI \end{array} \right\}. \tag{463}$$

The proof of Lemma 8.3 applies to the budget verified in the scalar certificate below. It gives a continuous value on the cube and a unique smooth optimizer on each open face. Write

$$u_i = \phi_i(U), \quad v_i = \phi_i(V), \quad w_i = \mathrm{tr}[A_i W], \quad z_i = \mathrm{tr}[A_i Z],$$

where $W, Z \succ 0$ are the dual multipliers. They obey

$$W = U\left(\rho I + g\sum_i \beta_i z_i B_i(U)\right) U,$$
$$Z = V\left(\rho I + g\sum_i \beta_i w_i B_i(V)\right) V, \qquad \mathrm{tr}[W+Z] = 1.$$

An endpoint move is available unless

$$g\beta_i u_i < 1 + y_i, \qquad g\beta_i v_i < 1 - y_i. \tag{464}$$

At a light state put

$$p_i = 1 + g\beta_i' v_i, \qquad q_i = 1 - g\beta_i' u_i,$$
$$c_i = \sqrt{\beta_i p_i / q_i}, \quad d_i = \sqrt{\beta_i q_i / p_i}, \quad \tau_i = \sqrt{p_i q_i / \beta_i}, \quad \gamma_i = (g\beta_i u_i v_i)^{-1}.$$

All these scales are positive. At an outer stationary point,

$$p_i w_i = q_i z_i, \qquad c_i w_i = d_i z_i. \tag{465}$$

For a physical direction h, fixed-b responses can be written as

$$\dot U = U M_a U, \quad \dot V = V M_b V, \qquad M_a = \sum_i c_i a_i^{\mathrm{resp}} A_i, \quad M_b = \sum_i d_i b_i^{\mathrm{resp}} A_i.$$

The response coefficients satisfy

$$a^{\mathrm{resp}} - \mathcal{V} b^{\mathrm{resp}} = \mathcal{U} a^{\mathrm{resp}} - b^{\mathrm{resp}} = \tau h,$$

where

$$\mathcal{U}_{ij} = g c_i c_j\,\mathrm{tr}[B_i(U) U A_j U], \qquad \mathcal{V}_{ij} = g d_i d_j\,\mathrm{tr}[B_i(V) V A_j V].$$

The transpose dual inequalities in Eq. (359) make this a k-dimensional graph when k coordinates are active. Set

$$E_a = \operatorname{diag}(c_i^2 u_i w_i), \quad E_b = \operatorname{diag}(d_i^2 v_i z_i), \quad E = \operatorname{diag}(E_a, E_b),$$
$$\xi = E^{1/2}(a^{\mathrm{resp}}, b^{\mathrm{resp}}), \qquad T = \begin{pmatrix} 0 & E_a^{1/2}\mathcal{V}E_b^{-1/2} \\ E_b^{1/2}\mathcal{U}E_a^{-1/2} & 0 \end{pmatrix}.$$

The response energy is

$$Q = \operatorname{tr}[W M_a U M_a] + \operatorname{tr}[Z M_b V M_b] = \xi^\top D_0 \xi.$$

Here D_0 is the normalized energy matrix before the diagonal padding in Section 8.5. The dual equations and

$$(D\phi_i(X)[H])^2 \le \phi_i(X)\operatorname{tr}[B_i(X) H X^{-1} H]$$

give

$$D_0 \succeq T^\top \Gamma T, \qquad \Gamma = \operatorname{diag}(\gamma_1, \dots, \gamma_k, \gamma_1, \dots, \gamma_k). \tag{466}$$

In each pair choose the orthonormal vectors

$$e_i = \frac{(\sqrt{E_{a,i}}, -\sqrt{E_{b,i}})}{\sqrt{E_{a,i} + E_{b,i}}}, \qquad f_i = \frac{(\sqrt{E_{b,i}}, \sqrt{E_{a,i}})}{\sqrt{E_{a,i} + E_{b,i}}}.$$

On the response graph,

$$f_i^\top (I - T)\xi = 0, \qquad Y_i := e_i^\top (I - T)\xi = \sqrt{p_i q_i S_i}\, h_i, \qquad S_i = v_i w_i + u_i z_i. \tag{467}$$

The second identity uses stationarity.

The joint energy–moment inequality.

Lemma 11.4. *Suppose $R \succeq 0$, $\operatorname{tr}[R^3] = 1$, and $\operatorname{rank}(R) \le r$. Put $\kappa = \operatorname{tr}[R^4]$. For every $S \succeq 0$ of trace one,*

$$r^{-1/3} \le \kappa \le 1, \qquad 0 \le e := \operatorname{tr}[RS] \le 1, \qquad e \le \frac{3 + \kappa}{4}. \tag{468}$$

Proof. The eigenvalues of R lie in $[0, 1]$, so $\kappa \le 1$ and $e \le 1$. The power-mean inequality gives $\kappa \ge r^{-1/3}$. For every $s \ge 0$,

$$s^4 - 4s + 3 = (s - 1)^2(s^2 + 2s + 3) \ge 0.$$

Consequently $\lambda_{\max}(R) \le (3 + \lambda_{\max}(R)^4)/4 \le (3 + \kappa)/4$, and $e \le \lambda_{\max}(R)$. □

For a fixed coordinate define

$$R_U = A_i^{1/2} U A_i^{1/2}/u_i, \qquad S_U = A_i^{1/2} W A_i^{1/2}/w_i,$$

with R_V, S_V defined analogously. Let $e_U = \operatorname{tr}[R_U S_U]$, $e_V = \operatorname{tr}[R_V S_V]$ and $\kappa_U = \operatorname{tr}[R_U^4]$, $\kappa_V = \operatorname{tr}[R_V^4]$. Set

$$\alpha_{\mathrm{loc},i} = \frac{c_i u_i}{c_i u_i + d_i v_i} \in (0, 1).$$

The exact local identities are

$$e_i^\top D_0 e_i = \alpha_{\mathrm{loc},i} e_U + (1 - \alpha_{\mathrm{loc},i}) e_V,$$

$$f_i^\top D_0 f_i = (1-\alpha_{\mathrm{loc},i})e_U + \alpha_{\mathrm{loc},i} e_V, \tag{469}$$
$$e_i^\top T e_i = -\frac{\alpha_{\mathrm{loc},i}\kappa_U + (1-\alpha_{\mathrm{loc},i})\kappa_V}{\gamma_i}, \qquad f_i^\top T f_i = -e_i^\top T e_i.$$

For example, $\mathcal{U}_{ii} = g c_i^2 u_i^2 \kappa_U$; substituting this and its V counterpart proves the last line. The diagonal entries of the raw energy before rotation are exactly e_U, e_V.

Fix weights $w_e, w_f > 0$, multipliers $h_0 \geq j_0$, and put $c_0 = (h_0 - j_0)/\gamma$. If $0 < \chi \leq r^{-1/3}$, equations (468)–(469) imply

$$\begin{aligned} &w_e e_i^\top D_0 e_i + w_f f_i^\top D_0 f_i + h_0 e_i^\top T e_i + j_0 f_i^\top T f_i \\ &\qquad \leq w_f + \max\left\{ w_e - c_0,\ \frac{3+\chi}{4} w_e - \chi c_0 \right\}. \end{aligned} \tag{470}$$

Indeed, bound the w_f contribution by w_f and use

$$w_e e - c_0 \kappa \leq \tfrac{3}{4} w_e + (\tfrac{1}{4} w_e - c_0)\kappa, \qquad \chi \leq \kappa \leq 1.$$

The two alternatives in (470) are its endpoint values. Thus the weighted comparison retains the dependence between the response energy and its fourth moment.

A two-branch projection certificate. Work in one coordinate pair. Let $0 < \mu < 1/2$, set $m_0 = 1/2$, and parameterize

$$\gamma = \frac{1+x}{x(x+\mu)}, \qquad 0 < x \leq \tfrac{1}{2}.$$

Write

$$V_0 = 1+x, \quad X = x(x+\mu), \quad F_0 = \mu + 2x + x^2, \quad b_0 = 2\gamma x, \quad w_f = \frac{1}{\mu\gamma + b_0 + b_0^2/(4\gamma)} = \frac{X}{V_0 F_0}.$$

Choose a positive row multiplier a_0. For each of the two pairs

$$(q_*, d_*) = (1,1), \qquad (q_*, d_*) = \left(\frac{3+\chi}{4}, \chi \right), \tag{471}$$

define

$$\begin{aligned} J_* &= F_0 - 2 d_* x X, & Y_* &= q_* V_0 + (a_0^2/4 - d_* a_0) X, \\ w_{e,*} &= \frac{J_*}{F_0 Y_*}, & F_{e,*} &= 1/w_{e,*}, \\ K_* &= \frac{a_0^2(1-m_0)/4 + m_0 \gamma F_{e,*}}{m_0\gamma + a_0 + a_0^2/(4\gamma) - F_{e,*}}. \end{aligned}$$

Set $w_e = \min_* w_{e,*}$, $K = \max_* K_*$, $h_0 = a_0 w_e$, and $j_0 = b_0 w_f$.

Lemma 11.5. *Assume*

$$\frac{13}{5} \leq a_0 \leq \frac{16}{5}, \quad \frac{5}{16} \leq \chi \leq \frac{69}{100}, \quad 0 < \mu < \frac{1}{2}, \quad 0 < x \leq \frac{1}{2}.$$

All denominators just displayed are positive and $h_0 > j_0$. Applied coordinatewise to the stationary response geometry, these weights give a covariance $\mathcal{A}$, positive definite on the response graph, such that

$$\mathrm{tr}[D_0 \mathcal{A}] - \sum_i \gamma_i \left(m_0 e_i^\top T \mathcal{A} T^\top e_i + \mu_i f_i^\top T \mathcal{A} T^\top f_i \right) \leq \sum_i K_i e_i^\top (I-T)\mathcal{A}(I-T)^\top e_i. \tag{472}$$

The total movement on the right before multiplication by K_i is strictly positive.

Proof. First verify the scalar denominators. We have $\gamma \geq 3$, $F_0 > 0$, and

$$J_* \geq F_0/2 > 0,$$

since

$$F_0 - 4x^2(x+\mu) = \mu(1-4x^2) + x(2+x-4x^2) \geq 0.$$

Furthermore

$$Y_*/V_0 = q_* + \frac{a_0^2/4 - d_* a_0}{\gamma} \geq q_* - d_*^2/\gamma > 0.$$

For $(q_*, d_*) = (1,1)$, $Y_* \leq V_0 \leq 3/2$, hence $F_{e,*} \leq 3$. For the rank branch, monotonicity in a_0 gives

$$Y_* \leq \frac{3(3+\chi)}{8} + \frac{1}{2}\left(\frac{64}{25} - \frac{16\chi}{5}\right) = \frac{481}{200} - \frac{49\chi}{40} \leq \frac{6471}{3200} < \frac{81}{40}.$$

Thus $F_{e,*} < 81/20$ in either branch. The denominator of K_* is greater than

$$\frac{3}{2} + \frac{13}{5} - \frac{81}{20} = \frac{1}{20}.$$

Its numerator is positive. Also

$$a_0 J_* - 2xY_* = a_0 F_0 - 2q_* x V_0 - \frac{a_0^2}{2} xX > 0.$$

The coefficient of μ in the last expression is at least $3a_0/5$; its remaining coefficient after removing x is at least $48/25$. Therefore $a_0 w_{e,*} > b_0 w_f$ for both branches, and the same holds for their minimum.

Let $\mathcal{D}$ be the diagonal matrix with pair weights $w_{e,i}, w_{f,i}$. If the columns of R form an orthonormal basis of the response graph, put

$$\mathcal{A} = R(R^\top \mathcal{D}^{-1} R)^{-1} R^\top, \qquad \mathcal{B} = \mathcal{D} - \mathcal{A}.$$

Then $\mathcal{A}, \mathcal{B} \succeq 0$, $\mathcal{A}\mathcal{D}^{-1}\mathcal{B} = 0$, and $\operatorname{tr}[\mathcal{D}^{-1}\mathcal{A}] = k$. For an e_i row $t_i = e_i^\top T$, complete the square using

$$C_i = \gamma_i \mathcal{B} + (m_0\gamma_i + K_i)\mathcal{A}, \qquad C_i^{-1} = \gamma_i^{-1}\mathcal{D}^{-1} + \big((m_0\gamma_i + K_i)^{-1} - \gamma_i^{-1}\big)\mathcal{D}^{-1}\mathcal{A}\mathcal{D}^{-1}.$$

The row inequality is

$$\begin{aligned}\gamma_i t_i \mathcal{B} t_i^\top + m_0\gamma_i t_i \mathcal{A} t_i^\top + K_i(e_i^\top - t_i)\mathcal{A}(e_i - t_i^\top) + h_{0,i} t_i e_i \\ \geq \frac{e_i^\top \mathcal{A} e_i}{w_{e,i}} - \frac{h_{0,i}^2}{4\gamma_i w_{e,i}}.\end{aligned}$$

This follows from the exact identity

$$K - \frac{(K - a_0/2)^2}{m_0\gamma + K} + \frac{a_0^2}{4\gamma} = 1/w_e.$$

Writing $F = 1/w_e$ and $d = m_0\gamma + a_0 + a_0^2/(4\gamma)$, differentiation gives

$$\frac{d}{dF}\frac{a_0^2(1-m_0)/4 + m_0\gamma F}{d - F} = \frac{(m_0\gamma + a_0/2)^2}{(d-F)^2} > 0.$$

Thus the minimum defining w_e corresponds to the maximum defining K. For an f_i row, the graph identity $f_i^\top T\mathcal{A} = f_i^\top \mathcal{A}$ and a square in the PSD matrix $\mathcal{B}$ give

$$\gamma_i f_i^\top T\mathcal{B}T^\top f_i + \mu_i\gamma_i f_i^\top \mathcal{A} f_i + j_{0,i} f_i^\top T f_i \geq \frac{f_i^\top \mathcal{A} f_i}{w_{f,i}} - \frac{j_{0,i}^2}{4\gamma_i w_{f,i}}.$$

Here the coefficient identity is $\mu\gamma + b_0 + b_0^2/(4\gamma) = 1/w_f$.

Write $S_* = (a_0^2 w_e + b_0^2 w_f)/(4\gamma)$ and $c_0 = (a_0 w_e - b_0 w_f)/\gamma$. For either pair in (471), the definition of $w_{e,*}$ gives

$$q_* w_{e,*} + w_f + \frac{a_0^2 w_{e,*} + b_0^2 w_f}{4\gamma} - d_* \frac{a_0 w_{e,*} - b_0 w_f}{\gamma} = 1.$$

Its coefficient of $w_{e,*}$ is positive. Replacing $w_{e,*}$ by their minimum therefore proves both inequalities

$$w_e + w_f + S_* - c_0 \leq 1, \qquad \frac{3+\chi}{4} w_e + w_f + S_* - \chi c_0 \leq 1.$$

Combine these with (470), sum the row completions, and use $\mathrm{tr}[\mathcal{D}^{-1}\mathcal{A}] = k$. The result is

$$\mathrm{tr}[\Gamma T\mathcal{B}T^\top] + \sum_i \gamma_i (m_0 e_i^\top T\mathcal{A}T^\top e_i + \mu_i f_i^\top T\mathcal{A}T^\top f_i) + \sum_i K_i e_i^\top (I-T)\mathcal{A}(I-T)^\top e_i \geq \mathrm{tr}[D_0\mathcal{D}].$$

Since $\mathrm{tr}[D_0\mathcal{B}] \geq \mathrm{tr}[\Gamma T\mathcal{B}T^\top]$ by (466), this proves (472). If movement vanished, the graph identity and positivity of $\mathcal{A}$ on its graph would make T the identity on a k-dimensional subspace. Then $\|T\|_F^2 \geq k$. But $\mathrm{tr}[D_0] \leq 2k$ and $\min_i \gamma_i > 2$ in (466) contradict this. □

Curvature and the scalar reduction. The functional-Hessian inequality in Eq. (341) is

$$D^2\phi_i(X)[H,H] \leq 2\left(\mathrm{tr}[B_i(X)HX^{-1}H] - \frac{(D\phi_i(X)[H])^2}{\phi_i(X)}\right).$$

At outer stationarity, the mixed functional–budget term uses only the f_i response direction. In the notation of (467) it is

$$2\lambda_i Y_i f_i^\top T\xi, \qquad \lambda_i = \beta_i' \sqrt{\frac{g\gamma_i}{\beta_i p_i q_i}}.$$

For any $0 < \mu_i < 1/2$, set $\eta_i = 2 - 4\mu_i$ and

$$L_i = -\beta_i'' - \frac{(\beta_i')^2}{\eta_i \beta_i}.$$

Young's inequality, charged only to the f_i direction, yields

$$D^2 P_\rho[h,h] \leq 4\left[Q - \sum_i \gamma_i \left(\tfrac{1}{2}(e_i^\top T\xi)^2 + \mu_i (f_i^\top T\xi)^2\right)\right] - g\sum_i L_i S_i h_i^2. \tag{473}$$

This is Eq. (434). No derivative of μ_i is taken. Lemma 11.5 now gives strict negative expected curvature whenever

$$gL_i > 4p_i q_i K_i \qquad \text{for every active coordinate.} \tag{474}$$

For the scalar comparison let $t = |y|$, $z = 1 - t^2$, and $\beta(y) = \sqrt{z}\, f(t)$. Define

$$N = zf' - tf, \qquad A = f - zf', \qquad \eta = 2 - 4\mu, \qquad H = -\eta f(zN' + tN) - N^2. \tag{475}$$

By interchanging U and V, we may assume $y = t \geq 0$. Then $\beta' = N/\sqrt{z}$ and $L = H/(\eta z^{3/2} f)$. At fixed product of the two normalized light variables, γ is fixed and pq is maximized when $g\beta u = 1+t$. To check this without any monotonicity assumption on K, write

$$\alpha_{\mathrm{loc}} = \frac{g\beta u}{1+t}, \quad \nu = \frac{g\beta v}{1-t}, \quad d = \alpha_{\mathrm{loc}}\nu, \quad pq = (1 + B\alpha_{\mathrm{loc}})(1 + Cd/\alpha_{\mathrm{loc}}),$$

where $B \geq 0$ and $C \leq 0$. Its derivative in α_{loc} is $B - Cd/\alpha_{\mathrm{loc}}^2 \geq 0$. At that maximizing pair,

$$q = \frac{A}{(1-t)f}, \qquad p = 1 + \frac{gN}{(1+t)\sqrt{z}\,\gamma}.$$

The actual x satisfies

$$G(t,x) = g^2 f^2 x^2 (x+\mu)^2 - z(1+x)^2 \leq 0. \tag{476}$$

We certify $G(t, 1/2) > 0$. Monotonicity of $x(x+\mu)/(1+x)$ then places all actual states in $0 < x < 1/2$.

For each branch (471), retain J_*, Y_* from the projection lemma above with $a_0 = a(t)$ and put

$$\begin{aligned} B_* &= \tfrac{1}{2} V_0 F_0 Y_* + \tfrac{1}{8} a_0^2 X J_*, \\ D_* &= \tfrac{1}{2} V_0^2 J_* + a_0 V_0 X J_* + \tfrac{1}{4} a_0^2 X^2 J_* - V_0 X F_0 Y_*, \\ W_* &= g(H D_* - 4\eta A N X B_*), \end{aligned} \tag{477}$$

$$P_*(t,x) = W_*^2 - 16\eta^2 A^2 (1+t)^2 z V_0^2 B_*^2. \tag{478}$$

Direct substitution gives $K_* = V_0 B_*/D_*$. All denominators are positive by Lemma 11.5. The budget checks below imply $W_* > 0$. Hence

$$G \leq 0 \implies P_{\mathrm{one}} > 0 \text{ and } P_{\mathrm{rank}} > 0 \tag{479}$$

implies (474) for both branches, and therefore for $K = \max K_*$.

Exact parameters and full-domain certificates. Let T_j denote the Chebyshev polynomial, $T_0(s) = 1$, $T_1(s) = s$, $T_{j+1}(s) = 2sT_j(s) - T_{j-1}(s)$. Use

$$f(t) = 1 + \sum_{j=1}^{12} c_j \big(T_j(2t^2 - 1) - (-1)^j\big). \tag{480}$$

All the following terminating decimals denote exact rational numbers.

j	c_j	j	c_j
1	−0.0259310343	7	−0.0019713066
2	−0.0422913686	8	−0.0010112036
3	−0.0216701975	9	−0.0004621211
4	−0.0124424432	10	−0.0001949353
5	−0.0067644498	11	−0.0000624130
6	−0.0037723907	12	−0.0000154478

Take

$$\mu(t) = 0.001 + 0.498(1-t)^2 + t(1-t)^2\big(-1.29237356 + 6.08566608t - 10t^2 + 4.53879333t^3\big), \tag{481}$$

$$a(t) = 2.66889770 + 1.60303187t - 1.35345314t^2. \tag{482}$$

The two choices of coupling and rank parameter are

$$(g, \chi) = \left(\frac{5403}{1000}, \frac{69}{100}\right) \quad (r \le 3), \qquad (g, \chi) = \left(\frac{113}{20}, \frac{5}{16}\right) \quad (r \le 32). \tag{483}$$

Indeed $3(69/100)^3 = 985527/10^6 < 1$ and $32(5/16)^3 = 125/128 < 1$.

Proposition 11.6. *For both choices in* (483)*, the budget is admissible and* (479) *holds on* $[0,1] \times [0,1/2]$.

Proof. The exact univariate coefficient tests give

$$f > 0, \quad -N/t > 0, \quad (f^2 - z)/t^2 > 0, \quad f + zf' > 0, \quad A > 0, \quad \mu > 10^{-4}, \quad \eta > 0, \quad H > 0,$$

$$13/5 < a < 16/5, \qquad G(t, 1/2) > 0.$$

Polynomial quotients are continued to zero. The first eight polynomials except $\mu - 10^{-4}$ have positive ordinary-degree Bernstein coefficients on $[0,1]$. The latter passes on four dyadic subintervals with maximum depth three. The upper bound on a passes after a single midpoint split; the lower bound passes without subdivision. Both $G(t, 1/2)$ polynomials pass without subdivision. Thus β is even and smooth inside the interval, satisfies $1 - y^2 \le \beta(y) \le 1$, and has positive response scales throughout the light region. The square-root endpoint factor is the only endpoint singularity.

For completeness, if a polynomial on the unit square has monomial coefficients M_{ij} and bidegree at most (n, m), its tensor Bernstein coefficients are

$$\mathcal{B}_{ab} = \sum_{i \le a,\, j \le b} M_{ij} \frac{\binom{a}{i}}{\binom{n}{i}} \frac{\binom{b}{j}}{\binom{m}{j}}. \tag{484}$$

On a dyadic rectangle, first make its affine substitution. The basis functions are nonnegative and sum to one. Thus strict positivity of all these coefficients proves strict positivity on the entire rectangle.

Use coordinates (t, u) with $x = u/2$. The polynomials $P_{\mathrm{one}}(t, u/2)$ and $P_{\mathrm{rank}}(t, u/2)$ each have bidegree $(156, 14)$; $G(t, u/2)$ has bidegree $(60, 4)$. Construct a partition recursively, starting from $[0,1]^2$ at depth zero. On each rectangle, compute the tensor Bernstein coefficients after its affine substitution, using Eq. (484). Retain the rectangle as a P leaf if every coefficient of both P polynomials is greater than 10^{-7}. Otherwise, retain it as a G leaf if every coefficient of G is greater than 10^{-5}. If neither test passes, split at the midpoint of the t interval at even depth and of the u interval at odd depth. Apply the same rule to both children at depth one greater, visiting the lower half first.

Exact rational arithmetic verifies that this construction terminates for both parameter choices, with the following counts.

Case	P leaves	G leaves	Total	Maximum (t, u) depths
$r \le 3$, $g = 5.403$	88	48	136	$(11, 11)$
$r \le 32$, $g = 5.65$	15	12	27	$(7, 6)$

Every split partitions its parent into two rectangles with disjoint interiors. Thus the retained rectangles cover the unit square; their areas sum to one in each case. A G leaf has no point with

$G \le 0$. Every point in the physical domain therefore lies in a P leaf, where both P polynomials exceed 10^{-7}, including on the boundary. This proves Eq. (479).

The polynomials are reconstructed from the displayed rational parameters. After clearing denominators, Eq. (484) and midpoint de Casteljau subdivision use only integer arithmetic. Direct affine substitution into the global monomial polynomials independently verifies every retained leaf, checking 503,430 bivariate coefficient inequalities across the two cases. □

From the certificate to signs.

Corollary 11.7. *Under the hypotheses of Theorem 11.23, the following bounds are available:*

$$r \le 3: \qquad \|\sum_i \sigma_i A_i\| \le 2\sqrt{\frac{5403}{1000}\, r^{1/3}\alpha},$$

$$r \le 32: \qquad \|\sum_i \sigma_i A_i\| \le 2\sqrt{\frac{113}{20}\, r^{1/3}\alpha}.$$

The signings may depend on the selected bound.

Proof. At every stationary light state, Proposition 11.6 gives $gL_i > 4p_iq_iK_i$. Lemma 11.5 and (473) supply a covariance with

$$\mathbb{E}[D^2 P_\rho[h,h]] \le -\sum_i (gL_i - 4p_iq_iK_i) S_i \mathbb{E}[h_i^2] < 0.$$

Choose a global minimizer of P_ρ with the fewest active coordinates. An available endpoint move contradicts this choice; otherwise the active Hessian is PSD at a stationary light point, contradicting the displayed inequality. Thus the minimizer is a vertex. The scalar trial $U = V = \ell I$ gives

$$P_\rho(0) \le 2\sqrt{g r^{1/3}\alpha + 2D\rho},$$

because $\phi_i(I) \le r^{1/3}\alpha$ and $\sum_i A_i = I$. Let $\rho \downarrow 0$ and select a signing occurring along a subsequence. There are only finitely many signings, so this proves the claim. No input matrix has been split into separately signed summands. □

All rank ranges and the exact decimal comparison. For $r = 1, 2$, use the exact trace-variance bound of Theorem 3.7

$$\|\sum_i \sigma_i A_i\| \le 2\sqrt{561/200}\, \|\sum_i \operatorname{tr}[A_i] A_i\|^{1/2}, \qquad \sum_i \operatorname{tr}[A_i] A_i \preceq r\alpha I.$$

Its squared coefficient relative to $\alpha \log(2r)$ is $(561/50)/\log 2$ for either integer rank. For $r = 3$, apply the first bound in Corollary 11.7. For $4 \le r \le 32$, apply its second bound. The function $r^{1/3}/\log(2r)$ decreases and then increases, so its maximum on $[4, 32]$ is at an endpoint. Those endpoint values are equal.

For $r \ge 32$, use Eq. (386): every integer $s \ge 5$ permits a signing with

$$\|\sum_i \sigma_i A_i\| \le \sqrt{6(s+1) r^{1/s}\alpha}.$$

Choose $s = \lfloor \log_2 r \rfloor$. On $2^s \le r \le 2^{s+1}$ the squared normalized coefficient has its maximum at an endpoint. The lower endpoint gives $12/\log 2$, while the upper gives

$$\frac{12}{\log 2}\frac{s+1}{s+2}\,2^{1/s} \le \frac{12}{\log 2}.$$

Indeed, for $s \ge 5$,

$$\left(1+\frac{1}{s+1}\right)^s \ge 1 + \frac{s}{s+1} + \frac{s(s-1)}{2(s+1)^2} = 2 + \frac{s^2-3s-2}{2(s+1)^2} > 2.$$

This proves the upper-endpoint comparison. The original individual higher-order bounds remain available; the dyadic choice is only a convenient uniform comparison.

The resulting exact squared bounds and convenient rational upper bounds are

Rank range	Squared coefficient	Certified rational upper bound
$1 \le r \le 2$	$\dfrac{561}{50\log 2}$	$\dfrac{112200}{6931}$
$r = 3$	$\dfrac{5403\,3^{1/3}}{250\log 6}$	$\dfrac{31169907}{1791700}$
$4 \le r \le 32$	$\dfrac{113\,4^{1/3}}{5\log 8}$	$\dfrac{17937733}{1039700}$
$r \ge 32$	$\dfrac{12}{\log 2}$	$\dfrac{120000}{6931}$

These use

$$\log 2 > 0.6931, \quad \log 6 > 1.7917, \quad \log 8 > 2.0794, \quad 3^{1/3} < 1.44225, \quad 4^{1/3} < 1.58741.$$

The root bounds are checked by integer cubes. For the logarithms, use 200 terms of

$$\log v = 2\sum_{j=0}^{N-1} \frac{z^{2j+1}}{2j+1} + R_N, \quad z = \frac{v-1}{v+1}, \quad 0 \le R_N \le \frac{2z^{2N+1}}{(2N+1)(1-z^2)}.$$

Every rational bound in the table is strictly smaller than $(4171/1000)^2$. The tightest comparison is

$$\left(\frac{4171}{1000}\right)^2 - \frac{31169907}{1791700} = \frac{7296997}{17917000000} > 0.$$

For nonintegral r, first apply the result to $\lfloor r \rfloor$ and then increase the logarithmic bound. Finally every signing obeys $-I \preceq \sum_{i=1}^N \sigma_i A_i \preceq I$, giving the cap by one. These comparisons give coefficient 4.171; Section 11.7 proves the stronger coefficient in Theorem 11.23.

11.6 Retaining the mixed budget and functional derivative

We now keep the mixed functional–budget derivative in the covariance comparison. The square-root budget suffices at every Schatten order used below. This improves the existence bound and also applies to rounding with two fixed matrix targets.

For an integer $q \ge 3$, use the functional $\phi_i(X) := \|A_i^{1/2} X A_i^{1/2}\|_{S_q}$, the budget $\beta(y) := \sqrt{1-y^2}$, and the coupling $g > 2$. Write $F_y(X) := g\sum_{i=1}^N \beta(y_i)\phi_i(X)A_i$. We use the potential, dual variables, and response normalization from Section 8.5, with g in place of ι. All sums over active coordinates below range over $i \in I$, where $I \subseteq [N]$ is the active set and $k := |I|$.

The full second variation. Let D_0 be the normalized response-energy matrix before diagonal padding, and let T be the normalized response operator. Put

$$Q := \xi^\top D_0 \xi, \qquad Y_i := e_i^\top (I - T)\xi, \qquad F_i := f_i^\top \xi = f_i^\top T\xi.$$

Eq. (361) gives $D_0 \succeq T^\top \Gamma T$, where Γ equals $\gamma_i I_2$ on pair i and $\gamma_i := (g\beta_i u_i v_i)^{-1}$. Define

$$c := q+1, \quad m_0 := \frac{q-1}{q+1}, \quad K_i := \frac{g(-\beta_i'')}{c p_i q_i}, \quad \ell_i := \frac{\beta_i'}{c}\sqrt{\frac{g\gamma_i}{\beta_i p_i q_i}}. \tag{485}$$

Here q_i is the response coefficient, distinct from the Schatten order q. At an outer stationary light point,

$$\begin{aligned} c^{-1} D^2 P_\rho[h,h] \le Q - \sum_{i\in I} m_0 \gamma_i [(e_i^\top T\xi)^2 + (f_i^\top T\xi)^2] \\ + 2\sum_{i\in I} \ell_i Y_i F_i - \sum_{i\in I} K_i Y_i^2. \end{aligned} \tag{486}$$

Indeed, the inverse terms contribute $2Q$. The functional Hessian bound in Eq. (341) contributes at most $(q-1)Q$ minus $(q-1)\sum_{i\in I}\gamma_i \|T\xi|_i\|^2$. Stationarity gives $c_i w_i = d_i z_i$, and the exact mixed term is $2\beta_i'\sqrt{g\gamma_i/(\beta_i p_i q_i)}Y_i F_i$. Finally, $Y_i^2 = p_i q_i (v_i w_i + u_i z_i) h_i^2$, so the budget term is $-cK_i Y_i^2$. Division by c proves the inequality.

For a normalized compressed matrix $R \succeq 0$ with $\mathrm{tr}[R^q] = 1$ and a density matrix $S \succeq 0$ with $\mathrm{tr}[S] = 1$, put $\kappa := \mathrm{tr}[R^{q+1}]$ and $e := \mathrm{tr}[RS]$. Then

$$0 \le \kappa \le 1, \qquad e \le \frac{q+\kappa}{q+1}. \tag{487}$$

The first assertion follows because the eigenvalues of R lie in $[0,1]$. The scalar inequality $x^{q+1} - (q+1)x + q \ge 0$, applied to the largest eigenvalue, gives the second assertion. Use (e_U, κ_U) and (e_V, κ_V) for the two local pairs of compressed matrices and density matrices, as in Eq. (469). For one coordinate put $\alpha_{\mathrm{loc}} := c_i u_i/(c_i u_i + d_i v_i)$ and $\delta := \gamma_i^{-1}$. The unpadded local identities are

$$\begin{aligned} (D_0)_{ee} &= \alpha_{\mathrm{loc}} e_U + (1-\alpha_{\mathrm{loc}}) e_V, & (D_0)_{ff} &= (1-\alpha_{\mathrm{loc}}) e_U + \alpha_{\mathrm{loc}} e_V, \\ T_{ee} &= -\delta[\alpha_{\mathrm{loc}}\kappa_U + (1-\alpha_{\mathrm{loc}})\kappa_V], & T_{ff} &= -T_{ee}, \\ T_{fe} &= \delta \frac{\alpha_{\mathrm{loc}}^2 \kappa_U - (1-\alpha_{\mathrm{loc}})^2 \kappa_V}{\sqrt{\alpha_{\mathrm{loc}}(1-\alpha_{\mathrm{loc}})}}. \end{aligned} \tag{488}$$

To verify the last identity, set $x := c_i u_i$, $v := d_i v_i$, and $\omega := c_i w_i = d_i z_i$. The response scales are $x\omega$ and $v\omega$, the local unnormalized response entries are $gx^2\kappa_U$ and $gv^2\kappa_V$, and $gxv = \delta$. Rotation by e_i, f_i gives the stated formula. This calculation does not assume that T is symmetric.

A covariance for the mixed term. Suppress the coordinate index and choose $\zeta \ge 0$. Define

$$\begin{aligned} d_* &:= m_0\gamma + K, & \nu &:= \frac{\ell}{1+\zeta\delta/2}, \\ Z_e &:= K + \delta(K^2+\nu^2), & Z_f &:= m_0\gamma + \zeta + \frac{\delta\zeta^2}{4} - \frac{\ell^2}{d_*}, \\ w_e &:= Z_e^{-1}, & w_f &:= Z_f^{-1}. \end{aligned} \tag{489}$$

These weights are positive. In fact, with $t := |y|$, $J := \ell^2/(K\gamma) = t^2/c$ and $v := K\delta$, we have

$$Z_f = \gamma[m_0 + \zeta\delta + \zeta^2\delta^2/4 - vJ/(m_0+v)] \ge \gamma(m_0 - J) \ge \gamma(q-2)/(q+1) > 0.$$

Set

$$\mathcal{E}_i := w_e(D_0)_{ee} + w_f(D_0)_{ff} + \delta[(K^2+\nu^2)w_e + \zeta^2 w_f/4] + 2Kw_eT_{ee} + 2\nu w_eT_{fe} + \zeta w_fT_{ff}. \tag{490}$$

Lemma 11.8. *If $\mathcal{E}_i < 1$ for every active coordinate at a stationary light point, a covariance of physical directions has strictly negative expected second variation.*

Proof. Let D have diagonal weights $w_{e,i}, w_{f,i}$ in the paired basis. For an orthonormal basis matrix R of the response graph, define

$$\Sigma := R(R^\top D^{-1}R)^{-1}R^\top, \qquad B := D - \Sigma.$$

Then $\Sigma, B \succeq 0$, $\Sigma D^{-1}B = 0$, and $\operatorname{tr}[D^{-1}\Sigma] = k$. The covariance Σ is positive definite on the response graph. Since the graph is parameterized by the physical directions, it gives a covariance of such directions.

For the row $a := e^\top T$, complete a square with $C := \gamma B + d_*\Sigma \succ 0$. Its inverse is

$$C^{-1} = \delta D^{-1} + (d_*^{-1} - \delta)D^{-1}\Sigma D^{-1}.$$

The resulting row inequality is

$$\begin{aligned}\gamma aBa^\top + m_0\gamma a\Sigma a^\top + K(e^\top - a)\Sigma(e - a^\top) - 2\ell(e^\top - a)\Sigma f + 2Kw_eae \\ \geq (K + \delta K^2)e^\top\Sigma e - 2\ell e^\top\Sigma f - \frac{\ell^2}{d_*}f^\top\Sigma f - \delta K^2 w_e.\end{aligned} \tag{491}$$

Indeed, its quadratic part is $aCa^\top$, and its linear vector is $KBe + \ell\Sigma f$. The two summands are orthogonal in the C^{-1} inner product because $BD^{-1}\Sigma = 0$. For the row $b := f^\top T$, the graph condition gives $b\Sigma = f^\top\Sigma$. Completing a square in B, with helper $\nu e + \zeta f/2$, gives

$$\begin{aligned}\gamma bBb^\top + m_0\gamma f^\top\Sigma f + 2\nu w_e be + \zeta w_f bf \\ \geq \delta\nu^2 e^\top\Sigma e + 2\nu(1 + \zeta\delta/2)e^\top\Sigma f + (m_0\gamma + \zeta + \delta\zeta^2/4)f^\top\Sigma f \\ - \delta(\nu^2 w_e + \zeta^2 w_f/4).\end{aligned} \tag{492}$$

The mixed coefficients cancel because $\nu(1 + \zeta\delta/2) = \ell$. The remaining coefficients are Z_e, Z_f, the local entries of D^{-1}. Sum the two row inequalities, use $D_0 \succeq T^\top\Gamma T$, and subtract the local costs. For a centered response with covariance Σ, this gives

$$\sum_{i\in I} K_i\mathbb{E}[Y_i^2] - 2\sum_{i\in I}\ell_i\mathbb{E}[Y_iF_i] - \mathbb{E}[Q] + \sum_{i\in I} m_0\gamma_i\mathbb{E}[\|T\xi|_i\|^2] \geq \sum_{i\in I}(1 - \mathcal{E}_i) > 0.$$

Eq. (486) completes the proof. All weights are chosen at the point and are not differentiated. □

Exact scalar comparisons. It suffices to take $t := y \geq 0$: exchanging U, V and reversing y changes the signs of both ℓ and T_{fe}. Define $h := g\beta u/(1+t)$ and $v_* := g\beta v/(1-t)$. Lightness gives $0 < h, v_* < 1$. Set

$$a := \frac{h}{1 - t + th}, \quad b := \frac{v_*}{1 + t - tv_*}, \quad y_0 := ab, \quad S := a + b, \quad z := 1 - t^2, \quad D_L := (1 - ta)(1 + tb).$$

The new variables range independently over $0 < a, b < 1$. Direct substitution gives

$$\alpha_{\mathrm{loc}} = \frac{a}{S}, \quad p_iq_i = \frac{1}{D_L}, \quad K = \frac{gD_L}{cz^{3/2}}, \quad \delta = \frac{z^{3/2}y_0}{gD_L}, \quad 2\ell T_{fe} = -\frac{2t(a^2\kappa_U - b^2\kappa_V)}{cS}. \tag{493}$$

The coefficients of e_U, e_V in the cost are positive. After Eq. (487), the cost is affine in (κ_U, κ_V). It therefore suffices to check the four pairs $(i,j) \in \{0,1\}^2$.

At order three take $g := 213/50$, $\zeta := 2$, and $\omega := 10^{-6}$. Put $\rho_w := w_f/w_e$. Then $1 - \mathcal{E} = w_e(K - \mathcal{R})$, where

$$\mathcal{R} := (D_0)_{ee} + \rho_w (D_0)_{ff} + 2KT_{ee} + \frac{2\ell T_{fe}}{1+\delta} + 2\rho_w T_{ff} + \delta\rho_w.$$

Use $t = (1-w^2)/(1+w^2)$, with $0 < w \le 1$, and define

$$\begin{aligned} Z &:= 1 + w^2, & T_0 &:= 1 - w^2, \qquad h_0 := 8w^3, \\ L &:= gZ(Z - T_0 a)(Z + T_0 b), & B_0 &:= L + h_0 y_0, \\ E_0 &:= (c + y_0) Z^2 B_0^2 + T_0^2 L^2, \\ F_0 &:= (q - 1 + y_0) Z^2 [(q-1)L^2 + 2c h_0 y_0 L + c h_0^2 y_0^2] - y_0 T_0^2 L^2. \end{aligned} \tag{494}$$

Here $q = 3$ and $c = 4$. At actual states all these quantities except T_0 are positive. Positivity of F_0 also follows from $Z_f > 0$. For a vertex (i,j) define

$$M_{ij} := ai + bj, \quad M^{(2)}_{ij} := a^2 i - b^2 j, \quad Q_{0,ij} := a(q+i) + b(q+j) - 2y_0 M_{ij}, \quad Q_{1,ij} := b(q+i) + a(q+j).$$

The four polynomial comparisons are

$$\begin{aligned} \Phi_{ij} := {} & cF_0 [Z B_0^2 ((1-\omega) LS - h_0 Q_{0,ij}) + 2T_0 h_0 L B_0 M^{(2)}_{ij}] \\ & - Z h_0 y_0 (q - 1 + y_0) E_0 L [L Q_{1,ij} + c h_0 y_0 (2M_{ij} + S)] \ge 0. \end{aligned} \tag{495}$$

Indeed, substitution in the cost gives the exact identity

$$\Phi_{ij} = c^2 F_0 Z B_0^2 h_0 S [(1-\omega) K - \mathcal{R}_{ij}].$$

Thus Eq. (495) implies $\mathcal{E}_i \le 1 - 10^{-6} K_i w_{e,i} < 1$.

For every integer $q \ge 5$, take $g := 4(q+1)/3$ and $\zeta := 0$. Set $d := q - 1$ and $g_0 := 1333(q+1)/1000 = (3999/4000)g$. Define

$$\begin{aligned} E_q &:= d^2 + y_0 (d - t^2), & F_q &:= (c + t^2 + y_0)(d + y_0), \\ \widetilde{Q}_{0,ij} &:= a(q+i) + b(q+j) - 2y_0 (ai + bj) - 2t(a^2 i - b^2 j), \\ \widetilde{Q}_{1,ij} &:= b(q+i) + a(q+j), \\ N_{ij} &:= c E_q \widetilde{Q}_{0,ij} + y_0 F_q \widetilde{Q}_{1,ij}, & D_q &:= c E_q D_L S. \end{aligned} \tag{496}$$

Now $\rho_w = y_0 F_q / (c E_q)$ is independent of g, and $\mathcal{R}_{ij} = N_{ij}/(c^2 E_q S)$. It suffices to prove

$$\Pi_{ij} := (g_0 D_q)^2 - (1 - t^2)^3 N_{ij}^2 \ge 0. \tag{497}$$

At actual states $D_q > 0$, so this comparison implies $g_0 D_q \ge (1-t^2)^{3/2} |N_{ij}|$. It therefore proves $\mathcal{E}_i \le 1 - K_i w_{e,i}/4000 < 1$ at the actual coupling g.

For completeness, both full-domain certificates are specified here. For coordinate degrees n_1, n_2, n_3, set $\mathcal{A} := \prod_{j=1}^{3} \{0, \dots, n_j\}$. A polynomial $P(x) := \sum_{\nu \in \mathcal{A}} M_\nu x^\nu$ has tensor Bernstein coefficients on the unit cube given by

$$B_\mu := \sum_{\nu \in \mathcal{A} : \nu \le \mu} M_\nu \prod_{j=1}^{3} \frac{\binom{\mu_j}{\nu_j}}{\binom{n_j}{\nu_j}}.$$

Nonnegative coefficients imply a nonnegative polynomial on the whole cube. On a rectangle first apply its affine substitution. The complete preorder subdivision words are

$$\texttt{TPUPVTPPTUPVTPUPVPPPP} \quad \text{and} \quad \texttt{TPP}.$$

Here `T`, `U`, and `V` bisect the first, second, and third coordinates, respectively; each is followed by its left and right subtrees, and `P` denotes a leaf. The first word covers $(w, a, b) \in [0,1]^3$ with 11 rectangles. On every rectangle all four polynomials in Eq. (495), of tridegree $(36, 8, 8)$, have nonnegative Bernstein coefficients. For the second word substitute $q = 5 + v$ in Eq. (497). The degree in v is eight. All 36 coefficient polynomials have nonnegative Bernstein coefficients on both resulting rectangles in (t, a, b). Since $v \geq 0$, this proves every higher order at once.

The two tests contain 131,868 and 44,578 coefficient inequalities, respectively. The trees are consumed completely and each has total volume one. The 4,376 zero coefficients are allowed: strictness at actual light states follows from the explicit row margins above. Apply the coefficient construction in Appendix B.1 to the displayed polynomials and the two complete trees. This gives the stated sign tests using exact rational arithmetic.

Fixed targets and subisotropic inputs.

Lemma 11.9. *Let $A_1, \dots, A_N \succeq 0$ be a finite family. Use either $(q, g) = (3, 213/50)$ or an integer $q \geq 5$ with $g = 4(q+1)/3$, and use $\beta(y) = \sqrt{1 - y^2}$. For any $y^0 \in [-1,1]^N$ and $U_0, V_0 \succ 0$, there is one signing satisfying*

$$M(\sigma) \preceq M(y^0) + U_0^{-1} + F_{y^0}(V_0), \tag{498}$$

$$-M(\sigma) \preceq -M(y^0) + V_0^{-1} + F_{y^0}(U_0). \tag{499}$$

Proof. Freeze the two right sides as H_+, H_-. In the potential replace the upper constraint targets by $bI + H_+$ and $bI + H_-$. The trial $b = 0$, $U = U_0$, $V = V_0$ is feasible at y^0. Adding the constraints bounds b below by $-\|H_+ + H_-\|/2$. Objective sublevels then bound b, U, V above, and the constraints bound their inverses above. These bounds are uniform over the cube. Thus attainment and continuity follow as in Lemma 8.3. The fixed targets have zero derivative. The same strict dual equations, invertible constraint derivative, and positive primal Hessian therefore give uniqueness and smoothness on each open face. The endpoint tests, response identities, and Eq. (486) are unchanged.

Choose a global minimizer with the fewest active coordinates. An available endpoint move contradicts this choice. Otherwise the point is light; if it is not a vertex, its active Hessian is positive semidefinite. The scalar certificates and Lemma 11.8 contradict this assertion. Hence a minimizing vertex exists. Its value is at most $\rho \operatorname{tr}[U_0 + V_0]$. At a vertex the coupling vanishes, so the two target inequalities hold with this additional scalar slack. Let $\rho \downarrow 0$ and choose a recurrent signing from the finite set of signings. This removes the slack and proves both bounds. □

Corollary 11.10. *Suppose $S := \sum_{i=1}^N A_i \preceq I_D$, $A_i \succeq 0$, $\|A_i\| \leq \alpha$, and $\operatorname{rank} A_i \leq r$, where $r \geq 1$. There are signs such that*

$$\|\sum_{i=1}^N \sigma_i A_i\| \leq \min\{\|S\|, C_* \sqrt{\alpha \log(2r)}\}, \qquad C_* := \sqrt{\frac{213 \cdot 2^{1/3}}{25 \log 2}} < 3.93531. \tag{500}$$

Proof. At $y^0 = 0$, use $U_0 = V_0 = \lambda I_D$ in Lemma 11.9. The estimate $\sum_{i=1}^N \phi_i(I_D) A_i \preceq r^{1/q} \alpha I_D$ and minimization over $\lambda > 0$ give

$$\|\sum_{i=1}^N \sigma_i A_i\| \leq 2\sqrt{g r^{1/q} \alpha}.$$

The signing may depend on the chosen Schatten order. For rank one, the theorem of Xie, Xu, and Zhu [XXZ22, Theorem 1.3] instead gives $3\|\sum_{i=1}^N A_i^2\|^{1/2} \le 3\sqrt{\alpha}$. For $2 \le r \le 32$, use order three. The function $r^{1/3}/\log(2r)$ decreases and then increases. Its endpoint ratio is $2^{4/3}/3 < 1$, so rank two gives the maximum, namely C_*^2. For $r \ge 32$, use $q := \lfloor \log_2 r \rfloor \ge 5$. The normalized squared bound again has its maximum at an endpoint of $[2^q, 2^{q+1}]$. Its lower endpoint is $32/(3\log 2)$; the upper endpoint is no larger because

$$(1 + 1/(q+1))^q \ge 1 + \frac{q}{q+1} + \frac{q(q-1)}{2(q+1)^2} = 2 + \frac{q^2 - 3q - 2}{2(q+1)^2} > 2.$$

Finally $9 < (213/25)(5/4) < (213/25)2^{1/3}$, and $2 \cdot 639^3 > 800^3$ proves $32/3 < (213/25)2^{1/3}$. Both other rank ranges therefore lie below C_*^2. The exact logarithm and cube-root enclosures give $15.4866493670 < C_*^2 < 15.4866493671 < 3.93531^2$. They follow from the logarithm and integer-power bounds in Appendix B.7. For nonintegral r, apply the result to $\lfloor r \rfloor$. Every signing lies between $-S$ and S, which supplies the cap. □

11.7 Rank-sensitive moments and quartic budgets

We improve the existence estimate using the mixed covariance from Section 11.6. The changes are a rank-sensitive order-two comparison, a different order-three coupling, and a quartic budget at every integer order at least five.

Use the notation of that section, with Schatten order $q \ge 2$ and $\beta(y) := \sqrt{1-y^2} f(|y|)$. The functional Hessian estimate remains valid at order two. Indeed, for positive scalars a, b,

$$\sum_{j=0}^{q-2} a^j b^{q-2-j} \le \frac{q-1}{2}(a^{q-2} + b^{q-2}) \le \frac{q-1}{2}(a^{q-1}/b + b^{q-1}/a).$$

The first comparison is equality when $q = 2$. Differentiating the Schatten power, diagonalizing the compressed matrix, and retaining the negative derivative square therefore gives Eq. (341) for every integer $q \ge 2$. The support-inverse comparison used there follows by conjugating $A_i^{1/2}(A_i^{1/2} X A_i^{1/2})^{-1} A_i^{1/2}$ by $X^{1/2}$: the result is an orthogonal projection. All inverses in this expression are restricted to the support.

The potential regularity, endpoint tests, and response construction in Section 8.5 use these functional estimates, positive light-state scales, and a positive budget on the open interval. The budget and scale checks below verify these conditions. In particular, Eq. (486) and Eq. (488) hold with the new budget derivatives. The row completions in Lemma 11.8 use only $Z_e, Z_f > 0$ and these identities. We verify that positivity below instead of using the square-root-specific lower bound for Z_f. Thus the same lemma applies once $\mathcal{E}_i < 1$ is established. We write $c := q + 1$, $d := q - 1$, and $m := d/c$ below.

There is also a lower bound for the moments. If $\operatorname{tr}[R^q] = 1$, the power-mean inequality gives

$$\kappa := \operatorname{tr}[R^{q+1}] \ge (\operatorname{rank} R)^{-1/q}.$$

Together with Eq. (487), this gives $\kappa_U, \kappa_V \in [1/2, 1]$ at order two whenever every input has rank at most four. The two other families of Schatten-norm functionals only use $\kappa_U, \kappa_V \in [0, 1]$.

Budgets and positive denominators. The exact choices are

Case	Schatten order	Coupling g	Row ζ	Moment interval
Low rank	2	159/50	4	$[1/2, 1]$
Middle	3	17/4	5/2	$[0, 1]$
Higher	integer $q \geq 5$	$119q/100 + 4/5$	2	$[0, 1]$

Use $\beta(y) := \sqrt{1 - y^2} f(|y|)$, where

$$\begin{aligned} f_{\mathrm{low}}(t) &:= 1 - \frac{7}{125}t^2 - \frac{63}{1000}t^4, \\ f_{\mathrm{mid}}(t) &:= 1, \\ f_{\mathrm{high}}(t) &:= 1 + \frac{1}{25}t^2 - \frac{11}{100}t^4. \end{aligned} \tag{501}$$

All displayed coefficients are exact rational numbers. For $t \in [0, 1]$, put

$$z := 1 - t^2, \quad N := zf' - tf, \quad H := f + 2tzf' - z^2 f'', \quad A := f - zf', \quad B := f + zf'. \tag{502}$$

Then $\beta' = N/\sqrt{z}$ and $-\beta'' = H/z^{3/2}$. The ordinary-degree Bernstein coefficients on the single interval $[0, 1]$ give the following strictly positive lower bounds.

Polynomial	Low budget	Higher budget
f	881/1000	93/100
$-N/t$	881/1000	23/25
$(f^2 - z)/t^2$	$776161/10^6$	8649/10000
A	881/1000	93/100
B	4041/5000	429/500
H	699/1000	3/4
$d_0^2 AB + d_0 fH - N^2$	1847457/2500000	164331/10000

In the last row $d_0 := 1$ for the low budget and $d_0 := 4$ for the higher budget. Polynomial quotients at zero have their continuous extensions. The degrees of these seven polynomials are $4, 4, 6, 5, 5, 6, 10$. These checks imply

$$\beta(0) = 1, \quad \beta(\pm 1) = 0, \quad z \leq \beta \leq 1, \quad -\beta'' > 0$$

on actual states. In particular, at $y = t \geq 0$, lightness gives

$$p_i > 1 + (1 - t)\beta'/\beta = \frac{B}{(1 + t)f} > 0, \qquad q_i \geq 1, \quad \gamma > g\beta/z \geq g.$$

Negative coordinates exchange U, V and reverse the signs of both ℓ and T_{fe}, leaving the local cost unchanged.

To retain the entire light domain, let $U_0 := g\beta u \in (0, 1 + t)$ and $V_0 := g\beta v \in (0, 1 - t)$. Set $L_{\log} := -\beta'/\beta$ and introduce $a, b \in (0, 1)$ by

$$\frac{U_0}{1 + L_{\log} U_0} = \frac{zf}{A} a, \qquad \frac{V_0}{1 - L_{\log} V_0} = \frac{zf}{B} b.$$

Both maps are increasing bijections onto the displayed intervals. Equivalently,

$$U_0 := \frac{zfa}{A + Na}, \qquad V_0 := \frac{zfb}{B - Nb}. \tag{503}$$

Their denominators are positive for actual $0 \le t < 1$ and $0 < a, b < 1$; for example $A + N = (1-t)f > 0$ and $B - N = (1+t)f > 0$. Let

$$P := AB, \qquad y_0 := ab, \qquad S := aB + bA, \qquad D_L := (A + Na)(B - Nb).$$

Direct substitution gives

$$\alpha_{\mathrm{loc}} = \frac{aB}{S}, \qquad p_i q_i = \frac{P}{D_L}, \qquad K = \frac{gHD_L}{cz^{3/2}P}, \qquad \delta = \frac{z^{3/2} f y_0}{gD_L}, \qquad K\delta = \frac{fHy_0}{cP}, \quad \frac{\ell^2}{K\gamma} = \frac{N^2}{cfH}. \tag{504}$$

Also

$$2\ell T_{fe} = \frac{2N}{cPS}(a^2B^2\kappa_U - b^2A^2\kappa_V). \tag{505}$$

We now justify all response denominators, including order two. Put $q_0 := K\delta$ and $J := \ell^2/(K\gamma)$. The part of Z_f/γ without the nonnegative row terms is

$$m - \frac{q_0 J}{m + q_0} = \frac{d^2P + y_0(dfH - N^2)}{c(dP + fHy_0)}. \tag{506}$$

The numerator is affine in $y_0 \in [0,1]$. At its endpoints it is d^2P and $d^2P + dfH - N^2$, both positive by the table. For higher orders the latter increases with $d \ge 4$. In the middle case it is $6 - t^2 \ge 5$. Thus $Z_f > 0$ and also $Z_e > 0$.

The coefficients of e_U, e_V in Eq. (490) are positive. Substituting Eq. (487) therefore bounds the local cost by an affine function of (κ_U, κ_V). It suffices to check its four corners: $\{1/2, 1\}^2$ in the low-rank case and $\{0,1\}^2$ in the other cases.

Explicit polynomial certificates. Set $\omega := 10^{-6}$ in all three cases. The certified inequality is

$$\mathcal{E} \le 1 - \omega K w_e < 1. \tag{507}$$

The positive row margin supplies strict curvature at every actual state. Writing $\rho_w := w_f/w_e$ and $\sigma := \zeta/2$, the exact cost identity is

$$1 - \mathcal{E} = w_e(K - \mathcal{R}), \tag{508}$$

$$\mathcal{R} = (D_0)_{ee} + \rho_w (D_0)_{ff} + 2KT_{ee} + \frac{2\ell T_{fe}}{1 + \sigma\delta} + 2\sigma\rho_w T_{ff} + \sigma^2\delta\rho_w. \tag{509}$$

Hence it is enough to show $(1-\omega)K - \mathcal{R} \ge 0$ at each moment corner.

The two quartic budgets. Use the half-angle coordinate

$$t := \frac{1 - w^2}{1 + w^2}, \qquad \sqrt{1 - t^2} = \frac{2w}{1 + w^2}, \qquad 0 < w \le 1. \tag{510}$$

The closed certificate cube is $(w, a, b) \in [0,1]^3$. Let $Z := 1 + w^2$. In the following formulas, use the polynomial lifts

$$\widehat{f} := Z^4 f(t), \quad \widehat{N} := Z^5 N(t), \quad \widehat{H} := Z^6 H(t), \quad \widehat{A} := Z^5 A(t), \quad \widehat{B} := Z^5 B(t).$$

They are polynomials in w. For the rest of this quartic calculation, denote these five lifted polynomials by f, N, H, A, B. Define

$$h_0 := 8w^3 Z^3, \quad P := AB, \quad S := aB + bA, \quad y_0 := ab,$$

$$D_L := (A+Na)(B-Nb), \quad L := gD_L, \quad B_0 := L+\sigma h_0 f y_0, \tag{511}$$

$$E_0 := (cP+fHy_0)fHB_0^2 + N^2PL^2, \tag{512}$$

$$F_0 := (dP+fHy_0)[dL^2+2\sigma c h_0 f y_0 L + c\sigma^2 h_0^2 f^2 y_0^2] - y_0 N^2 L^2. \tag{513}$$

All denominators below are positive at actual states. In particular $F_0 > 0$ follows by rewriting its terms without σ as $L^2[d^2P + y_0(dfH - N^2)]$ and using Eq. (506); its remaining terms are nonnegative. Substitution into the weights gives

$$K = \frac{HL}{ch_0P}, \qquad \delta = \frac{h_0 f y_0}{L}, \qquad \rho_w = \frac{y_0(dP+fHy_0)E_0L^2}{cP^2B_0^2F_0}. \tag{514}$$

For a corner (i,j), where i,j are the chosen moment endpoints, put

$$M_{ij} := aBi + bAj, \qquad M_{ij}^{(2)} := a^2B^2i - b^2A^2j,$$

$$Q_{0,ij} := P[aB(q+i)+bA(q+j)] - 2fHy_0M_{ij}, \qquad Q_{1,ij} := bA(q+i)+aB(q+j).$$

The four polynomial inequalities are $\Phi_{ij} \geq 0$, where

$$\begin{aligned}\Phi_{ij} := {} & cP^2F_0[B_0^2((1-\omega)HLS - h_0Q_{0,ij}) - 2h_0NM_{ij}^{(2)}LB_0] \\ & - h_0y_0(dP+fHy_0)E_0L[PLQ_{1,ij} + cPh_0fy_0(2\sigma M_{ij}+\sigma^2 S)].\end{aligned} \tag{515}$$

Positive-denominator clearance gives exactly

$$\Phi_{ij} = c^2P^3Sh_0B_0^2F_0[(1-\omega)K - \mathcal{R}_{ij}]. \tag{516}$$

For a direct check, multiply Eq. (509) by cPS after the moment substitution. Its value is

$$Q_{0,ij} + \frac{2NM_{ij}^{(2)}L}{B_0} + \rho_w[PQ_{1,ij} + \frac{cPh_0fy_0}{L}(2\sigma M_{ij}+\sigma^2 S)].$$

Eq. (514) now gives Eq. (516).

For the low-rank case substitute $q = 2$, $g = 159/50$, $\sigma = 2$, and $(i,j) \in \{1/2,1\}^2$. Each polynomial has tridegree $(182,8,8)$. For the higher-order case substitute $q = 5+v$, $g = 119q/100 + 4/5$, $\sigma = 1$, and $(i,j) \in \{0,1\}^2$. Each resulting polynomial has degree eight in v:

$$\Phi_{ij}(w,a,b;5+v) = \sum_{j_0=0}^{8} C_{ij,j_0}(w,a,b)v^{j_0}. \tag{517}$$

All 36 coefficient polynomials are certified nonnegative. This proves the scalar inequality for every integer $q \geq 5$ at once. No finite Schatten-order cutoff is used.

The middle-order square-root budget. A smaller polynomial represents this case. Put $q := 3$, $c := 4$, $d := 2$, $g := 17/4$, $\sigma := 5/4$, and let

$$Z := 1+w^2, \qquad T_0 := 1-w^2, \qquad h_0 := 8w^3, \qquad y_0 := ab, \qquad S := a+b,$$
$$L := gZ(Z-T_0a)(Z+T_0b), \qquad B_0 := L + \sigma h_0 y_0.$$

Define

$$E_0 := (c+y_0)Z^2B_0^2 + T_0^2L^2,$$

$$F_0 := (d+y_0)Z^2[dL^2 + 2\sigma c h_0 y_0 L + c\sigma^2 h_0^2 y_0^2] - y_0 T_0^2 L^2.$$

For $(i,j) \in \{0,1\}^2$, set

$$M_{ij} := ai+bj, \quad M_{ij}^{(2)} := a^2 i - b^2 j, \quad Q_{0,ij} := a(q+i)+b(q+j)-2y_0 M_{ij}, \quad Q_{1,ij} := b(q+i)+a(q+j).$$

The four polynomials are

$$\begin{aligned}\Psi_{ij} := {} & cF_0[ZB_0^2((1-\omega)LS - h_0 Q_{0,ij}) + 2T_0 h_0 L B_0 M_{ij}^{(2)}] \\ & - Zh_0 y_0 (d+y_0) E_0 L[LQ_{1,ij} + ch_0 y_0 (2\sigma M_{ij} + \sigma^2 S)].\end{aligned} \tag{518}$$

They have tridegree $(36,8,8)$, and direct substitution gives

$$\Psi_{ij} = c^2 F_0 Z B_0^2 h_0 S[(1-\omega)K - \mathcal{R}_{ij}].$$

Thus their nonnegativity also proves Eq. (507).

Complete exact verification. For a polynomial $P(x_1,x_2,x_3) = \sum_\nu M_\nu x^\nu$ with coordinate degrees at most (n_1,n_2,n_3), the tensor Bernstein coefficients on the unit cube are

$$B_\mu = \sum_{\nu \le \mu} M_\nu \prod_{j=1}^{3} \frac{\binom{\mu_j}{\nu_j}}{\binom{n_j}{\nu_j}}. \tag{519}$$

On another box first make its affine substitution. The basis functions are nonnegative and sum to one; nonnegative coefficients prove nonnegativity throughout the box.

The complete preorder tree words are

Low-rank quartic	`TPP`
Middle square-root	`TPVPTPP`
Higher-order quartic family	`TPTPP`

A letter `T`, `U`, or `V` bisects the first, second, or third coordinate and is followed by its left and right subtree. A `P` is an accepted leaf. The first coordinate is w in all cases. Every low-rank leaf checks four polynomials Eq. (515); every middle leaf checks four polynomials Eq. (518); every higher-order leaf checks all 36 polynomials Eq. (517).

Case	Polynomials	Boxes	Coefficient checks	Zeros
Low rank	4	2	118,584	2,544
Middle	4	4	47,952	1,376
Higher orders	36	3	1,455,948	29,004
Total	44	9	1,622,484	32,924

All tested coefficients are nonnegative. Every tree is consumed exactly, every split partitions its parent, and each family's leaf volumes sum exactly to one. These are full-cube certificates, including the boundaries. The 32,924 zero coefficients are accepted deliberately: strictness at actual states follows from the separately retained margin $\omega K w_e > 0$, not from claiming that every closed-cube coefficient is positive.

The polynomials and complete trees above determine every coefficient inequality. The rules in Appendix B.1 apply to each leaf at its stated coordinate degrees. Exact rational comparison of all 1,622,484 coefficients gives the nonnegativity and zero counts above.

Lemma 11.11. *For every stationary light state, Eq. (507) holds for the low-rank parameters when all input ranks are at most four, for the middle parameters without a rank restriction, and for the higher-order parameters for every integer $q \geq 5$ without a rank restriction.*

Proof. The univariate tests prove all budget, scale, and denominator conditions. The moment inequalities reduce each local cost to its four specified corners. The complete coefficient certificates and the positive factors in Eq. (516) or its middle-order counterpart prove $(1-\omega)K \geq \mathcal{R}$ at these corners. Affineness extends this to the actual moments. Eq. (508) proves the assertion. □

Lemma 11.12. *Let $A_1, \dots, A_N \succeq 0$ be a finite family. Use one of the three parameter and budget choices above. For the order-two choice, assume that every input has rank at most four. For any $y^0 \in [-1,1]^N$ and $U_0, V_0 \succ 0$, there is one signing satisfying*

$$M(\sigma) \preceq M(y^0) + U_0^{-1} + F_{y^0}(V_0), \tag{520}$$

$$-M(\sigma) \preceq -M(y^0) + V_0^{-1} + F_{y^0}(U_0). \tag{521}$$

Proof. Freeze the two right sides as H_+, H_-. In the potential replace the upper constraint targets by $bI + H_+$ and $bI + H_-$. The trial $b = 0$, $U = U_0$, $V = V_0$ is feasible at y^0. Adding the constraints bounds b below by $-\|H_+ + H_-\|/2$. Objective sublevels then bound b, U, V above, and the constraints bound their inverses above. These bounds are uniform over the cube. Thus attainment and continuity follow as in Lemma 8.3. The fixed targets have zero derivative. The same strict dual equations, invertible constraint derivative, and positive primal Hessian therefore give uniqueness and smoothness on each open face. The endpoint tests, response identities, and Eq. (486) are unchanged.

Choose a global minimizer with the fewest active coordinates. An available endpoint move contradicts this choice. Otherwise the point is light; if it is not a vertex, its active Hessian is positive semidefinite. Lemma 11.11 and the extended covariance comparison above contradict this assertion. Hence a minimizing vertex exists. Its value is at most $\rho \operatorname{tr}[U_0 + V_0]$. At a vertex the coupling vanishes, so the two target inequalities hold with this additional scalar slack. Let $\rho \downarrow 0$ and choose a recurrent signing from the finite set of signings. This removes the slack and proves both bounds. □

Corollary 11.13. *Suppose $S := \sum_{i=1}^N A_i \preceq I_D$, $A_i \succeq 0$, $\|A_i\| \leq \alpha$, and $\operatorname{rank} A_i \leq r$, where $r \geq 1$. There are signs such that*

$$\|\sum_{i=1}^N \sigma_i A_i\| \leq \min\{\|S\|, C_* \sqrt{\alpha \log(2r)}\}, \qquad C_* := \frac{3}{\sqrt{\log 2}} < 3.60337.$$

Proof. At $y^0 = 0$, put $U_0 = V_0 = \lambda I_D$ in Lemma 11.12. Since $\sum_{i=1}^N \phi_i(I_D) A_i \preceq r^{1/q} \alpha I_D$, minimization over $\lambda > 0$ gives

$$\|\sum_{i=1}^N \sigma_i A_i\| \leq 2\sqrt{g r^{1/q} \alpha}.$$

The order-two choice requires $r \leq 4$; the other choices apply at every rank. The signing may depend on the choice.

Work first with integer $r \geq 1$ and normalize squared discrepancy by $\alpha \log(2r)$. For fixed q,

$$\frac{d}{d \log r} \log\left(\frac{r^{1/q}}{\log(2r)}\right) = \frac{1}{q} - \frac{1}{\log(2r)}.$$

The expression decreases and then increases, so its maximum on a closed interval is at an endpoint.

For rank one, apply [XXZ22, Theorem 1.3] to independent uniform signs. It gives

$$\|\sum_{i=1}^{N}\sigma_i A_i\| \le 3\|\sum_{i=1}^{N} A_i^2\|^{1/2} \le 3\sqrt{\alpha},$$

since $A_i^2 \preceq \alpha A_i$. Its normalized square is $9/\log 2 = C_*^2$.

For $2 \le r \le 4$, use the order-two estimate. The endpoint $r = 2$ is the larger one, since $3\sqrt{2} > 4$. Its normalized square is

$$\frac{159\sqrt{2}}{25\log 2} < \frac{9}{\log 2},$$

with exact comparison

$$225^2 - 2\cdot 159^2 = 63 > 0. \tag{522}$$

For $4 \le r \le 32$, use the order-three estimate. The endpoint values are equal, giving

$$\frac{17\cdot 4^{1/3}}{3\log 2} < \frac{9}{\log 2}, \qquad 27^3 - 4\cdot 17^3 = 31 > 0.$$

For $32 \le r \le 63$, take $q := 5$ in the higher-order estimate. Its coupling is exactly $27/4$. Because $\log(2r) < 7\log 2 < 5$ here, the normalized expression decreases. At $r = 32$ it equals

$$\frac{27\cdot 32^{1/5}}{\log 64} = \frac{9}{\log 2}.$$

For $r \ge 64$, put $x := \log r$, $a_0 := \log 2$, $q := \lceil x\rceil \ge 5$, and $b_0 := 80/119$. Then

$$\frac{4g_q r^{1/q}}{\log(2r)} = \frac{119}{25}\frac{(q+b_0)e^{x/q}}{x+a_0}.$$

On $q-1 \le x \le q$, the last ratio has its maximum at an endpoint. At $x = q$ it is less than e, since $b_0 < a_0$. At $x = q-1$ use

$$\begin{aligned}(q-1+a_0)e^{1/q} - (q+b_0) &\ge a_0 - b_0 + \frac{a_0 - 1/2}{q} + \frac{a_0 - 1}{2q^2}\\ &> a_0 - b_0 + \frac{q-1}{6q^2} > 0.\end{aligned}$$

Here $e^{1/q} \ge 1 + 1/q + 1/(2q^2)$ and $a_0 > 2/3$ suffice. Thus the higher-order envelope is at most $119e/25$, strictly smaller than C_*^2. An elementary rational check is

$$e < \frac{68}{25}, \quad \log 2 < \frac{347}{500}, \qquad 119\cdot 68\cdot 347 < 225\cdot 25\cdot 500.$$

This proves $119e\log 2 < 225$. All arbitrarily large ranks are included; no finite rank scan is used.

The complete comparison is summarized by

Rank range	Upper bound for the normalized square
$r = 1$	$9/\log 2$
$2 \le r \le 4$	$159\sqrt{2}/(25\log 2) < 9/\log 2$
$4 \le r \le 32$	$17\cdot 4^{1/3}/(3\log 2) < 9/\log 2$
$32 \le r \le 63$	$9/\log 2$
$r \ge 64$	$119e/25 < 9/\log 2$

Exact rational evaluation gives

$$12.9842553680 < C_*^2 < 12.9842553681, \qquad 3.6033672263 < C_* < 3.6033672264 < 3.60337.$$

For the logarithm, use 100 terms of

$$\log v = 2\sum_{j=0}^{99} \frac{z^{2j+1}}{2j+1} + R, \quad z := \frac{v-1}{v+1}, \qquad 0 \le R \le \frac{2z^{201}}{201(1-z^2)},$$

with range reduction to $[1,2]$. For e, use the sum through $1/50!$ and the geometric remainder bound $1/(51!)(1-1/52)^{-1}$. All decisions and displayed decimal enclosures are checked as rational inequalities.

For nonintegral r, apply the result to $\lfloor r \rfloor$ and then increase the logarithmic bound. Finally every signing satisfies $-S \preceq \sum_{i=1}^N \sigma_i A_i \preceq S$, giving the cap by $\|S\|$. □

11.8 Sharper rank-sensitive budgets

We sharpen the existence coefficient using the response covariance of Section 11.6. The five budgets below use the same one-sign-per-input potential. The rank-two budget is rational, and the higher-order budget applies to every real Schatten order at least five. The formulas, parameters, and complete certificate trees are given below.

The differential inequality for real Schatten orders. The higher-order selection uses a real Schatten order. We give the derivative estimate explicitly because the earlier integer-order statement alone does not suffice.

Lemma 11.14. *For every real $s \ge 2$, nonzero $A \succeq 0$, and $X \succ 0$, put*

$$C = A^{1/2} X A^{1/2}, \qquad \phi(X) = (\operatorname{tr}[C^s])^{1/s}, \qquad B(X) = \frac{A^{1/2} C^{s-1} A^{1/2}}{\phi(X)^{s-1}}.$$

Powers and inverses on a support are restricted to that support. The functional ϕ is smooth, convex, and homogeneous of degree one. For every Hermitian H,

$$D\phi(X)[H] = \operatorname{tr}[B(X)H], \quad 0 \preceq B(X) \preceq A, \quad \operatorname{tr}[XB(X)] = \phi(X), \tag{523}$$

$$(D\phi(X)[H])^2 \le \phi(X) \operatorname{tr}[B(X) H X^{-1} H], \tag{524}$$

$$0 \le D^2\phi(X)[H,H] \le (s-1)[\operatorname{tr}[B(X)HX^{-1}H] - \frac{(D\phi(X)[H])^2}{\phi(X)}]. \tag{525}$$

Also, $\operatorname{tr}[WAXA] \le \phi(X) \operatorname{tr}[WA]$ *for* $W \succeq 0$.

Proof. The derivative follows by differentiating the trace. Since $C \preceq \phi I$, its spectral powers give $B \preceq A$ and the final assertion. Homogeneity gives the trace identity, and weighted Hilbert–Schmidt Cauchy–Schwarz gives (524). Convexity is the convexity of the Schatten norm applied to a linear compression.

For the real-order second derivative set $K = A^{1/2} H A^{1/2}$ and diagonalize C on its positive support, with eigenvalues λ_a. Continuous divided differences at repeated eigenvalues give

$$D^2\phi(C)[K,K] = \phi^{1-s} \sum_{a,b} \frac{\lambda_a^{s-1} - \lambda_b^{s-1}}{\lambda_a - \lambda_b} |K_{ab}|^2 - (s-1)\phi^{1-2s} (\operatorname{tr}[C^{s-1}K])^2. \tag{526}$$

For positive a, b and $s \geq 2$,

$$\frac{a^{s-1} - b^{s-1}}{a - b} \leq \frac{s-1}{2}\left(\frac{a^{s-1}}{b} + \frac{b^{s-1}}{a}\right).$$

Indeed, write $a = \sqrt{ab}e^{u}$, $b = \sqrt{ab}e^{-u}$ with $u \geq 0$, and $q = s - 1$. For $u > 0$ the assertion follows from

$$\frac{\sinh(qu)}{\sinh u} \leq q\cosh((q+1)u), \qquad \sinh(qu) \leq qu\cosh(qu) \leq q\sinh u\cosh((q+1)u).$$

The case $u = 0$ follows by a limit. Thus the positive term in Eq. (526) is at most

$$(s-1)\phi^{1-s}\operatorname{tr}[C^{s-1}KC^{-1}K].$$

The compression inequality $A^{1/2}C^{-1}A^{1/2} \preceq X^{-1}$ follows because conjugation by $X^{1/2}$ turns the left side into an orthogonal projection. Apply this inside the positive trace and retain the negative square in (526) to obtain (525). These arguments do not use integrality of s. □

Potential, response geometry, and mixed variation. Discard zero inputs and assign them arbitrary signs at the end. Fix $s \geq 2$ and a coupling $g > 2$. Apply Lemma 11.14 to each A_i to define ϕ_i, B_i. For an even budget β with $\beta(0) = 1$, $\beta(\pm 1) = 0$ and $\beta > 0$ inside, use

$$F_y(X) = g\sum_i \beta(y_i)\phi_i(X)A_i, \qquad M(y) = \sum_i y_i A_i,$$

with fixed contributions retained in M on every face. For $\rho > 0$ let

$$P_\rho(y) = \min_{b\in\mathbb{R},\ U,V\succ 0}\{b + \rho\operatorname{tr}[U+V] : \begin{array}{l} U^{-1} + M(y) + F_y(V) \preceq bI, \\ V^{-1} - M(y) + F_y(U) \preceq bI \end{array}\}. \tag{527}$$

The regularity proof in Section 8.5 uses exactly the properties of the Schatten-norm functional in Lemma 11.14. Compact sublevels and strict feasibility give attainment and a continuous value on the closed cube. The strict inverse-map Hessian and invertible positive linearization give a unique smooth optimizer on every open face. In particular, these assertions apply to real orders as well as integer ones.

Write $u_i = \phi_i(U)$, $v_i = \phi_i(V)$, $w_i = \operatorname{tr}[A_iW]$, $z_i = \operatorname{tr}[A_iZ]$ for the positive primal and dual scalars. The dual matrices satisfy

$$W = U(\rho I + g\sum_i \beta_i z_i B_i(U))U, \quad Z = V(\rho I + g\sum_i \beta_i w_i B_i(V))V, \quad \operatorname{tr}[W+Z] = 1.$$

They are positive definite and both primal constraints are equalities. A move to $+1$ is nonincreasing if $g\beta_i v_i \geq 1 - y_i$; a move to -1 is nonincreasing if $g\beta_i u_i \geq 1 + y_i$. The remaining state is *light*:

$$g\beta_i u_i < 1 + y_i, \qquad g\beta_i v_i < 1 - y_i. \tag{528}$$

Set

$$\begin{aligned} p_i &= 1 + g\beta_i' v_i, & q_i &= 1 - g\beta_i' u_i, \\ c_i &= \sqrt{\beta_i p_i/q_i}, & d_i &= \sqrt{\beta_i q_i/p_i}, \\ \tau_i &= \sqrt{p_i q_i/\beta_i}, & \gamma_i &= (g\beta_i u_i v_i)^{-1}. \end{aligned}$$

The budget checks below imply positivity of these scales and $\gamma_i > g > 2$. At outer stationarity, $p_i w_i = q_i z_i$ and $c_i w_i = d_i z_i$.

Holding b fixed, genuine first responses take the form $\dot{U} = UM_aU$, $\dot{V} = VM_bV$, with $M_a = \sum_i c_i a_i A_i$, $M_b = \sum_i d_i b_i A_i$. Their coefficients obey

$$a - \mathcal{V}b = \mathcal{U}a - b = \tau h, \quad \mathcal{U}_{ij} = gc_ic_j \operatorname{tr}[B_i(U)UA_jU], \quad \mathcal{V}_{ij} = gd_id_j \operatorname{tr}[B_i(V)VA_jV].$$

The matrices $\mathcal{U}, \mathcal{V}$ have nonnegative entries but need not be symmetric. The strict transposed dual comparisons imply a bijective response graph of dimension k_{act}, the active count, even when the inputs are linearly dependent. Define

$$E_a = \operatorname{diag}(c_i^2 u_i w_i), \quad E_b = \operatorname{diag}(d_i^2 v_i z_i), \quad E = \operatorname{diag}(E_a, E_b), \quad \xi = E^{1/2}(a, b),$$

$$T = \begin{pmatrix} 0 & E_a^{1/2}\mathcal{V}E_b^{-1/2} \\ E_b^{1/2}\mathcal{U}E_a^{-1/2} & 0 \end{pmatrix}.$$

Keep the *unpadded* energy matrix D_0, defined by

$$Q = \operatorname{tr}[WM_aUM_a] + \operatorname{tr}[ZM_bVM_b] = \xi^\top D_0 \xi.$$

The dual equations and (524) imply

$$D_0 \succeq T^\top \Gamma T, \qquad \Gamma|_{\text{pair } i} = \gamma_i I_2, \qquad \operatorname{tr}[D_0] \le 2k_{\mathrm{act}}. \tag{529}$$

In each coordinate pair use

$$e_i = \frac{(\sqrt{E_{a,i}}, -\sqrt{E_{b,i}})}{\sqrt{E_{a,i} + E_{b,i}}}, \qquad f_i = \frac{(\sqrt{E_{b,i}}, \sqrt{E_{a,i}})}{\sqrt{E_{a,i} + E_{b,i}}}.$$

On the response graph, stationarity gives

$$F_i := f_i^\top \xi = f_i^\top T\xi, \qquad Y_i := e_i^\top (I - T)\xi = \sqrt{p_i q_i S_i} h_i, \qquad S_i = v_i w_i + u_i z_i > 0. \tag{530}$$

Let $c = s + 1$, $d = s - 1$, $m = d/c$, and define the local scalars

$$k_i = \frac{g(-\beta_i'')}{cp_iq_i}, \qquad \ell_i = \frac{\beta_i'}{c}\sqrt{\frac{g\gamma_i}{\beta_i p_i q_i}}. \tag{531}$$

The subscripted k_i is not the active count. Differentiating the fixed-b comparison path and using (525) yields

$$\frac{1}{c} D^2 P_\rho[h, h] \le Q - \sum_i m\gamma_i[(e_i^\top T\xi)^2 + (f_i^\top T\xi)^2] + 2\sum_i \ell_i Y_i F_i - \sum_i k_i Y_i^2. \tag{532}$$

The inverse terms contribute $2Q$; the functional Hessian contributes at most $(s-1)Q$ minus its retained response squares. At stationarity the exact mixed term is $2\beta_i'\sqrt{g\gamma_i/(\beta_i p_i q_i)}Y_iF_i$ and the budget term is $g\beta_i'' S_i h_i^2$. Division by c gives (532). No Young inequality is applied to this mixed term.

Moments and the direct covariance. For each coordinate let

$$R_U = A_i^{1/2} U A_i^{1/2}/u_i, \qquad S_U = A_i^{1/2} W A_i^{1/2}/w_i,$$

and define the V, Z counterparts. Put $e_U = \operatorname{tr}[R_U S_U]$, $\kappa_U = \operatorname{tr}[R_U^{s+1}]$, and similarly for V. Since $\operatorname{tr}[R_U^s] = 1$ and $\operatorname{tr}[S_U] = 1$,

$$(\operatorname{rank} R_U)^{-1/s} \le \kappa_U \le 1, \qquad 0 \le e_U \le \frac{s+\kappa_U}{s+1}. \tag{533}$$

Power means give the lower bound. For the second bound, $x^{s+1} - (s+1)x + s \ge 0$ is the tangent inequality at one. Apply it at the largest eigenvalue of R_U, then use $e_U \le \lambda_{\max}(R_U)$ and $\lambda_{\max}(R_U)^{s+1} \le \kappa_U$. The V inequalities are identical.

Suppress an index and put $\alpha_{\mathrm{loc}} = c_i u_i/(c_i u_i + d_i v_i)$ and $\delta = 1/\gamma$. Rotation of the local response and energy blocks gives

$$(D_0)_{ee} = \alpha_{\mathrm{loc}} e_U + (1-\alpha_{\mathrm{loc}}) e_V, \qquad (D_0)_{ff} = (1-\alpha_{\mathrm{loc}}) e_U + \alpha_{\mathrm{loc}} e_V, \tag{534}$$

$$T_{ee} = -\delta[\alpha_{\mathrm{loc}}\kappa_U + (1-\alpha_{\mathrm{loc}})\kappa_V], \qquad T_{ff} = -T_{ee}, \tag{535}$$

$$T_{fe} = \delta \frac{\alpha_{\mathrm{loc}}^2 \kappa_U - (1-\alpha_{\mathrm{loc}})^2 \kappa_V}{\sqrt{\alpha_{\mathrm{loc}}(1-\alpha_{\mathrm{loc}})}}. \tag{536}$$

For a direct verification write $x_0 = c_i u_i$, $v_0 = d_i v_i$ and $\omega_0 = c_i w_i = d_i z_i$. Then $E_{a,i} = x_0\omega_0$, $E_{b,i} = v_0\omega_0$, $\mathcal{U}_{ii} = g x_0^2 \kappa_U$, $\mathcal{V}_{ii} = g v_0^2 \kappa_V$ and $g x_0 v_0 = \delta$. The raw energy diagonal is (e_U, e_V). These facts yield all three identities without a symmetry assumption on T.

Choose a nonnegative row parameter ζ, set $\sigma = \zeta/2$, and define

$$\nu = \frac{\ell}{1+\sigma\delta}, \qquad Z_e = k + \delta(k^2+\nu^2), \tag{537}$$

$$Z_f = m\gamma + \zeta + \sigma^2\delta - \frac{\ell^2}{m\gamma + k}, \qquad w_e = Z_e^{-1}, \qquad w_f = Z_f^{-1}. \tag{538}$$

Positivity for every selected budget is proved below. Define the local cost

$$\begin{aligned}\mathcal{E} = {} & w_e (D_0)_{ee} + w_f (D_0)_{ff} + \delta[(k^2+\nu^2) w_e + \sigma^2 w_f] \\ & + 2k w_e T_{ee} + 2\nu w_e T_{fe} + 2\sigma w_f T_{ff}.\end{aligned} \tag{539}$$

Lemma 11.15 (Direct mixed-term covariance)**.** *Suppose $Z_e, Z_f > 0$ and $\mathcal{E}_i < 1$ for every active coordinate at a stationary light state. There is a covariance of genuine physical responses with strictly negative expected second variation of P_ρ.*

Proof. Let D be the diagonal matrix of the pair weights $w_{e,i}, w_{f,i}$. If R has orthonormal columns spanning the response graph, set

$$\Sigma = R(R^\top D^{-1} R)^{-1} R^\top, \qquad B = D - \Sigma.$$

Then $\Sigma, B \succeq 0$, $\Sigma D^{-1}\Sigma = \Sigma$, $\Sigma D^{-1} B = 0$, and $\operatorname{tr}[D^{-1}\Sigma] = k_{\mathrm{act}}$. The covariance is positive definite on its graph and has a centered finite-support realization.

For an e-row $t_e = e^\top T$, complete the square with the auxiliary vector ke and $C_* = \gamma B + (m\gamma + k)\Sigma \succ 0$. The projection identities give

$$C_*^{-1} = \delta D^{-1} + ((m\gamma+k)^{-1} - \delta) D^{-1}\Sigma D^{-1}.$$

The resulting row inequality is

$$\gamma t_e B t_e^\top + m\gamma t_e \Sigma t_e^\top + k(e^\top - t_e)\Sigma(e - t_e^\top) - 2\ell(e^\top - t_e)\Sigma f + 2t_e D(ke)$$
$$\geq (k + \delta k^2)e^\top \Sigma e - 2\ell e^\top \Sigma f - \frac{\ell^2}{m\gamma + k} f^\top \Sigma f - \delta k^2 w_e.$$

It can be checked by minimizing the left side in t_e: its linear vector is $k\Sigma e - \ell\Sigma f - D(ke)$ and the displayed inverse gives the stated result.

For the f-row $t_f = f^\top T$, use $t_f\Sigma = f^\top\Sigma$ and auxiliary vector $\nu e + \sigma f$. Completing a square in the positive semidefinite B, without inverting it, gives

$$\gamma t_f B t_f^\top + m\gamma f^\top \Sigma f + 2t_f D(\nu e + \sigma f)$$
$$\geq \delta\nu^2 e^\top \Sigma e + 2\nu(1 + \sigma\delta)e^\top \Sigma f + (m\gamma + 2\sigma + \sigma^2\delta) f^\top \Sigma f - \delta(\nu^2 w_e + \sigma^2 w_f).$$

The off-diagonal coefficients cancel because $\nu(1 + \sigma\delta) = \ell$. The remaining diagonal coefficients are Z_e, Z_f, exactly the local entries of D^{-1}. Sum over all coordinates. The added linear terms are those in (539). Using (529) to pass from D to Σ yields

$$\sum_i k_i \mathbb{E} Y_i^2 - 2\sum_i \ell_i \mathbb{E}(Y_i F_i) - [\mathbb{E} Q - \sum_i m\gamma_i \mathbb{E}((e_i^\top T\xi)^2 + (f_i^\top T\xi)^2)] \geq \sum_i (1 - \mathcal{E}_i) > 0. \tag{540}$$

Equation (532) proves the conclusion. The proof is rowwise: T need not be symmetric, and D, Σ need not commute with Γ. All helpers are chosen pointwise and are never differentiated. □

The exact budgets and couplings. All terminating decimals in this section denote exact rational numbers. Write

$$\beta(y) = \sqrt{1 - y^2}\frac{P(|y|)}{Q(|y|)}.$$

The five cases are as follows. A moment interval $[\chi, 1]$ is used for each of κ_U, κ_V; the rank-free cases use $\chi = 0$.

Case	Order s	Coupling g	Row ζ	Moment/rank condition
Rank two	2	3	423/100	$\chi = 7071/10000$, rank ≤ 2
Rank three	2	63/20	197/50	$\chi = 5773/10000$, rank ≤ 3
Order three	3	79/20	12/5	$\chi = 5/16$, rank ≤ 32
Order four	4	103/20	5/2	$\chi = 0$, arbitrary rank
Higher orders	Real $s \geq 5$	$9s/8 + 803/1000$	2	$\chi = 0$, arbitrary rank

The required finite-rank comparisons are the exact rational inequalities

$$2(7071/10000)^2 < 1, \qquad 3(5773/10000)^2 < 1, \qquad 32(5/16)^3 = 125/128 < 1.$$

They imply the asserted moment intervals by (533).

For rank two the numerator and denominator are

$$P(t) = 1 - 1.84070402t^2 + 0.84514081t^4 + 0.01525213t^6 - 0.01630284t^8, \tag{541}$$
$$Q(t) = 1 - 1.75120238t^2 + 0.75880637t^4. \tag{542}$$

In the next three cases $Q = 1$ and, respectively,

$$P_{\text{rank3}}(t) = 1 - 0.0523t^2 - 0.0706t^4, \tag{543}$$

$$P_{\text{order3}}(t) = 1 - \tfrac{1}{25}t^2 - \tfrac{9}{100}t^4, \tag{544}$$
$$P_{\text{order4}}(t) = 1 - 0.01484449t^2 - 0.08000936t^4. \tag{545}$$

For every higher order use the same degree-24 polynomial and $Q = 1$:

$$P(t) = 1 + \sum_{j=1}^{12} c_j[T_j(2t^2 - 1) - (-1)^j], \tag{546}$$

where $T_0(x) = 1$, $T_1(x) = x$, and $T_{j+1}(x) = 2xT_j(x) - T_{j-1}(x)$. The coefficients are

j	c_j	j	c_j
1	−0.181400905612	7	−0.003407077546
2	−0.097022897608	8	−0.001720434829
3	−0.044428932542	9	−0.000796274005
4	−0.023147341672	10	−0.000328483808
5	−0.012136805760	11	−0.000106750170
6	−0.006532785802	12	−0.000023122913

The representation guarantees $P(0) = 1$. The pointwise row parameters are not differentiated.

Rational-budget identities and denominator control. Let $t = |y|$, $z = 1 - t^2$, and $f = P/Q$. Define the polynomial derivative numerators

$$D_1 = P'Q - PQ', \qquad N = zD_1 - tPQ, \tag{547}$$
$$A = PQ - zD_1, \qquad B = PQ + zD_1, \tag{548}$$
$$H = PQ^2 + 2tzQD_1 - z^2(D_1'Q - 2D_1Q'). \tag{549}$$

Then

$$\beta' = \frac{N}{\sqrt{z}Q^2}, \qquad -\beta'' = \frac{H}{z^{3/2}Q^3}. \tag{550}$$

These are numerators; for example the physical expression $f - zf'$ is A/Q^2, not A.

For each case the exact univariate checks prove

$$P > 0,\ Q > 0,\ -N/t > 0,\ (P^2 - zQ^2)/t^2 > 0,\ A > 0,\ B > 0,\ H > 0, \quad d_0^2AB + d_0PH - N^2 > 0 \tag{551}$$

on $[0, 1]$, where $d_0 = s - 1$ in a finite case and $d_0 = 4$ for the higher family. Quotients at zero are their polynomial continuations. The precise certificates are specified in Section 11.8. These imply

$$\beta(0) = 1, \quad \beta(\pm 1) = 0, \quad 1 - y^2 \le \beta(y) \le 1, \quad -\beta''(y) > 0 \quad (-1 < y < 1).$$

Indeed, $P/Q \ge \sqrt{z}$ and the derivative is nonpositive on $t \ge 0$. For positive $y = t$, lightness gives

$$p > 1 + (1 - t)\beta'/\beta = \frac{B}{(1 + t)PQ} > 0, \qquad q \ge 1, \qquad \gamma > g\beta/z \ge g > 2.$$

Negative coordinates exchange U, V and reverse the signs of both ℓ and T_{fe}, leaving the local cost unchanged. The budgets are smooth in the open interval and even; there is no differentiability issue at zero.

To retain the full light region, parameterize $U_0 = g\beta u$ and $V_0 = g\beta v$ by independent $a, b \in (0,1)$:

$$U_0 = \frac{zPQa}{A + Na}, \qquad V_0 = \frac{zPQb}{B - Nb}. \tag{552}$$

Each map is an increasing bijection onto $(0, 1+t)$ or $(0, 1-t)$, respectively. For example $A + N = (1-t)PQ$ and $B - N = (1+t)PQ$, and the derivative denominators and numerators are positive. Put

$$P_0 = AB, \quad y_0 = ab, \quad S = aB + bA, \quad D_L = (A + Na)(B - Nb).$$

Direct substitution, using the physical derivatives in (550), gives

$$\alpha_{\mathrm{loc}} = \frac{aB}{S}, \qquad pq = \frac{P_0}{D_L}, \tag{553}$$

$$k = \frac{gHD_L}{cz^{3/2}Q^3P_0}, \qquad \delta = \frac{z^{3/2}PQ^3y_0}{gD_L}, \tag{554}$$

$$k\delta = \frac{PHy_0}{cP_0}, \qquad J := \frac{\ell^2}{k\gamma} = \frac{N^2}{cPH}. \tag{555}$$

The lower off-diagonal mixed term simplifies to

$$2\ell T_{fe} = \frac{2N}{cP_0S}(a^2B^2\kappa_U - b^2A^2\kappa_V). \tag{556}$$

In particular this expression is rational and has no remaining square root.

For weight positivity put $q_0 = k\delta$. The part of Z_f/γ without its nonnegative row-parameter terms is exactly

$$m - \frac{q_0 J}{m + q_0} = \frac{d^2P_0 + y_0(dPH - N^2)}{c(dP_0 + PHy_0)}. \tag{557}$$

Its numerator is affine in $0 \le y_0 \le 1$. At the two endpoints it is d^2P_0 and $d^2P_0 + dPH - N^2$, both positive by (551). For the higher family the latter increases with $d \ge 4$. Thus $Z_f > 0$, and $Z_e > 0$ follows from $k > 0$. This proves every response denominator without requiring the stronger, unnecessary inequality $dPH - N^2 \ge 0$.

The coefficients of e_U, e_V in (539) are positive. Substituting (533) therefore bounds $\mathcal{E}$ by an affine function of $(\kappa_U, \kappa_V) \in [\chi, 1]^2$. It suffices to check its four corners.

The explicit polynomial comparison. Use the exact relative row margin

$$\omega = 10^{-6}, \qquad \mathcal{E} \le 1 - \omega k w_e < 1. \tag{558}$$

Writing $\rho_w = w_f/w_e$, the local cost satisfies

$$1 - \mathcal{E} = w_e(k - \mathcal{R}), \tag{559}$$

$$\mathcal{R} = (D_0)_{ee} + \rho_w(D_0)_{ff} + 2kT_{ee} + \frac{2\ell T_{fe}}{1 + \sigma\delta} + 2\sigma\rho_w T_{ff} + \sigma^2\delta\rho_w. \tag{560}$$

It is enough to prove $(1-\omega)k - \mathcal{R} \ge 0$ at each moment corner.

Use the half-angle coordinate

$$t = \frac{1 - w^2}{1 + w^2}, \qquad \sqrt{1 - t^2} = \frac{2w}{1 + w^2}, \qquad 0 < w \le 1. \tag{561}$$

The closed certificate cube is $(w,a,b)\in[0,1]^3$. Let $Z=1+w^2$, $p_0=\deg P$, $q_0=\deg Q$ and $h=p_0+q_0+1$. The selected parameters satisfy $p_0\geq q_0+1$. Define polynomial lifts

$$\widehat{P}=Z^{p_0}P(t),\quad \widehat{Q}=Z^{q_0}Q(t),\quad (\widehat{N},\widehat{A},\widehat{B})=Z^{h}(N,A,B)(t),\quad \widehat{H}=Z^{p_0+2q_0+2}H(t).$$

The lifting powers bound the actual degrees, so these are polynomials in w. For the rest of this section abbreviate the lifted quantities by $f=\widehat{P}$, $q=\widehat{Q}$, $N=\widehat{N}$, $A=\widehat{A}$, $B=\widehat{B}$, $H=\widehat{H}$. In particular f here is a lifted numerator, not the original rational factor P/Q. Set

$$h_0=8w^3Z^{p_0-q_0-1}q^3,\quad P_0=AB,\quad S=aB+bA,\quad y_0=ab,$$
$$D_L=(A+Na)(B-Nb),\quad L=gD_L,\quad B_0=L+\sigma h_0fy_0, \tag{562}$$
$$E_0=(cP_0+fHy_0)fHB_0^2+N^2P_0L^2, \tag{563}$$
$$F_0=(dP_0+fHy_0)(dL^2+2\sigma ch_0fy_0L+c\sigma^2h_0^2f^2y_0^2)-y_0N^2L^2. \tag{564}$$

At actual states all denominator factors are positive. For F_0, its terms without σ are $L^2[d^2P_0+y_0(dfH-N^2)]$, positive by the lifted version of (557); the remaining terms are nonnegative. Substitution gives

$$k=\frac{HL}{ch_0P_0},\qquad \delta=\frac{h_0fy_0}{L},\qquad \rho_w=\frac{y_0(dP_0+fHy_0)E_0L^2}{cP_0^2B_0^2F_0}. \tag{565}$$

For each corner $(i,j)\in\{\chi,1\}^2$, put

$$M_{ij}=aBi+bAj,\qquad M_{ij}^{(2)}=a^2B^2i-b^2A^2j,$$
$$Q_{0,ij}=P_0[aB(s+i)+bA(s+j)]-2fHy_0M_{ij},\qquad Q_{1,ij}=bA(s+i)+aB(s+j).$$

The four polynomial inequalities are $\Phi_{ij}\geq 0$, where

$$\begin{aligned}\Phi_{ij}=cP_0^2F_0\{B_0^2[(1-\omega)HLS-h_0Q_{0,ij}]-2h_0NM_{ij}^{(2)}LB_0\}\\ -h_0y_0(dP_0+fHy_0)E_0L\{P_0LQ_{1,ij}+cP_0h_0fy_0(2\sigma M_{ij}+\sigma^2S)\}.\end{aligned} \tag{566}$$

Positive-denominator clearance yields the exact identity

$$\boxed{\Phi_{ij}=c^2P_0^3Sh_0B_0^2F_0[(1-\omega)k-\mathcal{R}_{ij}].} \tag{567}$$

For a direct algebraic check, multiply (560) by cP_0S after the moment substitution. Its value is

$$Q_{0,ij}+\frac{2NM_{ij}^{(2)}L}{B_0}+\rho_w[P_0Q_{1,ij}+\frac{cP_0h_0fy_0}{L}(2\sigma M_{ij}+\sigma^2S)].$$

Now (565) gives (567). In the rational case the factor q^3 in h_0 is essential.

For the four finite cases substitute the parameters from Section 11.8. The resulting coordinate degrees are $(478,8,8)$ for rank two and $(182,8,8)$ for each of the other three cases. For the higher family substitute

$$s=5+v,\qquad g=\frac{9}{8}s+\frac{803}{1000},\qquad \sigma=1,\qquad (i,j)\in\{0,1\}^2.$$

Each polynomial has degree eight in v:

$$\Phi_{ij}(w,a,b;5+v)=\sum_{j_0=0}^{8}C_{ij,j_0}(w,a,b)v^{j_0}. \tag{568}$$

All 36 coefficient polynomials are certified nonnegative on the full cube. Since $v\geq 0$, this proves the full family for *every real* $s\geq 5$. Lemma 11.14, rather than a formal substitution into an integer-only matrix argument, justifies using these real orders.

Complete exact verification. For a polynomial $R(x_1, x_2, x_3) = \sum_\nu M_\nu x^\nu$ with coordinate degrees (n_1, n_2, n_3), its tensor Bernstein coefficients on the unit cube are

$$B_\mu = \sum_{\nu \le \mu} M_\nu \prod_{j=1}^{3} \frac{\binom{\mu_j}{\nu_j}}{\binom{n_j}{\nu_j}}. \tag{569}$$

On another box first make its affine substitution. The Bernstein basis is nonnegative and sums to one; therefore nonnegative coefficients prove nonnegativity on the whole box.

A complete preorder tree word encodes each polynomial's cover. A letter `T`, `U`, or `V` bisects the first, second, or third coordinate and is followed by both subtrees; `P` is an accepted leaf. The first coordinate is w. Write a corner label ab for $(i, j) = (\chi \text{ if } a = 0 \text{ else } 1, \chi \text{ if } b = 0 \text{ else } 1)$. All trees are given by this short table:

Case	Labels	Tree
Rank two	$00, 01, 10$	`P`
	11	`TPP`
Rank three	All four	`P`
Order three	01	`TPP`
	$00, 10, 11$	`P`
Order four	$00, 01$	`TPP`
	$10, 11$	`P`
Higher family	All corners, $j_0 = 0, \dots, 7$	`P`
	All corners, $j_0 = 8$	`TTTPPPP`

The last tree covers the four intervals $[0, 1/8]$, $[1/8, 1/4]$, $[1/4, 1/2]$ and $[1/2, 1]$ in w, with the full light-variable square. Every tree is consumed exactly; its boxes have disjoint interiors and their exact volumes sum to one. Different polynomials can use different covers of the same cube.

The eight univariate positivity tests in (551) use their ordinary polynomial degree in the one-dimensional version of (569). All have strictly positive coefficients without subdivision except the following:

Case	Test	Complete tree
Rank two	$d_0^2 AB + d_0 PH - N^2$	`TPTPTPTPTPTPTPP`
Higher family	B	`TPP`
Higher family	H	`TPTPP`

These cover 50 univariate leaves in total. For illustration, the rank-two tests give

$$P \ge 21163/6250000 > 0, \qquad Q \ge 760399/10^8 > 0.$$

Use Eqs. (854) and (855) for every polynomial in Eq. (551), including the polynomial continuations of the quotients and all denominator tests. The univariate degrees are at most 50.

Proposition 11.16. *For the parameters in Section 11.8, the budget and denominator tests (551) hold. All four polynomials (566) are nonnegative on the entire cube in each finite case. All 36 higher-order coefficient polynomials are nonnegative on the cube. Consequently Eq. (558) holds at every stationary light state in the stated rank and Schatten-order ranges.*

Proof. Reconstruct the rational polynomials from (547)–(566) and apply (569) on the displayed complete trees. The exact coefficient counts are

Case	Polynomials	Polynomial–leaf tests	Coefficient tests	Zeros
Rank two	4	5	193,995	3,475
Rank three	4	4	59,292	1,812
Order three	4	5	74,115	1,995
Order four	4	6	88,938	2,178
Higher family	36	48	2,925,720	52,584
Total	52	68	3,342,060	62,044

Every coefficient test is nonnegative. Each polynomial has its own complete volume-one cover. The univariate tests are strictly positive, so all divisions in the analytic argument have positive denominators. Equations (567) and (568) therefore imply $(1-\omega)k \geq \mathcal{R}$ at the moment corners and all required Schatten orders. Affineness extends this to actual local moments. Equation (559) proves (558).

The 62,044 zero coefficients are deliberately accepted on the closed certificate cubes. They are not claimed to be positive curvature margins. Strictness comes from $\omega k w_e > 0$ at every actual state, where $w > 0$, $a, b > 0$, and the weights are positive. □

The coefficient construction in Appendix B.1 applies directly to the polynomials and trees above. Expand each displayed polynomial over $\mathbb{Q}$, substitute the leaf's affine coordinates, and compare every Bernstein coefficient with zero. The univariate conditions use their separate strict thresholds. These are finite tests on complete boxes; the retained row margin supplies strict curvature.

From curvature to fixed-order signing bounds.

Corollary 11.17. *Let $A_1, \dots, A_N \succeq 0$ satisfy $S := \sum_i A_i \preceq I_D$, $\|A_i\| \leq \alpha$, and $\operatorname{rank} A_i \leq r$. Each following estimate is available, with a signing that may depend on the selected estimate:*

$$\|\sum_i \sigma_i A_i\| \leq 2\sqrt{3r^{1/2}\alpha}, \qquad r \leq 2, \tag{570}$$

$$\|\sum_i \sigma_i A_i\| \leq 2\sqrt{\tfrac{63}{20}r^{1/2}\alpha}, \qquad r \leq 3, \tag{571}$$

$$\|\sum_i \sigma_i A_i\| \leq 2\sqrt{\tfrac{79}{20}r^{1/3}\alpha}, \qquad r \leq 32, \tag{572}$$

$$\|\sum_i \sigma_i A_i\| \leq 2\sqrt{\tfrac{103}{20}r^{1/4}\alpha}, \qquad \text{all } r, \tag{573}$$

$$\|\sum_i \sigma_i A_i\| \leq 2\sqrt{(\tfrac{9}{8}s + \tfrac{803}{1000})r^{1/s}\alpha}, \qquad s \geq 5 \text{ real, all } r. \tag{574}$$

Proof. Proposition 11.16 and Lemma 11.15 exclude every nonvertex stationary light state as a local minimum on its open face. Choose a global minimizer of the continuous potential on the cube having the fewest active coordinates. An available endpoint move gives another minimizer with fewer active coordinates, a contradiction. Otherwise a nonvertex minimizer is stationary and light with positive semidefinite active Hessian, contradicting strict negative expected curvature. Hence this minimizer is a signing.

At zero, $\phi_i(I) = \|A_i\|_{S_s} \leq r^{1/s}\alpha$ and $S \preceq I_D$. The scalar trial $U = V = \lambda I$ gives

$$P_\rho(0) \leq \min_{\lambda>0}\{\lambda^{-1} + (gr^{1/s}\alpha + 2n\rho)\lambda\} = 2\sqrt{gr^{1/s}\alpha + 2n\rho}.$$

The potential dominates the signed matrix norm. Let $\rho \downarrow 0$ and pass to a signing occurring along a subsequence, possible because there are only finitely many signings. This gives each displayed estimate. No original matrix has been split into separately signed summands. □

The improved rank-one input. For rank one we use the variance bound of Song and Yue [SY26b]. Write $A_i = a_i a_i^*$ for $i = 1, \dots, N$ and $S := \sum_{i=1}^N A_i \preceq I_D$. Since $\|A_i\| \le \alpha$,

$$\sum_{i=1}^N A_i^2 \preceq \alpha \sum_{i=1}^N A_i = \alpha S \preceq \alpha I_D.$$

Apply their result to independent symmetric signs, whose means are zero and whose variances are one. There are signs $\sigma_1, \dots, \sigma_N$ such that

$$\|\sum_{i=1}^N \sigma_i A_i\| \le 2.176 \|\sum_{i=1}^N A_i^2\|^{1/2} \le 2.176\sqrt{\alpha}.$$

Hence rank one contributes at most $2.176^2/\log 2$ to the normalized squared coefficient. This argument applies directly when $S \preceq I_D$.

Every rank and the exact uniform coefficient. First suppose r is a positive integer and normalize squared discrepancy by $\alpha \log(2r)$. For each fixed Schatten order s,

$$\frac{d}{d \log r} \log\left(\frac{r^{1/s}}{\log(2r)}\right) = \frac{1}{s} - \frac{1}{\log(2r)}.$$

The function decreases and then increases, so its maximum on a closed rank interval is at an endpoint.

Rank one. Section 11.8 gives $2.176^2/\log 2 < 6\sqrt{2}/\log 2$.

Rank two. Equation (570) gives exactly

$$\frac{12\sqrt{2}}{\log 4} = \frac{6\sqrt{2}}{\log 2} = C_*^2.$$

This is the limiting rank in the completed comparison.

Rank three. The normalized square from (571) is $(63/5)\sqrt{3}/\log 6$. The integer inequality $3^{12} > 2^{19}$ implies $\log 6 > (31/12)\log 2$, so it is at most

$$\frac{756\sqrt{3}}{155 \log 2} < \frac{6\sqrt{2}}{\log 2}, \qquad 2 \cdot 930^2 - 3 \cdot 756^2 = 15192 > 0.$$

Ranks four through sixteen. Use (572). The endpoint ratio at 16 versus 4 is $(3/5)4^{1/3} < 1$, since $108 < 125$. Thus the normalized square is at most

$$\frac{794^{1/3}}{15 \log 2} < \frac{6\sqrt{2}}{\log 2}.$$

The final comparison is the exact integer inequality $2 \cdot 79^6 < 90^6$.

Ranks seventeen through eighty. Use (573) and enclose the interval by $[16, 81]$. At 16 the normalized square is

$$\frac{206}{25\log 2} < \frac{6\sqrt{2}}{\log 2}, \qquad 2\cdot 150^2 - 206^2 = 2564 > 0.$$

At 81 it is $(309/5)/\log 162$. From $3^{12} > 2^{19}$, $\log 162 = \log 2 + 4\log 3 > (22/3)\log 2$, and hence this endpoint is less than

$$\frac{927}{110\log 2} < \frac{6\sqrt{2}}{\log 2}, \qquad 2\cdot 660^2 - 927^2 = 11871 > 0.$$

Both endpoints are below the target.

All ranks at least eighty-one. Let $L = \log(2r)$ and choose the *real* Schatten order $s = L$. Since $\log 2 > 69/100$, the preceding lower bound on $\log 162$ gives $L > 5$. Put $a = 9/8$, $b = 803/1000$, and $\ell_0 = \log 2$. Equation (574) gives normalized square

$$4(a + b/L)e^{1-\ell_0/L}. \tag{575}$$

Set $u = \ell_0/L$ and $x = a\ell_0/b + u > 0$. The elementary inequality $xe^{-x} \le e^{-1}$ then gives

$$4(a + b/L)e^{1-\ell_0/L} \le \frac{4b}{\ell_0}e^{a\ell_0/b} = \frac{803}{250\log 2}2^{1125/803} < \frac{6\sqrt{2}}{\log 2}.$$

The last strict comparison is exactly the integer-power test

$$\boxed{803^{1606}2^{1447} < 1500^{1606}.} \tag{576}$$

Indeed cancellation of positive factors reduces it to $8032^{1447/1606} < 1500$. Integer exponentiation verifies (576). The proof treats every arbitrarily large rank, not a finite rank scan.

Rank range	Upper bound on the normalized square
$r = 1$	$2.176^2/\log 2 < C_*^2$
$r = 2$	$6\sqrt{2}/\log 2 = C_*^2$
$r = 3$	$756\sqrt{3}/(155\log 2) < C_*^2$
$4 \le r \le 16$	$794^{1/3}/(15\log 2) < C_*^2$
$17 \le r \le 80$	$\max\{206/25, 927/110\}/\log 2 < C_*^2$
$r \ge 81$	$8032^{1125/803}/(250\log 2) < C_*^2$

For nonintegral r apply the result to $\lfloor r \rfloor$ and increase the logarithmic bound. Every signing satisfies $-S \preceq \sum_i \sigma_i A_i \preceq S$. Thus the cap is $\|S\|$. The resulting bound is stated explicitly below.

Corollary 11.18. *Let $A_1, \dots, A_N \succeq 0$ satisfy $S := \sum_i A_i \preceq I_D$, $\|A_i\| \le \alpha$, and $\operatorname{rank} A_i \le r$, where $r \ge 1$ and $\alpha > 0$. There are signs, one per original matrix, such that*

$$\|\sum_i \sigma_i A_i\| \le \min\{\|S\|, C_*\sqrt{\alpha\log(2r)}\}, \qquad C_* := \sqrt{\frac{6\sqrt{2}}{\log 2}} < 3.49882.$$

Proof. The fixed-order estimates in Corollary 11.17, the rank-one estimate above, and the exact comparisons for all ranks give the square-root bound. For nonintegral r, apply the estimate with $\lfloor r \rfloor$ and increase the logarithmic bound. The inequalities $-S \preceq \sum_i \sigma_i A_i \preceq S$ give the cap. □

11.9 A further rank-sensitive budget

We refine the rank-dependent existence estimate of Section 11.8. The potential, differential inequality for real Schatten orders, response geometry, and direct mixed-term covariance remain those of Lemmas 11.14 and 11.15. We use a sharper energy–moment envelope at rank three and a new endpoint budget for every real Schatten order at least five. The exact parameters, polynomial definitions, and complete certificate trees are specified in this section.

The rank-three energy–moment envelope. For a normalized compressed matrix R and a trace-one dual density S, retain $\kappa = \operatorname{tr}[R^{s+1}]$ and $e_R = \operatorname{tr}[RS]$ as in Eq. (533). At order two and rank at most three we also have the following bound.

Lemma 11.19. *If $R \succeq 0$, $\operatorname{rank} R \le 3$, $\operatorname{tr}[R^2] = 1$ and $S \succeq 0$ has trace one, then, with $\kappa = \operatorname{tr}[R^3]$ and $e_R = \operatorname{tr}[RS]$,*

$$e_R \le \min\{\frac{2+\kappa}{3},\ \frac{6464}{24219} + \frac{243}{299}\kappa\}. \tag{577}$$

The upper boundary of this enclosing region over $0 \le \kappa \le 1$ has vertices

$$(\kappa, e_R) = (0, \frac{6464}{24219}), \quad (\frac{4841}{5805}, \frac{16451}{17415}), \quad (1,1). \tag{578}$$

Proof. Let x be the largest eigenvalue. Padding with zero eigenvalues when necessary, $x \in [1/\sqrt{3}, 1]$. The other two eigenvalues have squared sum $1 - x^2$. Power means imply

$$\kappa \ge h(x) := x^3 + \frac{1}{\sqrt{2}}(1-x^2)^{3/2}.$$

Writing $c_0 = 1/\sqrt{2}$, its derivatives on the interior are

$$h'(x) = 3x\big(x - c_0\sqrt{1-x^2}\big), \quad h''(x) = 6x - 3c_0\sqrt{1-x^2} + \frac{3c_0x^2}{\sqrt{1-x^2}} > 0.$$

For positivity use $c_0\sqrt{1-x^2} \le x$ when $x \ge 1/\sqrt{3}$. The tangent at $x_0 = 23/27$, whose other two equal eigenvalues are $10/27$, gives

$$\kappa \ge \frac{14167}{19683} + \frac{299}{243}(x - \frac{23}{27}).$$

Solving for x and using $e_R \le x$ proves the second line in (577); the first is (533). Their intersection is (578). Continuity includes $x = 1$. □

The rank-two ingredient is the analogous enclosing polygon

$$(0, \frac{1171}{3276}), \quad (\frac{1013}{1105}, \frac{3223}{3315}), \quad (1,1). \tag{579}$$

For reference its second supporting line is $1171/3276 + (169/252)\kappa$. It follows from the tangent to $x^3 + (1-x^2)^{3/2}$ at the spectrum $(12/13, 5/13)$. Enlarging the moment interval to $[0,1]$ is harmless. With positive energy coefficients an affine cost is maximized on the upper vertices, so nine pairs suffice for each of these two low-rank cases.

The covariance test. We use the notation $c = s + 1$, $d = s - 1$, $m = d/c$, k, ℓ, γ, $\delta = 1/\gamma$, and $\mathcal{E}$ of Eqs. (531)–(539). The pointwise row helper is $\zeta = 2\sigma$ and its weights are those of Eqs. (537)–(538). Whenever $Z_e, Z_f > 0$ and $\mathcal{E} < 1$ at every active light coordinate, Lemma 11.15 gives a physical direction of strictly negative curvature. The following calculations verify these conditions for the new choices, without changing that covariance argument.

Parameters. The following table separates the two new parameter sets from the inherited finite cases. All terminating decimals denote exact rationals.

Case	Order	Coupling g	Row ζ	Spectral condition
Rank two	2	149/50	427/100	Rank ≤ 2, (579)
Rank three (new)	2	157/50	197/50	Rank ≤ 3, (578)
Order three	3	79/20	12/5	$\kappa \in [5/16, 1]$, rank ≤ 32
Order four	4	103/20	5/2	$\kappa \in [0, 1]$, all ranks
Higher (new)	Real $s \ge 5$	$559s/500 + 4/5$	23211/10000	$\kappa \in [0, 1]$, all ranks

The rank-two budget is $\beta(y) = (1 - |y|)^{5/8} F_{\mathrm{low}}(|y|)$, where

$$F_{\mathrm{low}}(t) = 1 + \tfrac{5}{8}t - .082557t^2 + .088770t^3 - .107459t^4 + .164031t^5 - .113032t^6.$$

For the next three cases use $\beta(y) = \sqrt{1 - y^2} f(|y|)$ with, respectively,

$$\begin{aligned} f_{\mathrm{rank3}}(t) &= 1 - .0523t^2 - .0706t^4, \\ f_{\mathrm{order3}}(t) &= 1 - \tfrac{1}{25}t^2 - \tfrac{9}{100}t^4, \\ f_{\mathrm{order4}}(t) &= 1 - .01484449t^2 - .08000936t^4. \end{aligned}$$

The order-three moment cap is valid since $32(5/16)^3 = 125/128 < 1$.

For the new higher family use

$$\beta(y) = (1 - |y|)^{3/4} F(|y|), \qquad F(t) = 1 + \tfrac{3}{4}t + t^2 \sum_{j=0}^{15} c_j T_j(2t - 1), \tag{580}$$

where $T_0(x) = 1$, $T_1(x) = x$, $T_{j+1}(x) = 2xT_j(x) - T_{j-1}(x)$ and

j	c_j	j	c_j
0	0.326876138335	8	−0.001422358931
1	0.175082133981	9	−0.000823090271
2	0.039913633160	10	−0.000625660263
3	0.012725225728	11	−0.000361093856
4	0.001082794443	12	−0.000182692960
5	−0.001259824602	13	−0.000125859593
6	−0.002039031180	14	−0.000028732893
7	−0.001568317511	15	−0.000024913049

Thus F has degree 17. The displayed recurrence for T_j determines its rational monomial coefficients. Since $F(0) = 1$ and $F'(0) = p$ in either endpoint-power case, its even budget extension is C^2 with locally Lipschitz second derivative. A possible third-derivative jump at zero is not used by the existence proof.

Budget positivity and the light domain. We give the two budget representations separately, then a common polynomial comparison.

For $\beta(t) = (1-t)^p F(t)$, put $x = 1-t$, $z = 1-t^2$ and

$$\begin{aligned} N_1 &= xF' - pF, & H_1 &= p(1-p)F + 2pxF' - x^2F'', \\ N &= (1+t)N_1, & H &= (1+t)^2 H_1, \\ A &= (1-t)F - N, & B &= (1+t)F + N. \end{aligned} \tag{581}$$

Then $\beta' = \beta N/(zF)$ and $-\beta'' = \beta H/(z^2F)$. For a square-root budget $\beta(t) = \sqrt{z}f(t)$, instead put

$$N = zf' - tf, \quad H = f + 2tzf' - z^2 f'', \quad A = f - zf', \quad B = f + zf'. \tag{582}$$

Its derivatives are $\beta' = N/\sqrt{z}$ and $-\beta'' = H/z^{3/2}$.

Exact univariate Bernstein checks prove, in an endpoint-power case,

$$F,\ -N_1/t,\ H_1,\ A,\ B,\ d_0^2 AB + d_0 FH - N^2 > 0,$$

where $d_0 = 1$ for the low case and $d_0 = 4$ for the higher family. For $p = j/q$, substitute $t = 1 - w^q$. The polynomial

$$F(1-w^q) - w^{q-j}(2-w^q)$$

is divisible by $(1-w)^2$, and its quotient has positive Bernstein coefficients. This proves $\beta \geq z$. For the square-root cases the tests are

$$f,\ -N/t,\ (f^2-z)/t^2,\ A,\ B,\ H,\ d^2AB + dfH - N^2 > 0.$$

Divided polynomials use continuous extensions at zero. All five budgets therefore satisfy

$$\beta(0) = 1, \quad \beta(\pm 1) = 0, \quad 1 - y^2 \leq \beta(y) \leq 1, \quad -\beta'' > 0 \quad (|y| < 1). \tag{583}$$

Use each univariate polynomial's ordinary degree in the Bernstein test. In the new higher family, bisect $[0,1]$ once for $-N_1/t, H_1, A, B$ and the denominator test; use the unsplit interval for F and the budget-lower-bound quotient. All univariate tests in the other four cases use the unsplit interval. These are 40 leaves in total. Every Bernstein coefficient on these leaves is strictly positive, as determined by Eqs. (854)–(855).

Write $f = F$ in an endpoint-power case and keep f from (582) otherwise. The same full-domain change of variables works:

$$U_0 = g\beta u = \frac{zfa}{A + Na}, \qquad V_0 = g\beta v = \frac{zfb}{B - Nb}, \qquad 0 < a, b < 1. \tag{584}$$

The maps are increasing bijections onto $(0, 1+t)$ and $(0, 1-t)$. Their denominators are positive because $A + N = (1-t)f$ and $B - N = (1+t)f$. Set

$$P_0 = AB, \quad y_0 = ab, \quad S = aB + bA, \quad D_L = (A + Na)(B - Nb).$$

For both representations,

$$\alpha_{\mathrm{loc}} = \frac{aB}{S}, \quad pq = \frac{P_0}{D_L}, \quad k\delta = \frac{fHy_0}{cP_0}, \quad J := \frac{\ell^2}{k\gamma} = \frac{N^2}{cfH}, \quad 2\ell T_{fe} = \frac{2N}{cP_0 S}(a^2B^2\kappa_U - b^2A^2\kappa_V). \tag{585}$$

The part of Z_f/γ without its nonnegative row terms is exactly

$$m - \frac{(k\delta)J}{m + k\delta} = \frac{d^2P_0 + y_0(dfH - N^2)}{c(dP_0 + fHy_0)}. \tag{586}$$

Its numerator is affine in $y_0 \in [0,1]$. Both endpoint values are positive by the univariate tests. For the higher family the second endpoint increases with $d \geq 4$, since $P_0, fH > 0$. Thus $Z_f > 0$, and $Z_e > 0$ follows from $k > 0$. This does not assume the stronger inequality $dfH - N^2 \geq 0$.

The energy coefficients in (539) are positive. Consequently the low cases reduce to the nine pairs of their energy–moment vertices. The other cases reduce to the four corners $(\kappa_U, \kappa_V) \in \{\chi, 1\}^2$, with $e = (s + \kappa)/(s + 1)$ and $\chi = 5/16$ or zero.

The exact polynomial comparison. Fix the exact relative margin $\omega = 10^{-6}$ in every case. Writing $\rho_w = w_f/w_e$, the cost satisfies

$$\begin{aligned} 1 - \mathcal{E} &= w_e(k - \mathcal{R}), \\ \mathcal{R} &= (D_0)_{ee} + \rho_w(D_0)_{ff} + 2kT_{ee} + \frac{2\ell T_{fe}}{1 + \sigma\delta} + 2\sigma\rho_w T_{ff} + \sigma^2\delta\rho_w. \end{aligned} \tag{587}$$

It is enough to prove $(1 - \omega)k - \mathcal{R} \geq 0$ at every spectral vertex pair.

For an endpoint power $p = j/q$, use $t = 1 - w^q$ and replace f, N, H, A, B by their polynomial compositions. Set

$$h_0 = w^{2q-j}(2 - w^q)^2.$$

For a quartic square-root budget, use

$$t = \frac{1 - w^2}{1 + w^2}, \qquad \sqrt{1 - t^2} = \frac{2w}{1 + w^2}, \qquad Z = 1 + w^2.$$

Replace f, N, H, A, B by the polynomial lifts $Z^4 f, Z^5 N, Z^6 H, Z^5 A, Z^5 B$, evaluated at this t, and set $h_0 = 8w^3 Z^3$. In the following formulas these symbols mean the lifted/composed polynomials. In either case the closed certificate cube is $(w, a, b) \in [0,1]^3$.

Define

$$\begin{aligned} P_0 = AB, \quad S = aB + bA, \quad y_0 = ab, \\ D_L = (A + Na)(B - Nb), \quad L = gD_L, \quad B_0 = L + \sigma h_0 f y_0, \\ E_0 = (cP_0 + fHy_0)fHB_0^2 + N^2P_0L^2, \\ F_0 = (dP_0 + fHy_0)(dL^2 + 2\sigma c h_0 f y_0 L + c\sigma^2 h_0^2 f^2 y_0^2) - y_0 N^2 L^2. \end{aligned} \tag{588}$$

Every denominator factor is positive at actual states. In particular, the terms of F_0 without σ equal $L^2[d^2P_0 + y_0(dfH - N^2)] > 0$, and its other terms are nonnegative. Substitution gives

$$k = \frac{HL}{ch_0P_0}, \quad \delta = \frac{h_0 f y_0}{L}, \quad \rho_w = \frac{y_0(dP_0 + fHy_0)E_0L^2}{cP_0^2B_0^2F_0}. \tag{589}$$

For a vertex pair (κ_U, e_U), (κ_V, e_V) put

$$\begin{aligned} M &= aB\kappa_U + bA\kappa_V, & M_2 &= a^2B^2\kappa_U - b^2A^2\kappa_V, \\ Q_0 &= P_0(aBce_U + bAce_V) - 2fHy_0M, & Q_1 &= bAce_U + aBce_V. \end{aligned}$$

The polynomial to certify is

$$\Phi = cP_0^2F_0[B_0^2((1-\omega)HLS - h_0Q_0) - 2h_0NM_2LB_0] \\ - h_0y_0(dP_0 + fHy_0)E_0L[P_0LQ_1 + cP_0h_0fy_0(2\sigma M + \sigma^2 S)]. \tag{590}$$

Direct positive-denominator clearance gives the exact identity

$$\Phi = c^2P_0^3Sh_0B_0^2F_0[(1-\omega)k - \mathcal{R}]. \tag{591}$$

It can also be checked by multiplying (587)'s expression for $\mathcal{R}$ by cP_0S: the result is

$$Q_0 + \frac{2NM_2L}{B_0} + \rho_w[P_0Q_1 + \frac{cP_0h_0fy_0}{L}(2\sigma M + \sigma^2 S)].$$

For the higher family substitute $s = 5 + v$ and $g = 559s/500 + 4/5$, with $v \geq 0$. At its two spectral endpoints $ce = s + \kappa$ is polynomial in s. Each of the four polynomials has degree eight in v:

$$\Phi_{ij}(w, a, b; 5 + v) = \sum_{j_0=0}^{8} C_{ij,j_0}(w, a, b)v^{j_0}. \tag{592}$$

All 36 coefficient polynomials are nonnegative on the full cube. This proves every real order $s \geq 5$, not a finite order scan.

Complete coefficient checks and covers. For a polynomial $P(x) = \sum_\nu M_\nu x^\nu$ of coordinate degrees n_1, n_2, n_3, its tensor Bernstein coefficients on the unit cube are

$$B_\mu = \sum_{\nu \leq \mu} M_\nu \prod_{j=1}^{3} \frac{\binom{\mu_j}{\nu_j}}{\binom{n_j}{\nu_j}}. \tag{593}$$

On a subbox first make its affine substitution. Nonnegative coefficients imply nonnegativity on the whole box because the basis functions are nonnegative and sum to one.

The complete covers are extremely short. A letter T bisects w; P is an accepted leaf. In a finite case labels ij index the vertices in the displayed order. In the higher family ij_j_0 indexes the two moment corners and the coefficient of v^{j_0}.

Case	Labels	Complete tree
Low rank two	All nine	P
New rank three	All nine	P
Order three	01	TPP
	$00, 10, 11$	P
Order four	$00, 01$	TPP
	$10, 11$	P
Higher	$00\_8, 01\_8, 10\_0, 10\_8, 11\_0, 11\_8$	TPP
	All other labels	P

Every polynomial therefore has a complete cube cover, of exact volume one. The maximum w depth is one; neither light variable is subdivided. The coordinate degrees are $(1016, 8, 8)$ for low rank two and $(182, 8, 8)$ for each other finite case. In the higher family the degrees are $(1300, 8, 8)$ for $j_0 \leq 5$, $(1300, 7, 7)$ for $j_0 = 6, 7$, and $(1300, 6, 6)$ for $j_0 = 8$.

Case	Polynomials	Polynomial–leaf tests	Coefficients	Zeros
Low rank two	9	9	741,393	18,873
New rank three	9	9	133,407	4,077
Order three	4	5	74,115	1,995
Order four	4	6	88,938	2,178
New higher family	36	42	3,916,010	73,122
Total	62	71	4,953,863	100,245

All coefficient tests are nonnegative. The 100,245 zeros occur on the compactifying closed domains; they are not asserted to be strict margins. At actual states $h_0, a, b > 0$ and all denominators are positive. Equations (591) and (592) give

$$\mathcal{E} \le 1 - \omega k w_e < 1. \tag{594}$$

Thus Lemma 11.15 supplies strict negative curvature in every stated case.

Expand the polynomials from the displayed rational parameters and apply Eqs. (854) and (855) at every leaf of the complete trees. Use the coordinate degrees listed above and compare each coefficient with zero. This gives the stated counts and nonnegativity tests. The denominator-clearance identities follow by rational polynomial expansion, and the positive row margin below supplies strictness.

Proposition 11.20. *For each case above, the budget and covariance denominators are positive at every actual light state. The local cost obeys $\mathcal{E} \le 1 - 10^{-6} k w_e < 1$ there.*

Proof. The univariate Bernstein tests give the budget and denominator claims. Every listed complete tree covers its cube, and every tested Bernstein coefficient is nonnegative. Equation (591) therefore proves the comparison at each finite spectral vertex pair. For the higher family, the nonnegative coefficients of $v = s - 5$ in Eq. (592) prove the comparison for every real $s \ge 5$. The explicit positive relative margin gives strictness at actual light states even though coefficients may vanish on the closed certificate cubes. □

Corollary 11.21. *If $A_1, \dots, A_N \succeq 0$ satisfy $S := \sum_i A_i \preceq I_D$, $\|A_i\| \le \alpha$, and $\operatorname{rank} A_i \le r$, then a signing exists for each of the following bounds, possibly with different signs:*

$$\|\sum_i \sigma_i A_i\| \le 2\sqrt{\frac{149}{50} r^{1/2} \alpha} \qquad (r \le 2), \tag{595}$$

$$\|\sum_i \sigma_i A_i\| \le 2\sqrt{\frac{157}{50} r^{1/2} \alpha} \qquad (r \le 3), \tag{596}$$

$$\|\sum_i \sigma_i A_i\| \le 2\sqrt{\frac{79}{20} r^{1/3} \alpha} \qquad (r \le 32), \tag{597}$$

$$\|\sum_i \sigma_i A_i\| \le 2\sqrt{\frac{103}{20} r^{1/4} \alpha} \qquad (\textit{all } r), \tag{598}$$

$$\|\sum_i \sigma_i A_i\| \le 2\sqrt{(\frac{559}{500} s + \frac{4}{5}) r^{1/s} \alpha} \qquad (s \ge 5 \textit{ real, all } r). \tag{599}$$

Proof. Proposition 11.20 and Lemma 11.15 exclude every nonvertex stationary light state as a local minimum. The compact-minimum argument of Corollary 11.17 therefore gives a signing, one sign

per original matrix. At zero, $\phi_i(I) \le r^{1/s}\alpha$ and $S \preceq I_D$, so the scalar trial $U = V = \lambda I$ gives $P_\rho(0) \le 2\sqrt{gr^{1/s}\alpha + 2D\rho}$. Send ρ to zero along a recurrent signing. The rank-two case uses the first budget in the certificate; the other rows use the displayed parameter table. □

Corollary 11.22. *Under the hypotheses of Corollary 11.21, there are signs, one per original matrix, such that*

$$\|\sum_i \sigma_i A_i\| \le \min\{\|S\|, C_{\mathrm{new}}\sqrt{\alpha \log(2r)}\}, \qquad C_{\mathrm{new}} := \sqrt{\frac{16}{5\log 2} 2^{559/400}} < 3.48745.$$

Proof. For rank one, the variance bound of Song and Yue [SY26b], applied in Section 11.8, gives discrepancy at most $2.176\sqrt{\alpha}$ when $S \preceq I_D$ and contributes at most $2.176^2/\log 2$ to the normalized squared coefficient. For fixed s, the function $r^{1/s}/\log(2r)$ attains its maximum on a closed rank interval at an endpoint. The remaining finite comparisons, using the rows of Corollary 11.21, are

Rank range	Upper bound on normalized square	Rational upper bound
$r = 2$	$149\sqrt{2}/(25\log 2)$	12.16006221
$r = 3$	$314\sqrt{3}/(25\log 6)$	12.14145008
$4 \le r \le 16$	$79 \cdot 4^{1/3}/(15\log 2)$	12.06138097
$17 \le r \le 80$	endpoints $16, 81$ at order four	12.14719014

The rational logarithm, root, and exponential enclosures in Appendix B.7 verify every entry and place it strictly below C_{new}^2. For $r \ge 81$ put $L = \log(2r) > 5$ and choose the real order $s = L$. With $a = 559/500$, $b = 4/5$ and $\lambda = \log 2$, the normalized square is $4(a + b/L)e^{1-\lambda/L}$. Setting $x = a\lambda/b + \lambda/L$ and using $xe^{-x} \le e^{-1}$ gives

$$4(a + b/L)e^{1-\lambda/L} \le \frac{4b}{\lambda} e^{a\lambda/b} = C_{\mathrm{new}}^2.$$

The exact enclosure is $12.1622502785 < C_{\mathrm{new}}^2 < 12.1622502786$ and $C_{\mathrm{new}} < 3.48745$. Replace nonintegral r by its integer part. Finally $-S \preceq \sum_i \sigma_i A_i \preceq S$ supplies the cap by $\|S\|$. □

11.10 The rank-dependent signing theorem

The Schatten orders and rank comparisons in Section 11.9 gives logarithmic dependence on rank.

Theorem 11.23 (Formal version of Theorem 1.4)**.** *Let $A_1, \dots, A_N \in \mathbb{C}^{D\times D}$ be Hermitian positive semidefinite matrices with $\sum_{i=1}^N A_i = I_D$, $\|A_i\| \le \alpha$, and $\mathrm{rank}(A_i) \le r$, where $r \ge 1$ and $\alpha > 0$. There are signs $\sigma_1, \dots, \sigma_N \in \{-1, 1\}$ satisfying*

$$\|\sum_{i=1}^N \sigma_i A_i\| \le \min\{1, 3.48745\sqrt{\alpha \log(2r)}\}. \tag{600}$$

Proof. Discard zero inputs and replace a nonintegral rank bound by its integer part. Corollary 11.22 gives the square-root bound in Eq. (600). Every signing of an isotropic positive semidefinite family satisfies $-I_D \preceq \sum_i \sigma_i A_i \preceq I_D$. This gives the cap by one in Eq. (600). Assign arbitrary signs to the discarded zero inputs. □

11.11 Baseline construction and computational model

We count real arithmetic operations and comparisons, representing each complex entry by two real coordinates. Scalar functions, optimization, and spectral computations are evaluated to the finite accuracies specified below. The operation bound is independent of the smallest positive eigenvalue of an individual input.

We may replace the supplied rank bound by $\min(r, D)$. Define $k := \lceil \log_2 r \rceil$, computed by doubling and comparisons, and choose

$$p := 2k + 2, \qquad \iota := 16(p + 3), \qquad \delta_{\mathrm{cut}} := d := \tau := \frac{1}{10^6 N^2}, \qquad \rho := \frac{1}{10^6 D N^2}. \tag{601}$$

Isotropy guarantees $\alpha \geq 1/N$ and $\|A_i\| \leq 1$. The accuracies in Eq. (601) therefore depend only on N, D, r.

Discard inputs with $\mathrm{tr}[A_i] \leq \delta_{\mathrm{cut}}$, assign them arbitrary signs, and restore them at the end. Their total norm contribution is at most $N\delta_{\mathrm{cut}}$. Every retained input satisfies

$$\mathrm{tr}[A_i] > \delta_{\mathrm{cut}}, \qquad \|A_i\| > \delta_{\mathrm{cut}}/D. \tag{602}$$

The retained family satisfies $\sum_{i \in I} A_i \preceq I_D$.

Potential and uniform bounds. Use the potential in Eq. (343), with Schatten order p:

$$f_i(X) := (\mathrm{tr}[(A_i^{1/2} X A_i^{1/2})^p])^{1/p}, \qquad \beta(y) := \sqrt{1 - y^2},$$

$$P_\rho(y) := \min_{b, U, V \succ 0} b + \rho\, \mathrm{tr}[U + V]$$

subject to

$$U^{-1} + M(y) + \iota \sum_{i \in I} \beta(y_i) f_i(V) A_i \preceq bI,$$

$$V^{-1} - M(y) + \iota \sum_{i \in I} \beta(y_i) f_i(U) A_i \preceq bI.$$

Here I is the active set, and M includes the signed matrices already fixed. Start with every retained coordinate zero. Always $\|M(y)\| \leq 1$.

Set

$$B := 16(p + 3) + 2, \qquad a := \frac{1}{B + 1}, \qquad R := \frac{2}{\rho}. \tag{603}$$

On every face and for every y in its cube, the value is nonnegative and is less than B. To see the upper bound, take $U := V := \ell I$ with $\ell := 1/[8(p + 3)]$. Since $r^{1/p} \leq \sqrt{2}$, the inverse terms are $8(p + 3)I$, and each coupling term is at most $2\sqrt{2}I$. Thus $b := 8(p + 3) + 5$ is strictly feasible with matrix slack greater than one, and its objective is less than B.

At an optimum, the primal and dual equations imply

$$U, V \succeq aI, \qquad \rho\, \mathrm{tr}[U^2 + V^2] \leq 1, \qquad W, Z \succeq \rho a^2 I, \qquad \mathrm{tr}[W + Z] = 1. \tag{604}$$

In particular $U, V \preceq \rho^{-1/2} I \prec RI$. The upper bounds follow from the dual equations for the nonsymmetric derivative coupling.

11.12 Response conditioning and derivatives of the Schatten-norm functional

Let J be the fixed-b response operator from Section 8.5. Its inverse is positive and

$$J^*(W,Z) = \rho(I,I).$$

The strictly positive dual subeigenvector establishes positivity and invertibility. For a Hermitian pair Y, each component lies between $-\|Y\|_F I$ and $\|Y\|_F I$. Positivity therefore gives

$$-\frac{\|Y\|_F}{\rho}(W,Z) \preceq (J^*)^{-1}Y \preceq \frac{\|Y\|_F}{\rho}(W,Z).$$

If $-Q \preceq H \preceq Q$ and $Q \succeq 0$, then $\|H\|_F \le \|H\|_* \le \operatorname{tr}[Q]$, as follows by pairing the two inequalities with the positive and negative spectral projections of H. Applied to the pair and using $\operatorname{tr}[W+Z] = 1$, this proves

$$\|J^{-1}\|_{F\to F} = \|(J^*)^{-1}\|_{F\to F} \le 1/\rho. \tag{605}$$

The full KKT system is also polynomially conditioned. Its homogeneous-variable blocks are

$$-J\delta X - e\delta b = r_0, \qquad \mathcal{H}\delta X - J^*\delta Q = s_0, \qquad \langle e, \delta Q\rangle = t_0, \quad e := (I,I).$$

Here $\mathcal{H}$ is the Lagrangian Hessian in $X := (U,V)$. It satisfies

$$\mathcal{H} \succeq (2\rho/R)I. \tag{606}$$

For the inverse term in direction H, we have

$$2\operatorname{tr}[WU^{-1}HU^{-1}HU^{-1}] \ge 2\rho\operatorname{tr}[U^{-1}H^2] \ge (2\rho/R)\|H\|_F^2.$$

The first step uses $U^{-1}WU^{-1} \succeq \rho I$ from the dual equation. The second step uses $U \preceq RI$. The nonlinear functional Hessians are positive semidefinite. With $v := J^{-1}e$, eliminating $\delta X, \delta Q$ leaves the scalar denominator

$$\langle v, \mathcal{H}v\rangle \ge (2\rho/R)\frac{\|e\|_F^2}{\|J\|^2}. \tag{607}$$

All other terms use J^{-1}, $\mathcal{H}$, and v. Polynomial bounds on the entries and derivatives, proved next, therefore give a polynomial bound on the full inverse.

Derivatives of the Schatten-norm functional. For any $X \succeq aI$, put $C := A_i^{1/2}XA_i^{1/2}$ on its positive support. For a possibly complex perturbation H, define E by

$$A_i^{1/2}HA_i^{1/2} = C^{1/2}EC^{1/2}, \qquad \|E\| \le \|H\|/a.$$

The inequality follows from $A_i^{1/2}C^{-1}A_i^{1/2} \preceq X^{-1}$. Schatten Hölder gives

$$\|C^{1/2}EC^{1/2}\|_{S_p} \le \|E\|\|C\|_{S_p}.$$

Expanding the noncommutative product and bounding each trace by Hölder yields

$$\left|\frac{\operatorname{tr}[(C + C^{1/2}EC^{1/2})^p]}{\operatorname{tr}[C^p]} - 1\right| \le (1+\|E\|)^p - 1. \tag{608}$$

If $\|H\| \le a/(4p)$, the right side is less than $e^{1/4} - 1 < 1/2$. The positive p-th root consequently has an analytic continuation on that ball with absolute value at most $2f_i(X)$. Cauchy estimates and polarization give, for each fixed $j \le 4$,

$$\|D^j f_i(X)\| \le 2j!(\frac{4pj}{a})^j f_i(X) \le 4j!(\frac{4pj}{a})^j R\|A_i\|. \tag{609}$$

The derivative norm is the multilinear norm with Frobenius-norm directions. The derivative bound depends polynomially on p and is independent of the smallest positive eigenvalue of A_i.

On the truncated cube $|y_i| \le 1 - d/2$, the first three derivatives of β have polynomial bounds in $1/d$. Derivatives of inverse matrices have the usual bounds $j!\|X^{-1}\|^{j+1}$. Thus the KKT residual map $F(z, y)$, $z := (b, U, V, W, Z)$, and its derivatives through order three admit uniform polynomial bounds. The fourth derivative of f_i in Eq. (609) is included because stationarity already differentiates f_i once.

A common bound for the first three derivatives, in small neighborhoods of the optima, is

$$F_* := 10^{20} N(D+1)\iota(p+1)^4(R+1)(1+a^{-1})^5(1+d^{-1})^4. \tag{610}$$

To obtain Eq. (610), apply the product rule and the derivative bound for f_i in Eq. (609) on the neighborhood

$$U, V \succeq (a/2)I, \qquad U, V \preceq 2RI, \qquad \|W\|_F, \|Z\|_F \le 2.$$

The constant absorbs the fixed-order product and polarization coefficients. In particular, F_* bounds $\|J\|$, $\|\mathcal{H}\|$, and the joint derivative norms of F in (z, y).

To bound the derivatives of the value function, define

$$L := 1/\rho, \quad E_0 := 2D, \quad \mu := \rho/R,$$

$$K_b := \frac{F_*^2}{\mu}(1+LE_0+L^2E_0F_*), \quad K_X := L(1+E_0K_b), \quad K_Q := L(F_*K_X+1), \quad K_* := K_b+K_X+K_Q.$$

Eqs. (605)–(607) give $\|F_z^{-1}\| \le K_*$. Differentiating $F(z(y), y) = 0$ gives valid bounds

$$Z_1 := K_*F_*, \quad Z_2 := K_*F_*(1+Z_1)^2,$$

$$Z_3 := K_*F_*[(1+Z_1)^3 + 3(1+Z_1)Z_2].$$

Because the objective is linear in z, we can take

$$L_2 := (1+2D)Z_2 \ge \|\nabla^2 P_\rho\|, \qquad L_3 := (1+2D)Z_3 \ge \|D^3 P_\rho\|. \tag{611}$$

These bounds are explicit polynomials in N, D, p.

11.13 Quantitative gradient-or-curvature alternative

Assume all active coordinates satisfy $|y_i| \le 1 - d$ and both endpoint scores are negative, so the state is light. Use the scalar notation from Section 8.5:

$$u_i := f_i(U), \ v_i := f_i(V), \ w_i := \mathrm{tr}[A_iW], \ z_i := \mathrm{tr}[A_iZ],$$

$$p_i := 1 + \iota\beta_i' v_i, \quad q_i := 1 - \iota\beta_i' u_i, \quad g_i := p_iw_i - q_iz_i, \quad S_i := v_iw_i + u_iz_i.$$

The symbol p_i is a coordinate coefficient, distinct from the Schatten order p. Lightness implies

$$p_iq_i \ge 1/2, \quad p_i, q_i \le 1/d, \quad u_i, v_i \le 2/(\iota\sqrt{d}).$$

From Eqs. (602) and (604),

$$S_i \geq \frac{2\rho a^3 \delta_{\text{cut}}^2}{D}, \qquad p_i q_i S_i \geq \nu := \frac{\rho a^3 \delta_{\text{cut}}^2}{D}, \qquad |p_i u_i - q_i v_i| \leq R_0 := \frac{4}{\iota d^2}. \tag{612}$$

The covariance estimate in Section 8.5 also applies away from stationarity. Normalize its covariance $K \succeq 0$ so $\operatorname{tr}[K] = 1$. The exact identity

$$p_i^2 u_i w_i + q_i^2 v_i z_i = p_i q_i S_i + (p_i u_i - q_i v_i) g_i \tag{613}$$

accounts for the stationarity error. Insert this identity into the functional Hessian estimate and apply the normalized covariance bound. We have

$$\begin{aligned} \operatorname{tr}[\nabla^2 P_\rho K] &\leq -2(p+3) \sum_{i \in I} p_i q_i S_i K_{ii} + 2(p+3) \sum_{i \in I} (p_i u_i - q_i v_i) g_i K_{ii} \\ &\leq -2(p+3)\nu + 2(p+3) R_0 \|g\|_\infty. \end{aligned} \tag{614}$$

The second step uses Eq. (612) and the normalization $\sum_i K_{ii} = 1$.

Define

$$G := \frac{\nu}{2R_0}, \qquad \chi := (p+3)\nu. \tag{615}$$

Every truncated light state therefore satisfies

$$\|\nabla P_\rho\|_\infty > G \quad \text{or} \quad \lambda_{\min}(\nabla^2 P_\rho) \leq -\chi. \tag{616}$$

Both G and χ are inverse polynomials in N, D, p. The covariance establishes the alternative. The algorithm evaluates the gradient and Hessian through the KKT system.

Finite-accuracy values and derivatives. For each fixed y, the program defining P_ρ is convex in $O(D^2)$ real variables. Restrict the search to

$$(a/2)I \preceq U, V \preceq RI, \qquad -1 \leq b \leq B + 1.$$

All optimizers lie in its interior. The explicit strictly feasible point in Section 11.11 has inverse-polynomial distance to the box boundary and matrix slack greater than one. Interpolating it with an optimum gives, in an objective sublevel within δ of the optimum, a ball of radius inverse-polynomial in $N, D, p, 1/\delta$. Thus the logarithm of the outer-to-inner radius ratio is polynomial.

A separation oracle tests the two matrix constraints. A violating unit vector v produces the scalar convex constraint

$$v^* U^{-1} v + v^* M v + \iota \sum_{i \in I} \beta_i f_i(V)(v^* A_i v) - b \leq 0,$$

or its counterpart, and its gradient gives a separating hyperplane. An approximate spectral test gives weak separation; an uncertainty band is accepted as a small matrix-constraint error, repaired by increasing b. Values and gradients of the scalar functional are computable by matrix products and approximations of scalar roots. The functional and its gradient admit the matrix-product formulas

$$f_i(X)^p = \operatorname{tr}[(A_i X)^p], \qquad B_i(X) = \frac{(A_i X)^{p-1} A_i}{f_i(X)^{p-1}}. \tag{617}$$

For retained inputs in the search box, $f_i(X) \geq a\delta_{\mathrm{cut}}/(2D)$. Matrix powers require polynomially many arithmetic operations since $p = O(\log(2D))$. Scalar roots can be approximated by bisection using polynomial evaluations. The logarithms of the intermediate power magnitudes have polynomial bounds, which suffice for the finite-accuracy arithmetic count. Approximate spectral decompositions can be obtained by gap-independent dense routines such as Jacobi rotations, with scalar square roots likewise approximated.

Use the separation/ellipsoid method with objective bisection to evaluate the convex minimum. The standard primitive is described in Goemans's ellipsoid-method notes [Goe09]. Here the explicit bounded domain, slack, derivatives of f_i, and error repair supply the problem-specific oracle and radius assumptions.

To recover the optimizer accurately, fix its exact dual pair W_*, Z_*. On the bounded search domain the corresponding Lagrangian is strongly convex in (U, V), with modulus at least

$$\mu_{\mathrm{rec}} := 2\rho a^2/R^3.$$

Hence any feasible point whose objective is within δ of the optimum has $\|(U,V) - (U_*, V_*)\|_F \leq \sqrt{2\delta/\mu_{\mathrm{rec}}}$. Its b is close as well, by subtracting the linear trace term from the objective. Recover the dual multipliers from $J^*(W, Z) = \rho(I, I)$, then compute derivatives by differentiating the KKT equations and solving dense linear systems. The inverse estimates in Section 11.12 make every required accuracy attainable with polynomial work. Hessian entries can be recovered by solving for first and second parameter derivatives; there are polynomially many such solves.

Choose the primal solve and arithmetic tolerances sufficiently small, using the displayed conditioning bounds, to certify the value, gradient, Hessian, and endpoint-score error thresholds in Section 11.14. Weak feasibility errors are repaired before applying the recovery bound. The separation and spectral computations use the specified error margins.

11.14 Algorithm and outer-step count

Compute L_2, L_3, G, χ from Eqs. (611) and (615), and define

$$t_g := \min\{d/2, G/(4L_2)\}, \tag{618}$$
$$t_c := \min\{d/2, \chi/(4L_3)\}. \tag{619}$$

$$\Delta := \min\{Gt_g/4, \chi t_c^2/8\}. \tag{620}$$

The step lengths t_g, t_c and the decrease threshold Δ are explicitly computable. Moreover, Δ^{-1} is polynomial in N, D, p.

The four branches of Algorithm 2 have the following guarantees:

1. **Near-endpoint rounding.** If $1 - |y_i| \leq d$, fix the nearest sign and remove the coordinate's coupling term. The two constraints can be repaired by increasing b by at most $d\|A_i\| \leq d$, so the true potential increases by at most d.

2. **Endpoint test.** Evaluate both scores $s_i^+ := \iota\beta_i v_i - (1 - y_i)$ and $s_i^- := \iota\beta_i u_i - (1 + y_i)$, with absolute error at most τ. If either approximate score is at least $-\tau$, fix that sign. The corresponding true score is at least -2τ, so the true potential increases by at most 2τ. If no score triggers, every true score is negative: the state is certified light.

3. **Gradient step.** Compute $\widetilde{g}$ with $\|\widetilde{g} - g\|_\infty \leq G/4$. If some $|\widetilde{g}_i| \geq 3G/4$, move by $-t_g \operatorname{sign}(\widetilde{g}_i) e_i$. The directional derivative is at most $-G/2$. Taylor's theorem and Eq. (620) give a true potential decrease at least $Gt_g/4 \geq \Delta$.

4. **Curvature step.** Otherwise $\|g\|_\infty < G$, so Eq. (616) applies. Compute a Hessian approximation with operator error at most $\chi/8$, and a minimum-eigenvalue vector to Rayleigh accuracy $\chi/8$. Allowing a further small normalization error, it yields a direction h with $\|h\|_2 \le 1$ and $h^\top \nabla^2 P_\rho h \le -\chi/2$. Evaluate both $P_\rho(y + t_c h)$ and $P_\rho(y - t_c h)$, each to absolute accuracy at most $\Delta/8$, and take the one with smaller approximate value. The average of the two true values is at most
$$P_\rho(y) - \chi t_c^2/4 + L_3 t_c^3/6 \le P_\rho(y) - \chi t_c^2/8.$$
Approximate selection loses at most $\Delta/4$, so the accepted step decreases the true potential by at least $\Delta/2$.

Steps 3 and 4 remain in the cube and in the larger smooth region $|y_i| \le 1 - d/2$; a coordinate entering the endpoint strip is handled on the next iteration. Testing both orientations supplies the curvature decrease.

There are at most N endpoint events. Their total potential increase is at most $N(d + 2\tau)$. All other accepted moves decrease the nonnegative potential by at least $\Delta/2$. Therefore the number of iterations is at most
$$N + \lceil \frac{2(B + N(d + 2\tau))}{\Delta} \rceil. \tag{621}$$
This is polynomial in N, D, p. Every iteration performs polynomial-work finite-accuracy convex solves, matrix arithmetic, linear solves, and a possible Hessian eigenvector computation.

Final discrepancy. The acceptance margins in Section 11.14 guarantee a decrease of the true potential at each non-endpoint move. Initialization and endpoint accounting give for the retained inputs
$$\| \sum_{\substack{i=1 \\ \mathrm{tr}[A_i] > \delta_{\mathrm{cut}}}}^{N} \sigma_i A_i \| \le 2\sqrt{\iota r^{1/p} \alpha + 2D\rho} + N(d + 2\tau).$$
Restoring the discarded inputs and using $\sqrt{x + y} \le \sqrt{x} + \sqrt{y}$,
$$\| \sum_{i=1}^{N} \sigma_i A_i \| \le 8\sqrt{(p+3) r^{1/p} \alpha} + 2\sqrt{2D\rho} + N(d + 2\tau + \delta_{\mathrm{cut}}). \tag{622}$$
Put $L_r := \log(2r)$. Our integer choice of p satisfies
$$r^{1/p} \le \sqrt{2}, \qquad p + 3 \le \frac{2}{\log 2} \log r + 7 = \frac{2}{\log 2} L_r + 5 \le \frac{7}{\log 2} L_r.$$
Since $\sqrt{2} < 10/7$ and $\log 2 > 2/3$,
$$8\sqrt{(p+3) r^{1/p} \alpha} < 8\sqrt{15 \alpha L_r} < 31\sqrt{\alpha L_r},$$
where $64 \cdot 15 = 960 < 31^2$. The additive terms in Eq. (622) are less than $0.004/N$, while $\sqrt{\alpha L_r} \ge \sqrt{\log 2/N} > 0.004/N$. Thus Eq. (622) is strictly below $32\sqrt{\alpha L_r}$.

Finally every signing of the original PSD isotropic family lies between $-I$ and I, establishing the cap by one in Eq. (623).

11.15 The deterministic construction

We now turn the quantitative descent estimates in Section 11.14 into a deterministic algorithm by evaluating each step to a certified finite accuracy.

Theorem 11.24 (Deterministic polynomial time construction)**.** *Let $A_1, \dots, A_N \in \mathbb{C}^{D\times D}$ be Hermitian positive semidefinite matrices with $\sum_{i=1}^N A_i = I_D$, $\|A_i\| \le \alpha$, and $\operatorname{rank}(A_i) \le r$, where $r \ge 1$ and $\alpha > 0$. A deterministic algorithm uses a polynomial number of real arithmetic operations and comparisons in N and D and returns signs $\sigma_1, \dots, \sigma_N \in \{-1, 1\}$ satisfying*

$$\|\sum_{i=1}^N \sigma_i A_i\| \le \min\{1, 32\sqrt{\alpha \log(2r)}\}. \tag{623}$$

Proof. Choose the parameters in Eq. (601) and apply Algorithm 2. Section 11.13 implements every requested value, derivative, and endpoint evaluation with polynomial work. The iteration bound in Eq. (621) is polynomial in N, D, p, and $p = O(\log(2D))$ after truncating the rank bound to D. Thus the total arithmetic work is polynomial in N and D. Section 11.14 bounds the norm after restoring the discarded inputs by $32\sqrt{\alpha \log(2r)}$. Isotropy gives the additional bound by one, proving Eq. (623). □

Algorithm 2 uses the endpoint tests before taking a gradient or curvature step. Its evaluation tolerances are those of Section 11.14.

12 Deterministic algorithm for rank-dependent signing

Let $A_1, \dots, A_m \in \mathbb{C}^{n\times n}$ be Hermitian positive semidefinite matrices with $\sum_i A_i = I_n$, $\|A_i\| \le \alpha$, and $\operatorname{rank}(A_i) \le r$. Throughout this section, running time counts real arithmetic operations and comparisons. We use the functional estimates of Section 8.

12.1 Normalization and the maintained potential

Cap the integer rank bound by n. For $1 \le r \le 2$, use the algorithm of Theorem 9.1; its discrepancy and cost are verified in the proof of Theorem 12.15. For the remaining ranks, use the order-three budget from Section 11.5, and the square-root budget from Section 8.8, with parameters

rank range	p	$\beta(y)$	ι
3	3	$\sqrt{1-y^2}f(\lvert y\rvert)$	5.403
$4 \le r \le 32$	3	$\sqrt{1-y^2}f(\lvert y\rvert)$	5.65
$r \ge 33$	$\lfloor \log_2 r \rfloor$	$\sqrt{1-y^2}$	$3(p+1)/2$

(624)

Here f is the exact even polynomial in Eq. (480), with μ and a from Eqs. (481)–(482). For $r \ge 33$, determine $p \ge 5$ by comparing r with successive integer powers of two. This takes $O(\log n)$ operations and avoids an exact comparison at an integer value of $\log r$. For each run define

$$s^2 := \iota\alpha r^{1/p}, \qquad C_i := A_i/s.$$

If the stated discrepancy bound is at least one, arbitrary signs suffice. Otherwise $\alpha < 1$. Define $f_i(X) := \|C_i^{1/2} X C_i^{1/2}\|_{S_p}$. Then

$$\|C_i\| \le 1, \qquad \sum_{i=1}^m \|C_i\| C_i \preceq I/\iota, \qquad \sum_{i=1}^m f_i(I) C_i \preceq I/\iota, \qquad \sum_{i=1}^m C_i^2 \preceq I/\iota, \tag{625}$$

Algorithm 2 Deterministic square-root-budget signing

procedure RankSigning($A_1, \dots, A_N, r$)
  Set $r \leftarrow \min\{r, D\}$ and choose parameters from Eq. (601)
  Set $I \leftarrow \{i : \operatorname{tr}[A_i] > \delta_{\mathrm{cut}}\}$ and $y_i \leftarrow 0$ for $i \in I$
  Compute G, χ from Eq. (615)
  Compute t_g, t_c, Δ from Eqs. (618)–(620)
  while $I \neq \varnothing$ **do**
    if some $i \in I$ satisfies $1 - |y_i| \le d$ **then**
      $\sigma_i \leftarrow \operatorname{sign}(y_i)$, $y_i \leftarrow \sigma_i$, $I \leftarrow I \setminus \{i\}$
    else
      Evaluate all endpoint scores to error at most τ ▷ Section 11.14
      if some approximate score $\widetilde{s}_i^{\pm} \ge -\tau$ **then**
        $\sigma_i \leftarrow \pm 1$ for that score, $y_i \leftarrow \sigma_i$, $I \leftarrow I \setminus \{i\}$
      else
        Compute $\widetilde{g}$ with $\|\widetilde{g} - \nabla P_\rho(y)\|_\infty \le G/4$
        if some $|\widetilde{g}_i| \ge 3G/4$ **then** ▷ Eq. (616)
          $y \leftarrow y - t_g \operatorname{sign}(\widetilde{g}_i) e_i$
        else
          Compute h with $\|h\|_2 \le 1$ and $h^\top \nabla^2 P_\rho(y) h \le -\chi/2$
          Evaluate $P_\rho(y \pm t_c h)$ to error at most $\Delta/8$ ▷ Section 11.13
          $y \leftarrow$ the candidate with smaller approximate value
        end if
      end if
    end if
  end while
  Set $\sigma_i \leftarrow 1$ for every input discarded by the trace test
  Assemble $\sigma \leftarrow (\sigma_1, \dots, \sigma_N)$
  return σ
end procedure

These inequalities follow from $f_i(I) \le r^{1/p}\|C_i\|$ and $A_i^2 \preceq \alpha A_i$. An enclosing approximation to s with sufficiently small fixed relative error is enough; the final constant has slack for it. Since $\alpha \ge 1/m$, this requires only polylogarithmic scalar work.

Vectorize Hermitian matrices over the reals in dimension $d := n^2$. Precompute all matrix sums on a balanced binary tree of the inputs in $O(mn^2)$ operations. Initially the root group has coefficient zero. Maintain at most d fractional groups. Split these groups, copying their coefficients to their children. There are at most $2d$ fractional groups at a level. Round their coefficients while preserving their vector sum until at most d remain fractional. A group fixed at a sign is not split again; all its original inputs receive that sign.

Each level costs $O(d^3)$ operations. Find a maximal independent set of the at most $2d$ group columns. Express a nonbasis column in that basis, and move along the resulting dependence until a coefficient reaches a cube endpoint. If a basis column freezes, replace it by the participating nonbasis column; otherwise delete the frozen nonbasis column. Maintain a rank-revealing elimination factorization. Each dependence solve, basis replacement, or column deletion costs $O(d^2)$, including rank drops and simultaneous endpoint hits. Each move freezes at least one coefficient. There are at most $2d$ moves, in addition to the initial $O(d^3)$ elimination. Zero columns can be fixed directly. This count uses exact arithmetic and comparisons.

After $O(\log(2m))$ levels the coefficients satisfy

$$y \in [-1,1]^m, \qquad \sum_{i=1}^{m} y_i C_i = 0, \qquad k := |\{i : |y_i| < 1\}| \le n^2. \tag{626}$$

The cost is $\widetilde{O}(mn^2 + n^6)$. Retain the contribution of all fixed coordinates in M. Since $\beta(y) \le 1$, removing their coupling terms leaves the initial nonlinear coupling bounded by I.

Throughout this section $c > 0$ is a sufficiently small absolute constant, decreased a finite number of times if necessary. Set $L := \log(4n)$. The symbol P denotes a sufficiently large fixed polynomial in L; it may be enlarged a finite number of times, with degree independent of n. Defer each remaining input with $\operatorname{tr}[C_i] \le c/n^2$, preserving its fractional contribution in M and deleting its coupling term. Arbitrary final signs for these inputs cost at most $2c$ in norm. Every other input satisfies $\|C_i\| > c/n^3$.

An entropy-smoothed potential with a fixed reservoir. Set

$$\rho := c/n, \qquad \tau := c/L, \qquad \lambda := \frac{c}{n^2(1 + 16 A_n B_n)}.$$

The known fixed polynomials A_n, B_n are chosen in Lemma 12.2. In particular $\lambda^{-1} = n^2 \operatorname{polylog}(n)$ and $n^2\lambda \le c$. Use the budget selected in Eq. (624). Both budgets have $\beta(0) = 1$, vanish at the endpoints, and satisfy $0 \le \beta \le 1$. Set $\delta := c^2/(100L)$. On $|y| \le 1 - \delta/2$, their fixed-order derivatives and the ratios used below are polylogarithmically bounded. Moreover $-\beta'' \ge b_* > 0$ for an absolute b_*, as verified below. The reservoir floor and the relative estimates therefore lose only constants and logarithmic factors. For $X := (U, V) \succ 0$, let $\mathcal{C}(X, y)$ be block diagonal, with blocks

$$U^{-1} + M(y) + \iota \sum_{i \in I} \beta_i f_i(V) C_i, \qquad V^{-1} - M(y) + \iota \sum_{i \in I} \beta_i f_i(U) C_i.$$

Here I is the active set, while M includes all fixed and deferred contributions. Define

$$F(y) := \min_{U,V \succ 0} \{\rho \operatorname{tr}[U + V] + \tau \log \operatorname{tr}[\exp(\mathcal{C}(X, y)/\tau)]\} + \lambda \sum_{i \in I} \beta_i. \tag{627}$$

The minimization is convex: inversion and the coupling map are convex in Loewner order, and log-trace-exponential is convex and increasing in that order. At its optimizer let

$$Q := \operatorname{diag}(W, Z) = \frac{\exp(\mathcal{C}/\tau)}{\operatorname{tr}[\exp(\mathcal{C}/\tau)]}, \quad u_i := f_i(U),\ v_i := f_i(V),\ w_i := \operatorname{tr}[C_i W],\ z_i := \operatorname{tr}[C_i Z].$$

With $B_i(X) := Df_i(X)$ represented by its trace gradient, stationarity gives

$$\begin{aligned} W &= U(\rho I + \iota \sum_{i \in I} \beta_i z_i B_i(U))U, \\ Z &= V(\rho I + \iota \sum_{i \in I} \beta_i w_i B_i(V))V, \qquad \operatorname{tr}[W + Z] = 1. \end{aligned} \tag{628}$$

On a bounded potential sublevel there are absolute constants $a, b > 0$ with

$$aI \preceq U, V \preceq b\sqrt{n} I, \qquad W, Z \succeq \rho a^2 I. \tag{629}$$

Indeed $F \geq \|M\|$, and the soft maximum bounds each constraint block from above. Their inverse and coupling parts are positive, which bounds the inverses. Taking traces in Eq. (628) gives $\rho \operatorname{tr}[U^2 + V^2] \leq 1$. Existence, uniqueness, and local smoothness follow from these bounds and strict convexity with the positive optimal dual fixed. The same bounds hold on slightly larger bounded sublevels used for trial evaluations. The scalar initialization $U = V = \ell I$ and Eq. (625) give

$$F(y_{\text{initial}}) \leq 2\sqrt{1 + 2n\rho} + \tau \log(2n) + n^2\lambda = 2 + O(c). \tag{630}$$

The reservoir is added only at initialization.

Boundary events and deferred rounding. If $1 - |y_i| \leq \delta$, defer that coordinate: retain its fractional contribution in M and remove its coupling term and reservoir. The potential does not increase. For every other coordinate, test the two usual endpoint scores with absolute error at most $\eta := c/n^2$, accepting a score at least $-\eta$. Fixing its sign and removing its coupling term increases both constraint blocks, at the old primal point, by at most $2\eta\|C_i\|I$. Thus the true potential increase is at most 2η. There are at most n^2 such events. If no test passes, the true state is light and $|y_i| \leq 1 - \delta$.

All deferred near-endpoint coordinates can be rounded with norm error $O(c)$. Put $h_0 := c/(100L)$. For the set H of deferred near-endpoint coordinates with $f_i(I) > h_0$, Eq. (625) gives $\sum_{i \in H} C_i \preceq I/(\iota h_0)$, so nearest-sign rounding costs at most $\delta/(\iota h_0)$. For the remaining deferred near-endpoint coordinates, indexed by J, we have $\|C_i\| \leq h_0$. Independent endpoint signs with means y_i have centered increments $X_i = (\sigma_i - y_i)C_i$ satisfying

$$\|X_i\| \leq 2h_0, \qquad \sum_{i \in J} \mathbb{E}[X_i^2] = \sum_{i \in J} (1 - y_i^2) C_i^2 \preceq 2\delta I/\iota.$$

Matrix Bernstein therefore gives error $O(c)$ with positive probability. This rounding is made deterministic using the two-sided trace-exponential pessimistic estimator and conditional expectations. Its logarithmic moment matrices are matrix functions of each original C_i, so the construction applies to arbitrary rank. The required trace-exponential concavity is due to Tropp [Tro12, Corollary 3.3]. Choose the initial estimator strictly below the threshold for error Cc and reserve a fixed part of this gap for numerical evaluation. The estimator never increases, so the final signs satisfy the bound. A constant-factor uncertainty band in the $f_i(I)$ threshold only changes constants. Dense spectral evaluation of the conditional estimators costs at most $\widetilde{O}(n^6)$ over the at most n^2 deferred coordinates. The small-input error is added separately.

12.2 Descent directions and the outer-step bound

At a retained light state, define $p_i := 1 + \iota\beta_i' v_i$, $q_i := 1 - \iota\beta_i' u_i$, and $b_i := -\beta_i''$. The coupling and movement weights are

$$S_i := v_i w_i + u_i z_i, \tag{631}$$
$$R_i := p_i u_i - q_i v_i, \tag{632}$$
$$D_i := p_i^2 u_i w_i + q_i^2 v_i z_i, \tag{633}$$
$$t_i := p_i q_i S_i, \tag{634}$$
$$\omega_i := \iota b_i S_i + \lambda b_i. \tag{635}$$

Set $\Omega := \operatorname{diag}(\omega_i)$, $g := \nabla F$, $\ell_i := g_i - \lambda\beta_i'$, and $\Gamma_i := 1/(\iota\beta_i u_i v_i)$. For order three use $\mu_i := \mu(|y_i|)$ from Section 11.5; for $p \geq 5$ set $\mu_i := 0$. Set $\eta_i := p - 1 - (p+1)\mu_i$ and $L_{p,i} := b_i - (\beta_i')^2/(\eta_i \beta_i)$.

The outer derivative and the movement weight satisfy the exact identities

$$\ell_i = p_i w_i - q_i z_i, \tag{636}$$
$$D_i = t_i + R_i \ell_i. \tag{637}$$

These identities hold at every retained light state.

We first extend the joint energy–moment projection to a light state with a nonzero outer gradient. Let $K_{0,i}$ denote the maximum of the two coefficients in Lemma 11.5, with $\chi = 69/100$ for rank three and $\chi = 5/16$ for ranks four through thirty-two. Define

$$\zeta_i := \frac{(p_i u_i + q_i v_i)|\ell_i|}{D_i}. \tag{638}$$

Lemma 12.1 (Stability of the joint energy–moment projection)**.** *For order three there is an absolute constant $C > 0$ such that, for $0 < \delta_{\mathrm{stab}} \le 10^{-3}$, the response graph has a covariance $\mathcal{A}$, positive definite on that graph, satisfying*

$$\begin{aligned} \operatorname{tr}[D_0 \mathcal{A}] - \sum_i \Gamma_i\big((\tfrac{1}{2} - \delta_{\mathrm{stab}}) e_i^\top T \mathcal{A} T^\top e_i + \mu_i f_i^\top T \mathcal{A} T^\top f_i\big) \\ \le \sum_i K_i^* e_i^\top (I - T) \mathcal{A} (I - T)^\top e_i, \end{aligned} \tag{639}$$
$$K_i^* \le K_{0,i} + C(\delta_{\mathrm{stab}} + \zeta_i). \tag{640}$$

Here D_0 is the response-energy matrix before diagonal padding. The induced movement covariance is nonzero.

Proof. The Gibbs dual equations in Eq. (628) give the response graph and coercivity $D_0 \succeq T^\top \Gamma T$, as in Eq. (466). Form the orthonormal pair e_i, f_i from $E_{a,i} := c_i^2 u_i w_i$, $E_{b,i} := d_i^2 v_i z_i$, where $c_i^2 := \beta_i p_i / q_i$ and $d_i^2 := \beta_i q_i / p_i$. Direct substitution gives

$$e_i^\top (I - T)\xi = \sqrt{D_i}\, h_i, \qquad f_i^\top (I - T)\xi = 0, \tag{641}$$
$$(\sqrt{w_i/u_i}, \sqrt{z_i/v_i}) \cdot e_i = \frac{\ell_i}{\sqrt{D_i}}, \tag{642}$$
$$\|(\sqrt{w_i/u_i}, \sqrt{z_i/v_i})\|_2^2 = \frac{S_i}{u_i v_i}. \tag{643}$$

For example $E_{a,i} + E_{b,i} = \beta_i D_i / (p_i q_i)$, which gives the first and second identities.

Fix a coordinate and suppress its index. Normalize the input compressions as in Section 11.5, using C_i in place of A_i. Write $e_U = \operatorname{tr}[R_U S_U]$, $\kappa_U = \operatorname{tr}[R_U^4]$, and similarly for V. Set $\alpha_E := p^2 uw / D$ and $\bar\kappa := \alpha_E \kappa_U + (1 - \alpha_E)\kappa_V$. The exact identities away from stationarity are

$$\begin{aligned} e^\top D_0 e &= \alpha_E e_U + (1 - \alpha_E) e_V, \\ f^\top D_0 f &= (1 - \alpha_E) e_U + \alpha_E e_V \le 1, \\ -\Gamma e^\top T e &= \frac{pq(uz\kappa_U + vw\kappa_V)}{D} \\ &= \bar\kappa - \frac{\ell(pu\kappa_U - qv\kappa_V)}{D} \ge \bar\kappa - \zeta, \qquad f^\top T f = -e^\top T e. \end{aligned} \tag{644}$$

The last bound uses $0 \le \kappa_U, \kappa_V \le 1$. For $h_0 = a_0 w_e \ge j_0 = b_0 w_f$ and $c_0 = (h_0 - j_0)/\Gamma$, Lemma 11.4 therefore gives

$$w_e e^\top D_0 e + w_f f^\top D_0 f + h_0 e^\top T e + j_0 f^\top T f$$

$$\le w_f + \max\{w_e - (1-\zeta)c_0, (3+\chi)w_e/4 - (\chi-\zeta)c_0\}. \tag{645}$$

Thus the energy coefficients remain 1 and $(3+\chi)/4$; only the moment coefficients decrease by ζ.

We verify a uniform perturbation bound, including the boundary $\Gamma = \infty$. Put $\mathfrak{m} := 1/2 - \delta_{\mathrm{stab}}$ and use $\Gamma = (1+x)/(x(x+\mu))$ with $0 < x < 1/2$. The scalar certificates give $\mu > 10^{-4}$ and $13/5 < a_0 = a(|y_i|) < 16/5$. Set

$$V_0 = 1 + x, \quad X = x(x+\mu), \quad F_0 = \mu + 2x + x^2, \quad b_0 = 2\Gamma x, \quad w_f = X/(V_0 F_0).$$

When $0 \le \zeta \le 10^{-3}$, for each pair

$$(q_*, d_*) = (1, 1-\zeta), \qquad (q_*, d_*) = ((3+\chi)/4, \chi - \zeta),$$

define

$$\begin{aligned} J_* &= F_0 - 2d_* x X, & Y_* &= q_* V_0 + (a_0^2/4 - d_* a_0) X, \\ w_{e,*} &= J_*/(F_0 Y_*), & F_{e,*} &= 1/w_{e,*}, \\ K_*(\mathfrak{m}) &= \frac{a_0^2(1-\mathfrak{m})/4 + \mathfrak{m}\Gamma F_{e,*}}{\mathfrak{m}\Gamma + a_0 + a_0^2/(4\Gamma) - F_{e,*}}. \end{aligned}$$

Take $w_e = \min_* w_{e,*}$ and $K^* = \max_* K_*(\mathfrak{m})$. The denominator estimates in Lemma 11.5 give $J_* \ge F_0/2$ and $Y_*/V_0 \ge q_* - d_*^2/\Gamma > 0$. Decreasing d_* by ζ increases J_* by $2\zeta x X$ and Y_* by $\zeta a_0 X$. Since $2X \le 1$,

$$F_{e,*} \le 81/20 + (16/5)\zeta.$$

Consequently

$$\frac{\mathfrak{m}\Gamma + a_0 + a_0^2/(4\Gamma) - F_{e,*}}{\Gamma} \ge \frac{499}{1000} - \frac{81/20 + 16/5000 - 13/5}{3} = \frac{73}{5000} > \frac{1}{100}. \tag{646}$$

The ordering of the row multipliers also persists, because

$$a_0 J_* - 2x Y_* = a_0 F_0 - 2q_* x V_0 - a_0^2 x X/2 > 0.$$

This expression is independent of d_*, so the same estimate as in Lemma 11.5 applies. The row normalization and square completion are

$$\begin{aligned} q_* w_{e,*} + w_f + \frac{a_0^2 w_{e,*} + b_0^2 w_f}{4\Gamma} - d_* \frac{a_0 w_{e,*} - b_0 w_f}{\Gamma} &= 1, \\ K - \frac{(K - a_0/2)^2}{\mathfrak{m}\Gamma + K} + \frac{a_0^2}{4\Gamma} &= 1/w_e. \end{aligned}$$

The derivative of K with respect to F_e is $(\mathfrak{m}\Gamma + a_0/2)^2/(\mathfrak{m}\Gamma + a_0 + a_0^2/(4\Gamma) - F_e)^2 > 0$. Taking the smaller w_e makes both row normalizations at most one. Equation (645) and the row completions in Lemma 11.5 now give Eq. (639).

To bound the coefficient, clear denominators:

$$K_*(\mathfrak{m}) = \frac{V_0[(1-\mathfrak{m})a_0^2 X J_*/4 + \mathfrak{m} V_0 F_0 Y_*]}{\mathfrak{m} V_0^2 J_* + a_0 V_0 X J_* + a_0^2 X^2 J_*/4 - V_0 X F_0 Y_*}.$$

By Eq. (646), the denominator is at least $V_0^2 J_*/100 \ge \mu/200 > 1/2000000$, also at $x = 0$. On this parameter box the numerator has absolute value less than 16; its derivatives in $\mathfrak{m}, \zeta$ are less than

16, 8, respectively. The corresponding denominator derivatives are less than 4, 5. These bounds follow from $V_0 \le 3/2$, $X \le 1/2$, $F_0, J_* \le 7/4$, $Y_* \le 3$, $\partial_\zeta J_* = 2xX \le 1/2$, and $\partial_\zeta Y_* = a_0 X \le 8/5$. Both derivatives of K_* are therefore less than 10^{15} in absolute value. Taking the maximum preserves this Lipschitz bound, and gives $K_i^* \le K_{0,i} + 10^{15}(\delta_{\mathrm{stab}} + \zeta_i)$.

When $\zeta_i > 10^{-3}$, use the parameterized projection of Eq. (375) with zero retained energy, parameter $\theta = 3/2$, and coercivity $\Gamma_0 = 15/4$. Its coefficient is $45/19 < 3$, while the actual $\Gamma_i > \iota \ge 5.403$. The unpadded energy has diagonal entries at most one, so the same trace comparison applies. Subtracting the nonnegative retained energies preserves the inequality. Since $3 \le 3000\zeta_i$ and $K_{0,i} > 0$, the coefficient bound holds in this case as well.

These choices can be combined coordinatewise. If R spans the response graph and D contains the chosen row weights, use $\mathcal{A} = \mathsf{R}(\mathsf{R}^\top \mathsf{D}^{-1} \mathsf{R})^{-1} \mathsf{R}^\top$. Use actual coercivity on the perturbed rows and Γ_0 on the remaining rows when completing the complementary energy. This diagonal coercivity is no larger than Γ. The row normalizations are at most one, and $\operatorname{tr}[\mathsf{D}^{-1}\mathcal{A}] = k$, so the same trace proof gives Eq. (639). If movement vanished on the whole graph, T would be the identity there. Coercivity would then contradict $\operatorname{tr}[D_0] \le 2k$ and $\min_i \Gamma_i > 2$. □

The scalar certificates supply a uniform margin, which absorbs the loss in this perturbation. For order three they give an effectively computable absolute $d_0 > 0$ such that

$$\iota L_{3,i} - 4K_{0,i} p_i q_i \ge d_0. \tag{647}$$

Indeed, use the polynomials in Eq. (477), with $t = |y_i|$ distinct from the movement weight t_i. For each branch put $\mathcal{R}_* := 4\eta A(1+t)\sqrt{1-t^2}V_0 B_*$. At the maximizing light pair the exact factorization is

$$W_* - \mathcal{R}_* = \eta(1-t^2)^{3/2} f D_*[\iota L_3 - 4pqK_*].$$

The physical certificate leaves give $P_* = (W_* - \mathcal{R}_*)(W_* + \mathcal{R}_*) > 10^{-7}$. Choose positive rational upper bounds M_+ and M_0 for $W_* + \mathcal{R}_*$ and the displayed prefactor, uniformly over both branches and couplings. Absolute monomial coefficient sums and $\sqrt{1-t^2} \le 1$ give such bounds. Then $d_0 := 10^{-7}/(M_+ M_0)$ is valid. The G leaves contain no physical point. The fixed-product maximization of pq in Section 11.5 transfers this gap to every light state; it does not use outer stationarity. The same certificates give $b_i \ge b_* > 0$ and positive lower bounds on the scalar denominators. On the retained strip their required derivatives and ratios are polylogarithmically bounded.

For $p \ge 5$ retain the square-root budget and put $a_i := (p+1)\kappa(\Gamma_i, 0)$. The order-five certificate and the monotonic comparison in Section 8.8 give

$$\iota L_{p,i} - a_i p_i q_i \ge \frac{p+1}{30 M_5}, \tag{648}$$

where $M_5 := 1 + \|W_5^0(t, 2u/5)\|_1 + 12(1+2/5)$, $W_5^0(t,x) := 9((1-t^2/4)(1-x^3) + 6tx^2)$, and $\|\cdot\|_1$ is the sum of absolute monomial coefficients. Here $a_i \le P$ and $b_i \ge 1$. Fix $d_* := \min\{d_0, 1/(5M_5)\} > 0$.

The uniform scalar gap gives a descent direction at each retained light state: either a coordinate gradient exceeds its threshold, or the normalized Hessian has negative curvature.

Lemma 12.2 (A weighted gradient or curvature direction)**.** *There are known fixed polynomials A_n, B_n in $\log(4n)$ such that with*

$$\lambda = \frac{c}{n^2(1 + 16 A_n B_n)}, \qquad 0 < c \le 1, \tag{649}$$

at every retained light state either some coordinate satisfies

$$|g_i| > \frac{\sqrt{\omega_i}}{P\sqrt{n}},$$

or the Hessian satisfies

$$\lambda_{\min}(\Omega^{-1/2}\nabla^2 F\Omega^{-1/2}) \le -1/P.$$

The tests admit inverse-polynomial accuracy margins.

Proof. Along the comparison response curve, both constraint blocks are fixed. The log-trace-exponential is therefore constant. For order three, define $E_i := e_i^\top T\xi$ and $F_i := f_i^\top T\xi$. The functional Hessian estimate retains the negative energy $-2\Gamma_i(E_i^2+F_i^2)$. Absorb the f_i component of the mixed term with $\eta_i\Gamma_i F_i^2$, and its e_i component with $4\delta_{\mathrm{stab}}\Gamma_i E_i^2$. The response identities in Lemma 12.1 and scalar Young's inequality give

$$h^\top\nabla^2 F h \le 4[Q - \sum_i \Gamma_i((\tfrac12-\delta_{\mathrm{stab}})E_i^2+\mu_i F_i^2)] - \sum_i(\iota L_{3,i}S_i+\lambda b_i)h_i^2 + \sum_i \frac{(\beta_i')^2\ell_i^2}{4\delta_{\mathrm{stab}}\beta_i^2\Gamma_i D_i}h_i^2. \tag{650}$$

Using the full squared norm $S_i/(u_iv_i)$ for the f_i component only enlarges its bound.

The dual equations give the lower bound

$$w_i \ge \rho\,\mathrm{tr}[C_iU^2] \ge \rho a\,\mathrm{tr}[C_iU] \ge \rho a u_i. \tag{651}$$

The first step uses Eq. (628). The second uses $U \succeq aI$. The last uses the trace bound on the Schatten norm. The same argument gives $z_i \ge \rho a v_i$.

Consequently,

$$R_i^2 \le (p_iu_i+q_iv_i)^2 \le \frac{2D_i}{\rho a} \le CnD_i. \tag{652}$$

The first step uses $|p_iu_i - q_iv_i| \le p_iu_i + q_iv_i$. The second uses $(p_iu_i+q_iv_i)^2 \le 2p_i^2u_i^2+2q_i^2v_i^2$ and the dual-weight lower bounds. The last uses $\rho = c/n$.

The same lower bounds and the arithmetic–geometric mean inequality give

$$\frac{1}{\beta_i^2\Gamma_iD_i} \le \frac{\iota}{2\beta_i\rho a p_iq_i}. \tag{653}$$

The slope certificates give a positive lower bound on p_iq_i; for the square-root budget $p_iq_i \ge 1/2$. Hence the final coefficient in Eq. (650) is at most $nP\ell_i^2$ when $\delta_{\mathrm{stab}}^{-1} \le P$. From Eqs. (637) and (652), Young's inequality gives, for $0<v<1$,

$$D_i \le 2t_i + Cn\ell_i^2, \tag{654}$$

$$|R_i\ell_i| \le D_i\zeta_i \le vt_i + C_vn\ell_i^2, \tag{655}$$

where $C_v = O(1+v^{-1})$. For Eq. (655), substitute Eq. (654) into $D_i\zeta_i = (p_iu_i+q_iv_i)|\ell_i| \le \sqrt{CnD_i}|\ell_i|$ and apply Young again.

Apply Lemma 12.1 and use its response identities to transfer the covariance to physical directions. The leading positive term is $4K_{0,i}D_i$; its difference from $4K_{0,i}t_i$, and the projection losses, are bounded by $C\delta_{\mathrm{stab}}D_i + PD_i\zeta_i$. Choose δ_{stab} and v smaller than suitable fixed multiples of $d_*/(P(1+\sup_i p_iq_i))$. Their reciprocals are polylogarithmic. Eqs. (647) and (655) absorb these terms into half the scalar gap, apart from $nP\ell_i^2$.

For $p \ge 5$, the unperturbed projection gives a nonzero movement covariance K with coefficient $a_iD_i - \iota L_{p,i}S_i - \lambda b_i$. Its scalar gap and Eq. (655) give the same conclusion. Thus in every branch there is a nonzero $K \succeq 0$ and a known fixed polynomial A_n such that

$$\mathrm{tr}[K\nabla^2 F] \le \sum_i[-\tfrac12 d_*S_i - \lambda b_i + nA_n\ell_i^2]K_{ii}. \tag{656}$$

The bounds defining A_n are uniform for $0 \le n^2\lambda \le c$: the reservoir does not enter the inner equations and is nonnegative in the sublevel bound. Thus A_n can be fixed before choosing the reservoir in Eq. (649).

Choose a known polynomial $B_n \ge \sup_i(\beta_i')^2/b_i$ on the retained strip. Eq. (649) implies $2nA_n\lambda^2(\beta_i')^2 \le \lambda b_i/8$. Since $\ell_i = g_i - \lambda\beta_i'$, it follows that

$$\operatorname{tr}[K\nabla^2 F] \le \sum_i [-\tfrac{1}{2}d_* S_i - \tfrac{3}{4}\lambda b_i + 2nA_n g_i^2]K_{ii}.$$

Choose a known polylogarithmic H_n with $\omega_i \le H_n(d_* S_i + \lambda b_i)$, and enlarge P so that $P \ge 4H_n$ and $P^2 \ge 16A_nH_n$. If

$$|g_i| \le \frac{\sqrt{\omega_i}}{P\sqrt{n}} \quad \text{for every } i, \tag{657}$$

then $\operatorname{tr}[K\nabla^2 F] \le -(4H_n)^{-1}\operatorname{tr}[K\Omega]$. Since $K \ne 0$ and $\Omega \succ 0$, this gives

$$\lambda_{\min}(\Omega^{-1/2}\nabla^2 F\Omega^{-1/2}) \le -1/P. \tag{658}$$

Otherwise a coordinate gradient exceeds its threshold. Also

$$\Omega \succeq b_*\lambda I \succeq (Pn^2)^{-1}I, \qquad D_i \le P(\omega_i + ng_i^2). \tag{659}$$

The latter follows from Eq. (654) and the same reservoir estimate. These bounds give inverse-polynomial accuracy margins. The row weights and $\mu(|y_i|)$ enter the curvature proof. The algorithm obtains its direction from the normalized Hessian. The budgets have fixed degree and polylogarithmically bounded derivatives on the retained strip. The smaller reservoir therefore changes only constants and logarithmic factors in the finite-step estimates below. □

The same projection argument also controls the dimension of the negative spectral subspace. This will allow larger curvature moves when many coordinates remain active.

Lemma 12.3 (Many negative eigenvalues)**.** *After enlarging the fixed polynomial P, if Eq. (657) holds, then $\Omega^{-1/2}\nabla^2 F\Omega^{-1/2}$ has at least k/P eigenvalues at most $-1/P$. The conclusion admits inverse-polynomial accuracy margins.*

Proof. Write $\Pi_e := \sum_i e_i e_i^\top$ and $\Pi_f := \sum_i f_i f_i^\top$. Let D be the diagonal matrix of row weights used in the projection proof. Uniformly over its coordinatewise choices,

$$\mathsf{D}^{-1} \preceq C\Pi_e + C\Gamma\Pi_f \tag{660}$$

for an absolute constant C. For the perturbed order-three rows, the preceding formulas give $w_e^{-1} \le 5$ and $w_f^{-1} = \Gamma(\mu + 2x + x^2) \le 7\Gamma/4$. The fallback rows use fixed parameters and satisfy the same bound after increasing C. For the square-root rows, put $\Gamma = \vartheta(\vartheta+1)$ in the formulas of Section 8.8. They give $w_e^{-1} = O(1)$ and $w_f^{-1} = O(\vartheta+1) = O(\Gamma)$.

Restrict physical directions to any subspace of codimension $d \le k/4$. The response graph maps bijectively to physical directions, so its restricted graph has dimension $\ell = k - d$. Choose an orthonormal basis matrix R for this graph and set

$$C_R := R^\top \mathsf{D}^{-1} R, \qquad \mathcal{A} := RC_R^{-1}R^\top, \qquad \mathcal{B} := \mathsf{D} - \mathcal{A}.$$

Then $\mathcal{A}, \mathcal{B} \succeq 0$, $\mathcal{A}\mathsf{D}^{-1}\mathcal{B} = 0$, and $\operatorname{tr}[\mathsf{D}^{-1}\mathcal{A}] = k - d$. All row-square completions in Lemma 12.1 remain valid: the graph identity $\Pi_f(I - T)R = 0$ is preserved, and the complementary matrix is still

positive semidefinite. The sum of the row normalizations is at most k, whereas the projection trace is now $k-d$. Thus Eq. (639) holds with the additional term d on its right side. The square-root comparison has the same additional term. In the curvature estimate it is multiplied by 4 for order three and by $p+1$ for higher orders.

We next give a quantitative lower bound on movement. Coercivity and Eq. (660) imply

$$\operatorname{tr}[C_R] \le C\ell + C\|\Gamma^{1/2}\Pi_f T R\|_F^2 \le Ck + C\operatorname{tr}[D_0] \le C'k,$$
$$\|T\|_F^2 \le 2k/\gamma_0, \qquad \gamma_0 := \min_i \Gamma_i > 5.$$

For $\Pi := RR^\top$ we therefore have

$$\operatorname{tr}[(I-T)\Pi] \ge \ell - \sqrt{2k\ell/\gamma_0} \ge c_0 k,$$

because $3k/4 \le \ell \le k$. Applying Frobenius Cauchy–Schwarz to $(I-T)RC_R^{-1/2}$ and $RC_R^{1/2}$ gives

$$\operatorname{tr}[(I-T)\mathcal{A}(I-T)^\top] \ge c_1 k. \tag{661}$$

Let K be the induced covariance of physical directions. Its range lies in the prescribed subspace, and the response identities give $\sum_i D_i K_{ii} \ge c_1 k$.

Apply the proof of Lemma 12.2 to this restricted covariance. The drift, scalar Young, reservoir, and perturbation estimates are coordinatewise. The only new term is the projection trace defect just computed. Enlarging fixed polynomials gives

$$\operatorname{tr}[K\nabla^2 F] \le -P^{-1}\operatorname{tr}[K\Omega] + Pd, \qquad \operatorname{tr}[K\Omega] \ge k/P. \tag{662}$$

The second inequality follows from Eq. (661) and $D_i \le P\omega_i$, which follows from Eqs. (659) and (657).

Put $A := \Omega^{-1/2}\nabla^2 F\Omega^{-1/2}$ and let d count its eigenvalues below $-1/(2P)$. If $d \le k/4$, take the physical subspace obtained by applying $\Omega^{-1/2}$ to the orthogonal complement of these eigenspaces. The resulting K satisfies

$$\operatorname{tr}[K\nabla^2 F] \ge -(2P)^{-1}\operatorname{tr}[K\Omega].$$

Combining this with Eq. (662) gives $Pd \ge k/(2P^2)$. Thus $d \ge k/(2P^3)$. If $d > k/4$, this conclusion already holds after increasing P. Rename the fixed polynomials and reserve a fixed fraction of the negative spectral gap for finite accuracy. This proves the lemma. □

Relative remainders for the Schatten-norm functional. Finite outer steps require derivative bounds in the relative primal norm. For $X \succ 0$ and Hermitian H, define

$$\eta_X := \|X^{-1/2}HX^{-1/2}\|, \qquad q_i(H) := \operatorname{tr}[B_i(X)HX^{-1}H].$$

Lemma 12.4 (Relative derivatives of the Schatten-norm functional)**.** *There is an absolute constant C such that, for $|t|\eta_X \le 1/(Cp)$,*

$$\Big|\frac{d^2}{dt^2} f_i(X+tH)\Big| \le Cp^2 q_i(H), \qquad \Big|\frac{d^3}{dt^3} f_i(X+tH)\Big| \le Cp^3\eta_X q_i(H). \tag{663}$$

Proof. On the positive support, set $A := C_i^{1/2} X C_i^{1/2}$ and $K := C_i^{1/2} H C_i^{1/2}$. Then $K = A^{1/2}EA^{1/2}$ for a Hermitian E with $\|E\| \le \eta_X$. For nonnegative exponents summing to p and $j \ge 2$,

$$|\operatorname{tr}[A^{a_1}E\cdots A^{a_j}E]| \le \|E\|^{j-2}\operatorname{tr}[A^p E^2]. \tag{664}$$

To prove this, apply multivariable three-lines interpolation on the simplex of exponents a_j/p. At a vertex all positive powers occur in one factor A^p. Imaginary powers are unitary and commute with A^p. Weighted Hilbert–Schmidt Cauchy–Schwarz on the first and last E factors, and operator norms on the intervening factors, give the right side, uniformly over all imaginary translates. Three-lines interpolation extends this bound to the simplex. Restriction to the positive support handles singular inputs.

The coefficient of order $j \geq 2$ in $\operatorname{tr}[(A+tK)^p]$ is therefore at most $\binom{p}{j}\|E\|^{j-2}\operatorname{tr}[A^pE^2]$ in absolute value. The linear coefficient satisfies the ordinary and weighted Cauchy–Schwarz bounds. For $|t|\|E\| \leq 1/(Cp)$ the polynomial divided by $\operatorname{tr}[A^p]$ stays in a fixed neighborhood of one. Differentiate its analytic positive pth root three times to obtain Eq. (663). The support energy $\operatorname{tr}[A^pE^2]/f_i(X)^{p-1}$ is at most $q_i(H)$, by $C_i^{1/2}A^{-1}C_i^{1/2} \preceq X^{-1}$. All constants grow only polynomially in p. □

Let H and B denote the directional derivatives of the primal optimizer and Gibbs density, respectively, in direction h. Let $e(H)^2$ be the primal Lagrangian Hessian, including the positive functional Hessians, and put

$$\mathcal{E}(B)^2 := \tau \operatorname{tr}[BD\log_Q[B]], \qquad \theta := h^\top \Omega h.$$

The envelope identity is

$$h^\top \nabla^2 F h = -\theta + e(H)^2 + \mathcal{E}(B)^2 + 2\mathcal{M}, \qquad |\mathcal{M}| \leq P\sqrt{\theta}e(H). \tag{665}$$

Indeed differentiate the primal objective along its optimizer; primal stationarity cancels the second optimizer derivative, and the Gibbs variation contributes $\mathcal{E}(B)^2$. The remaining mixed term is the mixed derivative of the budget and functional, bounded by the first differential inequality for the functional and weighted Cauchy–Schwarz. Thus a nonpositive-curvature direction has $e(H)^2 + \mathcal{E}(B)^2 \leq P\theta$.

Define the relative response bound

$$R := \|h\|_\infty + \max\{\|U^{-1/2}H_UU^{-1/2}\|, \|V^{-1/2}H_VV^{-1/2}\|, \|W^{-1/2}B_WW^{-1/2}\|, \|Z^{-1/2}B_ZZ^{-1/2}\|\}.$$

For the primal relative norm use $U^{-1/2}WU^{-1/2} \succeq \rho a I$. For the dual norm, the coefficient of a relative matrix entry in the entropy metric is $\tau q_j q_l/\mathcal{L}(q_j, q_l)$, where $\mathcal{L}$ is the logarithmic mean. This is at least $\tau \min(q_j, q_l) \geq \tau\rho a^2$. Hence

$$R \leq \|h\|_\infty + P\sqrt{n\theta} \quad \text{in a nonpositive-curvature direction.} \tag{666}$$

Use the straight comparison paths $X_t := X + tH$, $Q_t := Q + tB$, and $y + th$. When $|t|R \leq 1/P$, they remain positive and comparable to their base points, and the active coordinates stay away from endpoints. The Gibbs equation and its derivative imply that

$$E_t := \mathcal{C}(X_t, y + th) - \tau \log Q_t - (b_0 + tb_1)I$$

vanishes to first order. There is a PSD absolute majorant J_t with

$$-J_t \preceq E_t \preceq J_t, \quad \|J_t\| \leq PR^2t^2, \quad \operatorname{tr}[Q_tJ_t] \leq P\theta t^2. \tag{667}$$

The inverse quadratic and the log-resolvent quadratic have their usual positive majorants. For the new quadratic term from the functional use Eq. (663). Its dual-weighted sum is controlled by

$$\iota \sum_{i\in I} \beta_i z_i q_i(H_U) \leq \operatorname{tr}[WU^{-1}H_UU^{-1}H_UU^{-1}],$$

and its exchanged counterpart, by Eq. (628). Its operator norm is controlled by

$$q_i(H_U) \le f_i(U)\|U^{-1/2} H_U U^{-1/2}\|^2$$

and the bounded coupling matrix. Mixed budget and functional terms use scalar Young and the first differential inequality for the functional. Integrating these second-derivative majorants twice proves Eq. (667). All functional comparisons use the relative estimate at the base point.

The saddle objective on these straight paths has third derivative bounded by $PR\theta$. The cubic term from the functional is bounded by Eq. (663); mixed terms use the same majorants; derivatives of the dual pair use $|\operatorname{tr}[BJ]| \le R \operatorname{tr}[QJ]$ for $J \succeq 0$; and the entropy cubic is bounded by its positive resolvent integral. Its second derivative at zero equals $h^\top \nabla^2 F h$, by stationarity and differentiated stationarity. Golden–Thompson and $e^x \le 1 + x + x^2$ for $|x| \le 1/2$ give

$$\begin{aligned} &\tau \log \operatorname{tr}[\exp(\log Q_t + E_t/\tau)] - \operatorname{tr}[Q_t E_t] \\ &\quad \le \tau^{-1} \operatorname{tr}[Q_t E_t^2] \\ &\quad \le \tau^{-1} \|J_t\| \operatorname{tr}[Q_t J_t] \\ &\quad \le PR^2 \theta t^4. \end{aligned} \tag{668}$$

The first step applies the exponential remainder bound. For the second, factor $E_t = J_t^{1/2} T J_t^{1/2}$ with $\|T\| \le 1$; then $E_t^2 \preceq \|J_t\| J_t$. The last step uses Eq. (667), absorbing τ^{-1} into P. Evaluating the primal objective at X_t therefore proves

$$F(y + th) \le F(y) + t g^\top h + \tfrac{1}{2} t^2 h^\top \nabla^2 F h + PR\theta |t|^3 + PR^2 \theta t^4. \tag{669}$$

For a direction of arbitrary curvature, the same proof holds with θ in the remainder bounds replaced by an upper bound on $\theta + e(H)^2 + \mathcal{E}(B)^2$.

The relative remainder converts the descent alternative into a uniform potential decrease. Boundary events contribute only the allowance already specified in Section 12.1.

Lemma 12.5 (Number of outer steps)**.** *With the step lengths specified below, every accepted gradient or curvature move decreases the potential by at least* $1/(Pn^2)$*. The construction uses* $\widetilde{O}(n^2)$ *outer iterations, including at most* n^2 *boundary events.*

Proof. In the curvature branch choose $h^\top \Omega h = 1$ and $h^\top \nabla^2 F h \le -1/P$, reserving a fixed part of the margin for approximation. Since $\Omega \succeq (Pn^2)^{-1} I$, $\|h\|_\infty \le Pn$. Eq. (666) gives $R \le Pn$. Choose $t = c/(Pn)$, enlarging P to preserve endpoint distance $\delta/2$. Evaluate both orientations. Their linear terms cancel in the average, so Eq. (669) gives a true decrease at least $1/(Pn^2)$. Both orientations are evaluated deterministically.

For a coordinate gradient step put $\theta = \omega_i$ and $g = g_i$. The movement-weight bound in Lemma 12.2, which holds without the small-gradient assumption, gives $D_i \le P(\theta + ng^2)$. Pairing the differentiated primal and dual saddle equations gives

$$e(H)^2 + \mathcal{E}(B)^2 \le P(\theta + ng^2) =: T, \qquad |F''| \le PT, \qquad R \le 1 + P\sqrt{nT}.$$

In this forcing estimate

$$\operatorname{tr}[W(p_i C_i)^2] + \operatorname{tr}[Z(q_i C_i)^2] \le a^{-1} D_i,$$

because $C_i^2 \preceq a^{-1} C_i U C_i$ and its V counterpart. Entropy Cauchy–Schwarz bounds its pairing with the dual response; the coupling forcing has magnitude at most $P\sqrt{\theta} e(H)$. Take a step against g of length a sufficiently small fixed multiple of

$$\min\{\delta,\ |g|/(PT),\ [P(1 + \sqrt{nT})]^{-1}\}.$$

Eq. (669), with T replacing the remainder weight, gives decrease at least a constant times this length times $|g|$. Failure of the coordinate-gradient test in Lemma 12.2 and $\theta \geq 1/(Pn^2)$ imply $|g| \geq 1/(Pn^{3/2})$ and $g^2/T \geq 1/(Pn)$. Each of the three candidate products is therefore at least $1/(Pn^2)$.

There are at most n^2 boundary events with total increase $O(c)$. All other moves decrease the nonnegative potential by $1/(Pn^2)$; Eq. (630) therefore bounds the number of iterations by $\widetilde{O}(n^2)$. Evaluate values to a small fixed fraction of the proved decrease and use separated thresholds for gradient and curvature tests. Their smallest relevant magnitudes are inverse polynomials, so these margins are attainable and each accepted move retains the stated true decrease. □

12.3 Solving the inner optimization problem

Fix an outer state y and omit the reservoir, which is constant in $X = (U, V)$. Define

$$\phi(X) := \rho \operatorname{tr}[U + V] + \tau \log \operatorname{tr}[\exp(\mathcal{C}(X, y)/\tau)].$$

The optimizer bounds in Eq. (629) permit fixed absolute constants $a_{\min}, B > 0$ such that the optimizer $X_* := (U_*, V_*)$ lies in the interior of

$$\mathcal{D} := \{(U, V) : a_{\min} I \preceq U, V \preceq B\sqrt{n} I\}.$$

Use the Frobenius pair norm on this space. The constants are uniform on the bounded outer sublevels, including those used for trial values.

The optimal dual density controls the objective gap in the Frobenius norm. This estimate determines the restart length of the inner solver.

Lemma 12.6 (Quadratic growth of the inner objective)**.** *For every $X \in \mathcal{D}$,*

$$\phi(X) - \phi(X_*) \geq c_g \|X - X_*\|_F^2, \qquad c_g := \frac{\rho}{B\sqrt{n}}. \tag{670}$$

Moreover $\nabla\phi$ is L_0-Lipschitz on $\mathcal{D}$, where

$$L_0 := 2{a_{\min}}^{-3} + \frac{p-1}{a_{\min}} + \tau^{-1}({a_{\min}}^{-2} + 1)^2 = \widetilde{O}(1). \tag{671}$$

Proof. Fix the optimal Gibbs density $Q_* = \operatorname{diag}(W_*, Z_*)$. The Gibbs variational formula gives

$$\phi(X) \geq \rho \operatorname{tr}[U + V] + \operatorname{tr}[Q_* \mathcal{C}(X, y)] - \tau \operatorname{tr}[Q_* \log Q_*],$$

with equality at X_*. The right side has zero primal gradient at X_*. Its coupling terms are convex and have nonnegative Bregman remainders. For $H := U - U_*$, the exact inverse remainder is

$$U^{-1} - U_*^{-1} + U_*^{-1} H U_*^{-1} = U_*^{-1} H U^{-1} H U_*^{-1}.$$

Eq. (628) gives $U_*^{-1} W_* U_*^{-1} \succeq \rho I$. Pairing the remainder with W_* therefore gives at least

$$\rho \operatorname{tr}[H U^{-1} H] \geq \frac{\rho}{B\sqrt{n}} \|H\|_F^2.$$

The exchanged block has the same bound. The linear terms cancel by stationarity, proving Eq. (670).

For smoothness, Schatten duality gives $B_i(X) \succeq 0$ and $\operatorname{tr}[B_i(X)] \le f_i(I)$. Thus, for $X \succeq a_{\min} I$ and Hermitian H,

$$|Df_i(X)[H]| \le f_i(I)\|H\|, \qquad 0 \le D^2 f_i(X)[H,H] \le \frac{p-1}{a_{\min}} f_i(I)\|H\|^2.$$

The second bound follows from the Hessian bound for f_i by bounding $HX^{-1}H \preceq {a_{\min}}^{-1}\|H\|^2 I$. Since $\iota \sum_{i\in I} \beta_i f_i(I) C_i \preceq I$, the coupling blocks have first and second operator derivative bounds independent of the active count. Together with the inverse derivatives, this gives

$$\|D\mathcal{C}[H]\| \le ({a_{\min}}^{-2} + 1)\|H\|_F, \qquad \|D^2\mathcal{C}[H,H]\| \le (2{a_{\min}}^{-3} + \frac{p-1}{a_{\min}})\|H\|_F^2.$$

The Hessian of $\tau \log \operatorname{tr}[\exp(A/\tau)]$ in direction E is at most $\tau^{-1}\|E\|^2$: its Gibbs covariance formula has logarithmic-mean coefficients bounded by their arithmetic means. The chain rule proves Eq. (671) throughout $\mathcal{D}$. □

An accelerated inner solve with feasible gradient queries. We use a restarted projected scheme. Start a block at $x_0 \in \mathcal{D}$, with $z_0 := x_0$ and $A_0 := 0$. Choose $a_j > 0$ satisfying $L_0 a_j^2 = A_j + a_j$, set $A_{j+1} := A_j + a_j$, and define

$$y_j := (A_j x_j + a_j z_j)/A_{j+1}, \tag{672}$$
$$g_j := \nabla\phi(y_j), \tag{673}$$
$$z_{j+1} := \operatorname{proj}_{\mathcal{D}}(z_j - a_j g_j), \tag{674}$$
$$x_{j+1} := (A_j x_j + a_j z_{j+1})/A_{j+1}. \tag{675}$$

All gradient queries remain in $\mathcal{D}$. Projection clips the eigenvalues of the two Hermitian blocks to $[a_{\min}, B\sqrt{n}]$.

Projection optimality implies

$$a_j\langle g_j, z_{j+1} - X_*\rangle \le \tfrac{1}{2}(\|z_j - X_*\|_F^2 - \|z_{j+1} - X_*\|_F^2 - \|z_{j+1} - z_j\|_F^2).$$

Combining this with convexity at y_j, smoothness from y_j to x_{j+1}, and $L_0 a_j^2 = A_{j+1}$ gives

$$\begin{aligned} A_{j+1}(\phi(x_{j+1}) - \phi(X_*)) + \tfrac{1}{2}\|z_{j+1} - X_*\|_F^2 \\ \le A_j(\phi(x_j) - \phi(X_*)) + \tfrac{1}{2}\|z_j - X_*\|_F^2. \end{aligned} \tag{676}$$

Since $A_N \ge (N+1)^2/(4L_0)$, a block satisfies

$$\phi(x_N) - \phi(X_*) \le \frac{2L_0\|x_0 - X_*\|_F^2}{(N+1)^2}.$$

By Lemma 12.6, taking $N = \lceil\sqrt{4L_0/c_g}\rceil = \widetilde{O}(n^{3/4})$ halves the objective gap. Restart with $x_0 = z_0 = x_N$. A scalar feasible start has polynomially bounded gap, so $O(\log n)$ blocks suffice for any prescribed inverse-polynomial objective accuracy. Eq. (670) then supplies the required optimizer accuracy.

A value/gradient oracle costs

$$\widetilde{O}(kn^\omega + n^4). \tag{677}$$

Compute the matrix powers defining the functionals and their gradients by matrix multiplication, then form the constraint blocks and Gibbs density. The n^4 term retains elementary gap-independent spectral approximation and also covers each projection. Since $p = O(\log n)$ and the trace cutoff bounds $f_i(X)$ below by an inverse polynomial, these matrix powers and scalar roots require only polylogarithmic precision lengths. One cold solve therefore costs

$$\begin{aligned}\widetilde{O}(n^{3/4}(kn^{\omega} + n^4)) &= \widetilde{O}(n^{3/4}(n^{\omega+2} + n^4)) \\ &= \widetilde{O}(n^{3/4}n^{\omega+2}) \\ &= \widetilde{O}(n^{\omega+11/4}).\end{aligned} \tag{678}$$

The first step uses $k \le n^2$. The second uses $\omega \ge 2$. The last step follows from simple algebra.

The same rate holds at the finite accuracies used by the outer algorithm. An approximate gradient or projection adds its residual, times polynomially bounded factors, to Eq. (676). Allocate a fixed fraction of the target gap over all iterations and restarts. Inverse-polynomial residuals suffice, because the box diameter, L_0, c_g^{-1}, and the iteration count are polynomially bounded. Spectral clipping uses an interior safety margin so the numerical queries stay in a slightly enlarged fixed box with the same growth and smoothness bounds. All numerical gradient queries therefore remain positive definite.

12.4 Constructing directions and updating the optimizer

For at most n active coordinates, normalized matrix powers supply the direction required by Lemma 12.2. For larger active sets, we use a bounded spectral filter and control all four matrix responses.

Lemma 12.7 (A direction from normalized matrix powers)**.** *Let $A \in \mathbb{R}^{k\times k}$ be symmetric with $\|A\| \le M$ and $\lambda_{\min}(A) \le -4\gamma < 0$. A deterministic algorithm finds a nonzero v with $v^{\top}Av \le -\gamma\|v\|_2^2$ using $O(k^{\omega}\log(kM/\gamma))$ arithmetic operations and comparisons.*

Proof. Set $B := I - A/(2M) \succ 0$ and $q := (1+\gamma/M)/(1+2\gamma/M) < 1$. Choose a power of two N with

$$kq^{2N} \le \frac{\gamma}{2(M+2\gamma)},$$

and compute $Y := B^N$ up to a positive scalar factor by repeated squaring. Repeated doubling and scalar squaring find such N in $O(\log(kM/\gamma))$ steps, since a polynomially large N suffices.

An eigenvalue λ of A receives weight $(1-\lambda/(2M))^{2N}$. There is an eigenvalue at most -4γ. Compared with its weight, every eigenvalue greater than -2γ has weight at most q^{2N}. Hence

$$\frac{\operatorname{tr}[Y^{\top}AY]}{\operatorname{tr}[Y^{\top}Y]} \le -2\gamma + (M+2\gamma)kq^{2N} \le -\tfrac{3}{2}\gamma. \tag{679}$$

The first step separates eigenvalues at the threshold -2γ and uses their relative weights. The second uses the choice of N.

Compute AY and all nonzero-column Rayleigh quotients. At least one column is the required v. The multiplication costs $O(k^{\omega})$ and the remaining dot products cost $O(k^2)$.

Rescale each square by its largest absolute entry. After an initial rescaling, a symmetric matrix with largest absolute entry one has a square with largest absolute entry between one and k. Thus all rescaling divisors are nonzero and controlled. These rescalings use ordinary arithmetic operations and comparisons. □

Local correction and charging the cold solves. Successive optimizers are evaluated by local correction. A cold solve is used at initialization, after a failed boundary correction, and after each larger curvature move defined below.

Define $z := (X, Q)$, with $\operatorname{tr}[Q] = 1$, and let Π_0 be the orthogonal projection onto the trace-zero block-Hermitian space. Use the saddle function and its residual

$$\mathcal{L}_y(X, Q) := \rho \operatorname{tr}[U + V] + \operatorname{tr}[Q\mathcal{C}(X, y)] - \tau \operatorname{tr}[Q \log Q], \tag{680}$$

$$\mathcal{R}_y(z) := (\nabla_X \mathcal{L}_y, -\Pi_0 \nabla_Q \mathcal{L}_y). \tag{681}$$

The scalar reservoir is constant in this system. At a root put

$$E := \nabla^2_{XX} \mathcal{L}_y, \qquad B := \tau \Pi_0 D \log_Q \Pi_0, \qquad \mathcal{D} := \operatorname{diag}(E, B).$$

Here $\mathcal{D}$ denotes the block metric for the saddle system. The squared metric norm is $e(H)^2 + \mathcal{E}(K)^2$ for a primal variation H and a trace-zero dual variation K. The Jacobian has blocks $[E, J^*; -J, B]$, where $J := \Pi_0 D_X \mathcal{C}$. Consequently

$$\|\mathcal{D}^{1/2} (D\mathcal{R}_y)^{-1} \mathcal{D}^{1/2}\| \le 1. \tag{682}$$

Indeed the symmetric part of its normalized matrix is the identity. The skew-adjoint off-diagonal blocks cancel in this quadratic form.

The inverse terms supply a preconditioner for the primal Schur complement. This avoids assembling a matrix-space Hessian for a single correction or response.

Lemma 12.8 (Preconditioning the inner Hessian)**.** *At a corrected root define the positive definite operator* E_0 *by its quadratic form*

$$\begin{aligned}\langle H, \mathsf{E}_0 H\rangle := {} & 2 \operatorname{tr}[W U^{-1} H_U U^{-1} H_U U^{-1}] \\ & + 2 \operatorname{tr}[Z V^{-1} H_V V^{-1} H_V V^{-1}].\end{aligned} \tag{683}$$

For the inner objective ϕ*,*

$$\mathsf{E}_0 \preceq \nabla^2 \phi \preceq P \mathsf{E}_0. \tag{684}$$

Its inverse can be applied in $\widetilde{O}(n^3)$ *operations after* $\widetilde{O}(n^4)$ *spectral setup. These bounds require no lower bound on a positive eigenvalue of an individual input.*

Proof. Put $q_i(H) := \operatorname{tr}[B_i(U) H U^{-1} H]$. The functional Hessian inequality and Eq. (628) give

$$0 \le D^2 f_i(U)[H, H] \le (p-1) q_i(H), \qquad \iota \sum_i \beta_i z_i q_i(H) \le \operatorname{tr}[W U^{-1} H U^{-1} H U^{-1}].$$

Use the exchanged identities for V. The fixed-dual Hessian therefore satisfies $\mathsf{E}_0 \preceq E \preceq (p+1)\mathsf{E}_0/2$.

On the maintained sublevel, the positive coupling matrices $S_U := \iota \sum_i \beta_i f_i(U) C_i$ and S_V have bounded operator norm. Indeed, each constraint block is bounded above by the smoothed objective; subtract its signed matrix contribution and its positive inverse term. Write $S_U \preceq CI$ and put $K_U(H) := \iota \sum_i \beta_i D f_i(U)[H] C_i$. The operator convexity of the square gives

$$K_U(H)^2 \preceq C\iota \sum_i \beta_i \frac{D f_i(U)[H]^2}{f_i(U)} C_i.$$

For completeness, apply the unital positive map with weights $\iota\beta_i f_i(U)C_i/C$ to the scalar squares, adding the positive remainder $I - S_U/C$ with scalar value zero. The first differential inequality for the functional $Df_i(U)[H]^2 \le f_i(U)q_i(H)$ now implies

$$\operatorname{tr}[ZK_U(H)^2] \le C\iota \sum_i \beta_i z_i q_i(H) \le C\langle H_U, \mathsf{E}_{0,U} H_U\rangle/2.$$

Also $U \succeq aI$ gives

$$\operatorname{tr}[W(U^{-1}HU^{-1})^2] \le (2a)^{-1}\langle H, \mathsf{E}_{0,U} H\rangle.$$

Apply $(A+B)^2 \preceq 2A^2 + 2B^2$ to the two constraint derivatives and exchange U, V. It follows that $\operatorname{tr}[Q(D_X\mathcal{C}[H])^2] \le C'\langle H, \mathsf{E}_0 H\rangle$. The Gibbs covariance in direction K is nonnegative and at most $\tau^{-1}\operatorname{tr}[QK^2]$. The chain rule and $p, \tau^{-1} \le P$ prove Eq. (684).

For the inverse action, put $T := U^{-1/2}WU^{-1/2}$ and $\widehat{H} := U^{-1/2}HU^{-1/2}$. Then $\langle H, \mathsf{E}_{0,U} H\rangle = 2\operatorname{tr}[T\widehat{H}^2]$, and $\mathsf{E}_{0,U} H = G$ is equivalent to

$$T\widehat{H} + \widehat{H}T = U^{1/2}GU^{1/2}. \tag{685}$$

In an eigenbasis of T, divide entry ij by $t_i + t_j$ and undo the congruence. The root bounds control these denominators and all basis changes by polynomial factors. Do the same for V. Gap-independent spectral approximation to inverse-polynomial accuracy fits $\widetilde{O}(n^4)$ setup; subsequent basis changes and entrywise divisions cost $\widetilde{O}(n^3)$ per action. □

The normalized inverse bound controls a Newton correction near the current root. Relative derivative estimates determine its radius.

Lemma 12.9 (A preconditioned local correction)**.** *At every root on the maintained bounded sublevels there is a known fixed polynomial P in $\log(4n)$ such that the following holds. Starting within distance $1/(P\sqrt{n})$ of the root in its metric, a capped Newton correction reaches any prescribed inverse-polynomial residual in*

$$V(n,k) := \widetilde{O}(kn^\omega + n^4) \tag{686}$$

arithmetic operations. An accepted correction is certified against the true residual. An unsuccessful call has the same cost cap and returns failure.

Proof. At a root, the inverse part of E and Eq. (628) imply that, for $A = U, V$ and the corresponding block Q_A, $T_A := A^{-1/2}Q_A A^{-1/2}$ and $\widehat{H}_A := A^{-1/2}H_A A^{-1/2}$ satisfy

$$T_A \succeq \rho a I, \qquad e(H)^2 \ge 2\operatorname{tr}[T_A \widehat{H}_A^2]. \tag{687}$$

Thus both $\|\widehat{H}_A\|$ and the singular-value norm $\|T_A^{1/2}\widehat{H}_A T_A^{-1/2}\|$ are at most $P\sqrt{n}e(H)$. The latter estimate follows by bounding its squared Frobenius norm by $\lambda_{\min}(T_A)^{-1}\operatorname{tr}[T_A\widehat{H}_A^2]$. It is needed for the exterior factors of the inverse-Hessian words; these factors retain their matrix order. The entropy metric similarly gives

$$\|Q^{-1/2}KQ^{-1/2}\| \le P\sqrt{n}\mathcal{E}(K).$$

These inequalities keep all matrix blocks positive in the claimed ball.

We now bound the derivatives of the joint system. The inverse terms are bounded by the two relative primal norms just displayed. The logarithm terms use the resolvent formula for $D\log_Q$; a

relative dual variation bounds its quadratic form derivative by twice that relative norm. For the nonlinear coupling, its dual-weighted quadratic form obeys

$$\iota \sum_i \beta_i z_i q_i(H_U) \le \operatorname{tr}[W U^{-1} H_U U^{-1} H_U U^{-1}],$$

with the exchanged bound for V. Eq. (663) therefore bounds the cubic term from the functional by $P\sqrt{n}e(H)^3$. Polarization gives the corresponding trilinear bound in the E norm. For a mixed dual variation use $|\operatorname{tr}[KC_i]| \le \|Q_A^{-1/2} K_A Q_A^{-1/2}\| \operatorname{tr}[Q_A C_i]$. This gives the same $P\sqrt{n}$ bound for the mixed coupling and inverse terms. These derivatives are taken directly from the saddle function.

The bounds extend to the entire smaller ball with the metric at its center fixed. Expand inverses and logarithm resolvents in relative perturbations. For each functional expand $\operatorname{tr}[(C_i^{1/2}(X+H)C_i^{1/2})^p]$ and its positive pth root as in the proof of Lemma 12.4. The degree-j terms with $j \ge 2$ are bounded by the base quadratic energy times $(Cp\|X^{-1/2}HX^{-1/2}\|)^{j-2}$, with polynomial factors for any fixed number of derivatives. Polarization on the metric unit ball bounds their mixed derivatives. Their convergent sums, and the relative dual bound, give

$$\|\mathcal{D}_*^{-1/2}(D\mathcal{R}_y(z) - D\mathcal{R}_y(z_*))\mathcal{D}_*^{-1/2}\| \le P\sqrt{n}\|z - z_*\|_{\mathcal{D}_*}. \tag{688}$$

Decreasing the radius makes this at most $1/4$. The trace-word bounds apply throughout this ball. Eq. (682), the Neumann series, and the integral Taylor remainder now give the usual Newton error bound $\|z^+ - z_*\|_{\mathcal{D}_*} \le P\sqrt{n}\|z - z_*\|_{\mathcal{D}_*}^2$. Inexact evaluations add their prescribed smaller error floor. A fixed $O(\log n)$ iteration cap is more than sufficient.

All matrix entries and inverse bounds in this neighborhood are polynomial. Impose positivity, magnitude, and residual caps before accepting a call. The root lies strictly inside enlarged versions of the fixed primal box and the dual spectral bounds. Within these boxes, a sufficiently small Frobenius residual certifies root accuracy: the symmetric part of the saddle Jacobian is uniformly positive definite with an inverse-polynomial lower bound along the segment to the root. Pair the residual difference with that segment and use Cauchy–Schwarz. The same argument certifies accepted calls that started outside the guaranteed neighborhood.

Eliminate the entropy block. The primal Schur complement is $E + J^* B^{-1} J$, which equals $\nabla^2\phi$ at the root. Lemma 12.8 preconditions this matrix with a polylogarithmic condition number. The same comparison holds throughout the correction ball. To see this, use Eq. (688) to compare E, B with their root values and to bound $\|B_*^{-1/2}(J - J_*)E_*^{-1/2}\|$. These estimates bound the Schur complement above and below by polylogarithmic multiples of $\mathsf{E}_{0,*}$. The relative primal bound, its nonsymmetric similarity bound in Eq. (687), and relative dual comparability also compare the inverse-term preconditioner at the starting point with $\mathsf{E}_{0,*}$. It may therefore be frozen during the correction.

Apply preconditioned Richardson iteration with a step $1/P$ after rescaling the lower spectral bound to one. The preconditioned operator is self-adjoint in the preconditioner metric and has spectrum in $[1, P]$. Each iteration contracts the error by at most $1 - 1/P$, so a polylogarithmic number of actions attains the prescribed accuracy. No square root of the matrix-space preconditioner is required.

For one action, Eq. (696) applies each functional Hessian using $\widetilde{O}(n^\omega)$ operations. All coupling factors, Jacobians, and their adjoints therefore cost $\widetilde{O}(kn^\omega)$. The entropy inverse and Gibbs actions use divided differences in cached spectral bases, with trace centering. Together with the preconditioner they cost $\widetilde{O}(n^3)$ per action after $\widetilde{O}(n^4)$ setup. The Newton and linear-iteration counts are polylogarithmic, giving Eq. (686). Retain the positivity, magnitude, and true-residual checks

above. Outside the guaranteed ball the same finite action cap returns failure unless these checks certify success. This proves both the cost and the cap. □

A first-order predictor places the next optimizer within the correction radius. The step lengths from Lemma 12.5 control its residual.

Lemma 12.10 (Correction after an ordinary outer move)**.** *Every short gradient or curvature trial of Lemma 12.5 can be evaluated from the current corrected optimizer using a constant number of the local corrections in Lemma 12.9, up to logarithmic factors.*

Proof. Let $z'_h = (H, K)$ be the differentiated root and use the predictor $z_p(t) := z + tz'_h$. Put $T = P\theta$ in the curvature branch and use $T = P(\theta + ng^2)$ in the coordinate-gradient branch, increasing P when necessary. Eq. (665) and the coordinate forcing estimate above give $\|z'_h\|_{\mathcal{D}} \le \sqrt{T}$. Set $b_h := \|h\|_\infty$ and $R_* := b_h + P\sqrt{nT}$, enlarged to bound all the relative primal norms, including the nonsymmetric similarities in the proof of Lemma 12.9, and the relative dual norms.

The predictor residual has zero constant and linear terms. The same inverse, logarithm, and functional expansions give

$$\|\mathcal{R}_{y+th}(z_p(t))\|_{\mathcal{D}^{-1}} \le Pt^2 R_* \sqrt{T}. \tag{689}$$

We record the parameter estimates used here. The pure coupling forcing has coefficients $\beta''_i h_i^2$; its positive majorant has operator norm at most Pb_h^2 and dual trace at most $P\theta$. The entropy variance bound therefore gives dual energy norm at most $Pb_h\sqrt{\theta}$. Mixed coupling terms are bounded by the first differential inequality for the functional and the dual-weighted q_i estimate in the proof of Lemma 12.9. Primal mixed terms have the same bound by Cauchy–Schwarz. The signed matrix M is affine, so it has no second parameter derivative. These estimates prove Eq. (689) by integration.

The explicit coefficient changes in the Jacobian are at most $P|t|(b_h + \sqrt{n\theta})$ in the normalized metric. For the coupling-Jacobian block this follows by pairing a dual variation with $\sum_i \beta'_i h_i Df_i(H)C_i$, using its relative dual norm and weighted Cauchy–Schwarz. The resulting bound is $P\sqrt{n}\sqrt{\theta}$ per unit t, independently of the active count. The functional-Hessian coefficient change is bounded by Pb_h per unit t. Combining these with the Jacobian estimate in Lemma 12.9, a chord correction around the predictor is contractive provided

$$P|t|R_* \le c_0, \qquad P\sqrt{n}t^2 R_* \sqrt{T} \le c_0$$

for a sufficiently small absolute $c_0 > 0$. The corrected root is within $Pt^2 R_* \sqrt{T}$ of the predictor. Convexity identifies it with the optimizer of the new outer state.

In the curvature branch $T \le P$, $b_h \le Pn$, and $|t| \le 1/(Pn)$. Both requirements follow by enlarging the fixed P. In the gradient branch $b_h = 1$ and the prescribed step obeys $|t| \le 1/[P(1 + \sqrt{nT})]$, which gives both requirements as well. Shrinking these steps by fixed or polylogarithmic factors preserves the proved $\widetilde{O}(n^2)$ outer-step count. Separated residual and value tests absorb the predictor and coordinate-rounding errors. □

At a boundary event, an objective gap bounds the distance to the new root. This allows failed local corrections to be charged to potential decrease.

Lemma 12.11 (A gap sufficient for a warm correction)**.** *There is a known $h_* = 1/(Pn)$ such that, for a fixed outer state, if a primal point X in the maintained box satisfies $\phi(X) - \phi(X_*) \le h_*$, then $(X, Q(X))$ lies in the correction neighborhood of Lemma 12.9. Here $Q(X)$ is the Gibbs density of $\mathcal{C}(X, y)$.*

Proof. Let Q_* be the optimal Gibbs density and set $d := \phi(X) - \phi(X_*)$. The Gibbs variational identity and primal stationarity give the exact decomposition

$$d = \mathcal{L}_y(X, Q_*) - \mathcal{L}_y(X_*, Q_*) + \tau D(Q_* \| Q(X)),$$

where $D(P\|Q) := \operatorname{tr}[P(\log P - \log Q)]$. Both terms are nonnegative. Define $R := U_*^{-1/2} U U_*^{-1/2}$ and $T_* := U_*^{-1/2} W_* U_*^{-1/2} \succeq \rho a I$. The inverse Bregman remainder equals $\operatorname{tr}[T_*(R-I)R^{-1}(R-I)]$. If R has an eigenvalue outside $[1/2, 2]$, it is at least $\rho a/2$. Thus $d \le 1/(Pn)$ forces both primal relative matrices into this interval.

Likewise $\lambda_{\min}(Q_*) \ge \rho a^2$. If $Q(X)$ is not between $Q_*/2$ and $2Q_*$, a unit vector gives probabilities q and q_* with $q/q_* \notin [1/2, 2]$ and $q_* \ge \rho a^2$. The quantum relative entropy is at least the binary relative entropy of this measurement, which is at least $c_1 q_*$ for an absolute $c_1 > 0$. This measurement inequality also follows directly from the Gibbs variational formula and Golden–Thompson applied to a scalar multiple of the rank-one projection. Taking $d \le 1/(Pn)$ sufficiently small therefore forces $Q(X)$ between $Q_*/2$ and $2Q_*$.

Within these comparisons, the inverse remainder controls the inverse part of $e(X - X_*)^2$. The coupling part is at most P times that part, by the functional Hessian inequality and Eq. (628) at the root. The logarithm resolvent formula on the segment between the two densities gives

$$e(X - X_*)^2 + \mathcal{E}(Q(X) - Q_*)^2 \le Pd.$$

Choose the fixed polynomial in h_* larger than the polynomials in this bound and in the correction radius. Then Pd is smaller than that radius squared. This proves the claim. All comparisons allow fixed slack for finite-accuracy Gibbs evaluation. □

The boundary correction rule first tests the local solver and falls back to a cold solve on failure.

Lemma 12.12 (Number of cold solves)**.** *The boundary correction rule uses at most $\widetilde{O}(n)$ cold solves and at most n^2 local correction attempts, in addition to the initial cold solve.*

Proof. After a boundary event, keep the old primal point and compute the Gibbs density for the new state. First attempt the capped local correction. If it fails, use the cold solver of Eq. (678). On failure, Lemma 12.11 implies that the gap is at least $h_*/2$, with the factor two reserved for numerical errors.

Let F_- and F_+ be the exact optimized potentials immediately before and after this event, and define $e_j := 2\eta$ for a regular endpoint event and $e_j := 0$ for a deferred deletion. Evaluation at the old primal point, including the new reservoir, gives

$$0 \le d_j \le F_- - F_+ + e_j.$$

The coupling and reservoir deletion are nonincreasing at that point; the endpoint score repair is exactly the allowance e_j already charged above. Between boundary events the accepted outer moves decrease the potential. Since $F \ge 0$, $F_{\text{initial}} = O(1)$, and $\sum_j e_j = O(c)$, telescoping gives

$$\sum_j d_j \le F_{\text{initial}} + \sum_j e_j = O(1).$$

Consequently at most $\widetilde{O}(n)$ boundary events require a cold solve. There are still at most n^2 local correction attempts. A failed attempt is charged before its cold solve, and the initial cold solve is charged separately. □

Larger curvature moves. For $k > n$, Lemma 12.3 supplies a large negative subspace. We use it to control coordinates and matrix responses simultaneously.

Lemma 12.13 (A deterministic direction with bounded responses)**.** *At a retained light state satisfying Eq. (657), a deterministic construction finds h such that*

$$h^\top \Omega h = 1, \qquad h^\top \nabla^2 F h \le -1/P, \qquad R \le Pn/\sqrt{k}. \tag{690}$$

Including the Hessian and response columns, its cost is $\widetilde{O}(n^{2\omega})$.

Proof. Put $A := \Omega^{-1/2}\nabla^2 F \Omega^{-1/2}$. Choose $\gamma = 1/P$ so that at least k/P eigenvalues of A are at most $-\gamma$, with slack for approximation. A bounded symmetric filter Y can be computed with

$$\|Y\| \le 2, \qquad \operatorname{tr}[Y^2] \ge k/P,$$

and arbitrarily small inverse-polynomial leakage outside the band $(-\infty, -\gamma/2]$. Here is an explicit construction. Take a polynomial bound $M \ge \max\{\|A\|, \gamma\}$, start at $Z_0 = (A + 3\gamma I/4)/(2M)$, and iterate $Z_{j+1} = (3Z_j - Z_j^3)/2$. Scalar eigenvalues stay in $[-1, 1]$. Those corresponding to $A \le -\gamma$ converge to -1, and those corresponding to $A \ge -\gamma/2$ converge to 1. Error ϵ takes $O(\log(M/\gamma) + \log\log(1/\epsilon))$ iterations. Set $Y = (I - Z_j)/2$. The transition band already has negative curvature. The scalar derivative is at most $3/2$; the same bound holds for the Frobenius norm of the matrix derivative by divided differences. Thus inverse-polynomial rounding errors suffice throughout this construction. The cost is $\widetilde{O}(k^\omega)$.

Let $\mathcal{T}_1, \dots, \mathcal{T}_4$ be the real-linear maps from h to the four relative optimizer responses in the definition of R. Completing the square in Eq. (665) and using the primal and entropy lower bounds gives

$$\|\mathcal{T}_j \Omega^{-1/2} Y\|_{2\to F} \le P\sqrt{n}. \tag{691}$$

Tighten the filter leakage so that its contribution is smaller than one in this inequality. This is possible because the unfiltered response maps have polynomial norm. For $B_{ja} := \mathcal{T}_j \Omega^{-1/2} Y e_a$,

$$\Big\| \sum_a B_{ja}^2 \Big\| \le Pn^2. \tag{692}$$

Indeed, for a unit vector v, the real-linear map $H \mapsto Hv$ has norm at most one and rank at most $2n$. Its composition with the map in Eq. (691) has squared Frobenius norm at most $2n(P\sqrt{n})^2$. This equals $\sum_a \|B_{ja} v\|^2$. Also every row of $\Omega^{-1/2} Y$ has Euclidean norm at most Pn, by Eq. (659).

We give the deterministic selection explicitly. Choose $\pi = 1/P$ with $\operatorname{tr}[Y^2] \ge \pi k$, a scale $L = Pn$ bounding the preceding row norms, and $s = 1/(Pn)$ with $s^2 \|\sum_a B_{ja}^2\| \le 1$. For independent uniform signs ξ_a, used only to define conditional expectations, consider

$$\begin{aligned} \Phi(\xi) := \|Y\xi\|^2 &- \frac{\pi}{100} \sum_i \cosh((\Omega^{-1/2} Y \xi)_i / L) \\ &- \frac{\pi k}{1000 n} \sum_{j=1}^{4} \sum_{\epsilon = \pm 1} \operatorname{tr}[\exp\Big(\epsilon s \sum_a \xi_a B_{ja}\Big)]. \end{aligned} \tag{693}$$

Use exact conditional expectations for the first two terms. For each matrix penalty use the pessimistic estimator

$$\operatorname{tr}[\exp\Big(\epsilon s S_j + \frac{s^2}{2} \sum_{a \in J} B_{ja}^2\Big)], \qquad S_j := \sum_{a \notin J} \xi_a B_{ja},$$

where J is the remaining set of signs. Its children average to at most the current value, by the trace-exponential concavity argument used for Eq. (185). Initially the scalar penalties are at most $e^{1/2}$ each and the matrix penalties are at most $ne^{1/2}$ each. The initial lower estimator is therefore at least $\pi k/2$. Choose the child with the larger lower estimator at every sign. Errors at most $\pi/100$ per comparison retain final value at least $\pi k/4$.

Write $z = Y\xi$. Positivity of the penalties and $\|z\|^2 \le 4k$ give

$$\|z\|^2 \ge k/P, \qquad \|\Omega^{-1/2}z\|_\infty \le Pn, \qquad \|\mathcal{T}_j\Omega^{-1/2}z\| \le Pn.$$

Normalize $h = \Omega^{-1/2}z/\|z\|$. The negative band and the tightened leakage give Eq. (690).

Section 12.5 forms the Hessian and all response columns in $\widetilde{O}(n^{2\omega})$ work. Composing the four response arrays with Y has the same bound after padding dimensions to $O(n^2)$. The variance matrices cost $\widetilde{O}(kn^\omega)$, and scalar conditional choices cost $\widetilde{O}(k^2)$. Each matrix-exponential comparison costs $\widetilde{O}(n^\omega)$: positivity of the lower estimator bounds both signed exponentials, the remaining variance has bounded norm, and a single increment has bounded norm. Factorial Taylor evaluation therefore has polylogarithmic degree, as in the proof of Eq. (192). All k comparisons fit $\widetilde{O}(kn^\omega + k^2) = \widetilde{O}(n^{2\omega})$. All selections and their finite-accuracy checks are deterministic. □

Lemma 12.14 (Progress with many active coordinates)**.** *When $k > n$ and Eq. (657) holds, a curvature move gives certified potential decrease at least $k/(Pn^2)$. In a phase $K/2 < k \le K$ with $K > n$, there are at most $\widetilde{O}(n^2/K)$ such moves. Each accepted move is followed by a cold inner solve.*

Proof. Use Lemma 12.13 and a dyadic step within fixed factors of $\sqrt{k}/(Pn)$. Increase the fixed polynomial so that $|t|R$ preserves the endpoint strip and all positivity margins. Since $h^\top\Omega h = 1$, the cubic and quartic terms in Eq. (669) use less than half the negative quadratic term. Average the two orientations to cancel the linear term. One orientation decreases the potential by $k/(Pn^2)$. The proof of that equation bounds the primal objective at the explicit point $X + tH$. Evaluate these two primal values and compare with the old corrected value at a small fraction of the decrease margin. Thus acceptance certifies true decrease before the new optimizer is known. Run the cold solver after acceptance; no local-correction guarantee across this larger step is needed.

The initial potential is $O(1)$, it is nonnegative, and all boundary increases sum to $O(c)$. Every other accepted move decreases it. In the stated phase each larger curvature move decreases it by at least $K/(Pn^2)$ after changing P by a constant factor. This proves the count. Ordinary gradient moves and boundary events are counted separately. □

12.5 Running time and discrepancy guarantee

We use separate implementations according to the active count. First, at a fixed corrected root, all coordinate response systems share the same coupling factors, spectral bases, and preconditioner. Solving the k systems separately by Lemma 12.8 forms the outer Hessian and its response columns in

$$\widetilde{O}(n^4 + k^2n^\omega + kn^3). \tag{694}$$

There is only one spectral setup at this root. Each right-hand side takes a polylogarithmic number of actions costing $\widetilde{O}(kn^\omega + n^3)$. The mixed-block inner products cost $O(k^2n^2)$, already included. Symmetrize the computed outer Hessian and tighten the response errors so that its normalized operator error is below a fixed fraction of the curvature gap. All normalization and inverse bounds are polynomial. For $k \le n$, this cost, the matrix-power direction, and a constant number of local trial corrections together cost

$$\widetilde{O}(n^{\omega+2}). \tag{695}$$

Indeed $k^2 n^\omega \le n^{\omega+2}$ and $kn^3 \le n^4$.

For $k > n$, the larger curvature moves need all response columns. We assemble the sum of the functional Hessians in batches. For fixed positive X, define

$$F_{i,a} := (C_i X)^a C_i, \qquad f_i := \operatorname{tr}[(C_i X)^p]^{1/p}, \qquad G_i := f_i^{1-p} F_{i,p-1}.$$

Then $G_i = B_i(X)$ and direct differentiation gives

$$DB_i(X)[H] = f_i^{1-p} \sum_{a=0}^{p-2} F_{i,a} H F_{i,p-2-a} - \frac{p-1}{f_i} \operatorname{tr}[G_i H] G_i. \tag{696}$$

All factors cost $\widetilde{O}(kn^\omega)$ to form. For each a, the entries of a weighted sum of the first terms have the form

$$\mathcal{T}_{uv,st} = \sum_{i \in I} d_i (F_{i,a})_{us} (F_{i,p-2-a})_{tv}.$$

Stack the first factors in an n^2-by-k matrix with row index (u,s), and the second factors in a k-by-n^2 matrix with column index (t,v). One product gives all entries; reshuffle the indices afterwards. Padding $k \le n^2$ to n^2 costs $O(n^{2\omega})$. The rank-one terms are weighted Gram products with the same bound. There are only $O(p)$ batches. Complex Hermitian inputs are handled by restricting the resulting complex linear map to a real Hermitian basis; the sparse basis conversion costs $O(n^4)$.

The coupling Jacobians are similarly products of input-column and gradient-column matrices. The Gibbs derivative has its spectral divided-difference representation on matrix space. Forming it and composing it with the primal and outer Jacobians costs $\widetilde{O}(n^{2\omega})$. Thus the primal Hessian and the mixed and outer derivative blocks can be assembled in

$$\widetilde{O}(kn^\omega + n^{2\omega}) = \widetilde{O}(n^{\omega+2} + n^{2\omega}) = \widetilde{O}(n^{2\omega}). \tag{697}$$

The first step uses $k \le n^2$. The second uses $\omega \ge 2$.

At the optimizer, Eq. (670) bounds the primal Hessian below by $2c_g I$. Compute a symmetric approximation retaining a fixed fraction of this margin. Recursive block inversion and the outer Schur complement cost $\widetilde{O}(n^{2\omega})$; all principal Schur complements of the positive definite primal matrix are positive definite. Optimizer and derivative errors can be made small enough for the separated outer tests because the inverse norm is polynomially bounded. The same assembled blocks serve every response direction.

The two derivative implementations and the phase count give the following deterministic construction.

Theorem 12.15 (Formal version of Theorem 1.6)**.** *Let $A_1, \dots, A_m \in \mathbb{C}^{n \times n}$ be Hermitian positive semidefinite matrices with $\sum_{i=1}^m A_i = I_n$, $\|A_i\| \le \alpha$, and $\operatorname{rank}(A_i) \le r$. A deterministic algorithm returns one sign per original matrix satisfying*

$$\|\sum_{i=1}^m \sigma_i A_i\| \le \min\{1, \frac{4171}{1000}\sqrt{\alpha \log(2r)}\}$$

using $\widetilde{O}(mn^2 + n^7)$ real arithmetic operations and comparisons with ordinary dense matrix multiplication. More generally, for every admissible square matrix multiplication exponent $\omega \in [2,3]$, the operation bound is

$$\widetilde{O}(mn^2 + n^{\omega+4}).$$

The current bound $\omega < 2.371177$ of Dupont et al. [DEK+26] gives $\widetilde{O}(mn^2 + n^{6.371177})$, and hence the rounded bound $\widetilde{O}(mn^2 + n^{51/8})$. The tilde hides fixed powers of logarithms of $2mn$. Scalar functions, spectral computations, and optimization are evaluated to specified finite accuracies, with their evaluation costs included. The operation bound is independent of the inverse of an individual positive input eigenvalue.

Proof. Use the separated gradient test in Lemma 12.2. A violating coordinate takes the short step in Lemma 12.5, evaluated by Lemma 12.10. If all coordinates pass and $k > n$, use Lemma 12.14 and then a cold inner solve. If $k \le n$, form the small Hessian by Eq. (694), apply Lemma 12.7, and use the original short curvature step. Approximate the normalized Hessian to operator error below a fixed fraction of its negative gap, normalize the direction in the Ω metric, and certify the true decrease. The bounds $\Omega \succeq (Pn^2)^{-1} I$ and polynomial response norms make every required tolerance inverse polynomial.

At each boundary event first attempt the capped local correction of Lemma 12.9; on failure use the cold solver. The proof of Lemma 12.12 is unchanged, because the larger curvature moves also decrease the optimized potential. Thus there are at most $\widetilde{O}(n)$ failed boundary corrections requiring a cold solve. The cold solves after larger curvature moves are counted separately below.

All short moves and boundary attempts number $\widetilde{O}(n^2)$: short moves decrease the potential by at least $1/(Pn^2)$, larger moves only help, and the total boundary allowance remains $O(c)$. By Eq. (686), the coordinate-gradient moves and boundary attempts together cost

$$\widetilde{O}(n^2(n^{\omega+2} + n^4)) = \widetilde{O}(n^{\omega+4}).$$

By Eq. (695), all curvature moves with $k \le n$ also cost $\widetilde{O}(n^{\omega+4})$.

Partition the range $k > n$ into dyadic phases $K/2 < k \le K$. Lemma 12.14 gives $I_K = \widetilde{O}(n^2/K)$ larger curvature moves per phase. Their Hessians, response columns, filters, and deterministic selections cost

$$I_K \widetilde{O}(n^{2\omega}) = \widetilde{O}(n^{2\omega+2}/K) = \widetilde{O}(n^{2\omega+1}). \tag{698}$$

The last step uses $K > n$. Retaining the active count in Eq. (677), the cold solve after each such move costs $\widetilde{O}(n^{3/4}(Kn^{\omega} + n^4))$. Hence its phase total is

$$\widetilde{O}(n^{\omega+11/4} + n^{27/4}/K) = \widetilde{O}(n^{\omega+11/4} + n^{23/4}). \tag{699}$$

The initial cold solve and the $\widetilde{O}(n)$ cold solves after failed boundary corrections contribute at most $\widetilde{O}(n^{\omega+15/4} + n^{23/4})$.

There are $O(\log n)$ phases. Adding preprocessing and final rounding, the total work satisfies

$$\begin{aligned} \mathcal{W} &\le \widetilde{O}(mn^2 + n^6 + n^{\omega+4} + n^{2\omega+1} + n^{\omega+15/4} + n^{23/4}) \\ &= \widetilde{O}(mn^2 + n^{\omega+4}). \end{aligned} \tag{700}$$

The last equality uses $2 \le \omega \le 3$, which implies $6 \le \omega + 4$, $2\omega + 1 \le \omega + 4$, and $23/4 \le \omega + 4$. Every trial, failed capped correction, mandatory cold solve, and final rounding is included.

Ordinary multiplication gives $\widetilde{O}(mn^2 + n^7)$. The current bound $\omega < 2.371177$ of Dupont et al. [DEK+26] gives $\widetilde{O}(mn^2 + n^{6.371177})$.

We next verify the discrepancy coefficient. For each Schatten run, initialization, score errors, small-input restoration, and deferred rounding give normalized discrepancy $2 + O(c)$. The budget change preserves these estimates: its value at zero is one, its range is $[0, 1]$, and its endpoint value is zero. Its derivatives on the retained strip have the same polylogarithmic bounds needed by the relative remainder and finite-step arguments. Choose the absolute allowances, scale approximation,

and final norm-evaluation accuracy so that their combined relative loss is at most 10^{-5}. The selected run then has discrepancy at most

$$(1+10^{-5})2\sqrt{\iota r^{1/p}\alpha}.$$

All scale and norm evaluations can use inverse-polynomial accuracy, since $\alpha \geq 1/m$. The Schatten order is fixed by the rank, so only one Schatten run is required.

For $1 \leq r \leq 2$, Theorem 9.1 gives a signing with discrepancy at most $3.34963\sqrt{r\alpha}$, because $\operatorname{tr}[A_i] \leq r\alpha$ and $\sum_{i=1}^{m} \operatorname{tr}[A_i]A_i \preceq r\alpha I$. Its stated operation bound is dominated by the bound of the present theorem for every $\omega \in [2,3]$. In this range

$$(3.34963)^2 r/\log(2r) < 16.189.$$

The endpoint comparisons and the joint large-rank envelope in Section 11.5, using the rounded algorithmic trace constant in the first row, give the following upper bounds, rounded upward to six decimal places:

rank range	squared leading coefficient
$1 \leq r \leq 2$	16.189000
$r = 3$	17.396834
$4 \leq r \leq 32$	17.252797
$r \geq 33$	17.313519

(701)

In particular,

$$\begin{aligned} 17.396834(1+10^{-5})^2 &= 17.3971819384196834 \\ &< 17.397241 \\ &= (4171/1000)^2. \end{aligned} \tag{702}$$

The first step expands the numerical allowance. The second compares exact decimals, and the last evaluates the squared target coefficient.

This absorbs the chosen numerical and rounding allowances and proves coefficient 4.171 in Theorem 12.15 with the same arithmetic exponents. Every signing of the original isotropic PSD family has norm at most one. This proves the stated discrepancy and operation bounds. □

Algorithm 3 records the construction. All trial values and directions use the accuracy margins in Section 12.2.

13 Simultaneously spectrally thin spanning trees

A positive definite normalized Laplacian certifies connected support. We first record an all-block thinning argument and a deterministic construction. We then improve the existence bound by retaining a positive lower bound in only one weighting while controlling the upper bound in every weighting. The resulting spanning tree keeps all original edge weights.

13.1 Biased rounding at a fixed Schatten order

For an integer $\nu \geq 6$, define

$$\mathcal{C}_{r,\nu}^2 := \frac{137}{25}(\nu+1)r^{1/\nu}. \tag{703}$$

The even polynomial used in the budget is

$$\begin{aligned} f(t) := 1 &- 0.03382t^2 + 0.00994t^4 + 0.11342t^6 - 2.05274t^8 \\ &+ 7.10462t^{10} - 12.36505t^{12} + 10.74952t^{14} - 3.82102t^{16}. \end{aligned} \tag{704}$$

Algorithm 3 Deterministic rank-dependent signing

procedure RankSigning($(A_i)_{i=1}^m, \alpha, r$)
  $r \leftarrow \min\{r, n\}$.
  if $(4171/1000)\sqrt{\alpha \log(2r)} \geq 1$ **then**
    $\sigma_i \leftarrow 1$ for every i.
    return σ.
  end if
  if $r \leq 2$ **then**
    Compute σ by Theorem 9.1.
    return σ.
  end if
  $(p, \beta, \iota) \leftarrow$ the order, budget, and parameter in Eq. (624).
  $s \leftarrow \sqrt{\iota \alpha r^{1/p}}$; $C_i \leftarrow A_i/s$ for every i.
  Compute the reduced state in Eq. (626).
  Defer inputs with $\operatorname{tr}[C_i] \leq c/n^2$ and initialize Eq. (627).
  $\delta \leftarrow c^2/(100L)$; $\eta \leftarrow c/n^2$.
  while an active coordinate remains **do**
    $k \leftarrow$ the number of active coordinates.
    For $1 - |y_i| > \delta$, evaluate endpoint scores to absolute error at most η. ▷ Section 12.1
    if some $1 - |y_i| \leq \delta$ or an endpoint score is at least $-\eta$ **then**
      Update the active set and M as in Section 12.1.
      Attempt local correction; on failure, use Eq. (678). ▷ Lemma 12.12
    else
      Test $|g_i| \leq \sqrt{\omega_i}/(P\sqrt{n})$ for every active i. ▷ Lemma 12.2
      if a coordinate violates the test **then**
        Take its short gradient step using Lemma 12.10.
        Accept certified decrease at least $1/(Pn^2)$.
      else if $k > n$ **then**
        Compute the direction in Lemma 12.13.
        Test both primal predictors and accept decrease at least $k/(Pn^2)$.
        Correct the accepted state by the cold solver in Eq. (678).
      else
        Form the small Hessian by Eq. (694).
        Use Lemma 12.7 and the short curvature step.
        Apply local trial corrections and accept decrease at least $1/(Pn^2)$.
      end if
    end if
  end while
  Round deferred coordinates as in Section 12.1.
  Restore one sign per original input and certify its norm upper bound.
  return σ.
end procedure

Every decimal is an exact rational. Define the biased rounding factor by

$$h(p) := (p(1-p)/4)^{1/4}\sqrt{f(1-2p)}. \tag{705}$$

Use the budget and coupling

$$\beta(y) := \sqrt{1-y^2}\, f(y), \qquad \iota_\nu := \frac{137}{100}(\nu+1).$$

The modification is chosen to reduce the coupling at the biased starting points while retaining the curvature inequality on every light face.

Lemma 13.1 (The adjusted scalar budget)**.** *For every integer $\nu \geq 6$, this budget satisfies the positivity, endpoint, slope, and strict curvature requirements of the coordinate-dependent projection argument in Section 8.8.*

Proof. Set $t := |y|$ and $z := 1 - t^2$, and define polynomials

$$N := zf' - tf, \qquad A := f - zf', \qquad H_\nu := -f(zN' + tN) - \frac{N^2}{\nu - 1}.$$

Then $\beta' = N/\sqrt{z}$ for $y = t$ and

$$L_\nu := -\beta'' - \frac{(\beta')^2}{(\nu-1)\beta} = \frac{H_\nu}{z^{3/2} f}.$$

Direct exact Bernstein expansion on $[0,1]$ gives strictly positive coefficients for each of

$$f, \quad -N/t, \quad (f^2 - z)/t^2, \quad f + zf', \quad A, \quad H_6. \tag{706}$$

The quotients have their polynomial continuations at zero. Thus $z \leq \beta \leq 1$, β is positive and smooth in the open interval, and $\beta(\pm 1) = 0$. Also $-\beta'(t)(1-t)/\beta(t) < 1$, giving positive response coefficients. The budget is even, so exchanging the two matrix variables handles negative y.

Define the lightness parameters by

$$a := \frac{\iota_\nu \beta u}{1+t}, \qquad b := \frac{\iota_\nu \beta v}{1-t}, \qquad 0 < a, b < 1,$$

where $u = f_i(U)$ and $v = f_i(V)$ are the local functional values for the coordinate under consideration. With $\ell := -\beta'/\beta \geq 0$, the response coefficients and coercivity are

$$\mathfrak{p} := 1 - \ell(1-t)b, \quad \mathfrak{q} := 1 + \ell(1+t)a, \quad \Gamma := \frac{\iota_\nu f}{\sqrt{z}\, ab}.$$

Eq. (377), with the energy-retention parameter $\mu = 0$ and $\kappa(\theta(\theta+1)) = \theta^2(\theta+1)/(\theta^3-1)$, reduces strict negative curvature to the scalar inequality

$$\iota_\nu L_\nu > (\nu+1)\kappa(\Gamma)\mathfrak{p}\mathfrak{q}. \tag{707}$$

Both factors on the right increase with a at fixed t, b: $\mathfrak{q}$ increases, whereas Γ decreases and κ is decreasing. We may therefore set $a = 1$. Set $x := 1/\theta$, where $\Gamma = \theta(\theta+1)$. Since $\beta \geq z$,

$$\Gamma \geq \iota_\nu \geq \frac{959}{100} > \frac{35}{4}, \qquad 0 < x < \frac{2}{5}.$$

At $a = 1$, the condition $b \leq 1$ becomes

$$D_\nu(t,x) := \iota_\nu^2 f(t)^2 x^4 - z(1+x)^2 \leq 0. \tag{708}$$

Substitution and multiplication by positive factors reduce Eq. (707) to

$$W_\nu := \iota_\nu[H_\nu(1-x^3) - (\nu+1)ANx^2] > (\nu+1)A(1+x)\sqrt{z}(1+t).$$

The left side is positive by Eq. (706). For the base order $\nu = 6$ define the polynomial

$$P_6 := W_6^2 - 49A^2(1+x)^2 z(1+t)^2. \tag{709}$$

Its bidegree is $(68, 6)$. The exact certificate partitions $[0,1] \times [0, 2/5]$ into 52 dyadic rectangles. On 33 rectangles, all tensor Bernstein coefficients of P_6 are strictly positive. On the remaining 19, all coefficients of D_6 are strictly positive, so those rectangles contain no admissible point. This proves Eq. (707) at order six over the entire admissible domain.

For $\nu \geq 6$, we have $H_\nu \geq H_6$, $N \leq 0$, and $x < 1$; hence

$$\begin{aligned}
\frac{W_\nu}{\nu+1} &= \frac{137}{100}[H_\nu(1-x^3) - (\nu+1)ANx^2] \\
&\geq \frac{137}{100}[H_6(1-x^3) - (\nu+1)ANx^2] \\
&\geq \frac{137}{100}[H_6(1-x^3) - 7ANx^2] \\
&= \frac{W_6}{7}.
\end{aligned} \tag{710}$$

The first step substitutes ι_ν. The second uses $H_\nu \geq H_6$ and $x < 1$. The third uses $A > 0$, $N \leq 0$, and $\nu + 1 \geq 7$. The last step is the definition of W_6.

Moreover $D_\nu \leq 0$ implies $D_6 \leq 0$. The base certificate therefore proves the same strict inequality for every integer $\nu \geq 6$.

Appendix B.5 gives the subdivision rule and the exact Bernstein tests. The terminal coefficient tests use exact affine monomial substitution. □

Starting the signing potential at a constant fractional vector gives rounding around an arbitrary mean pS.

Lemma 13.2 (Biased subisotropic rounding)**.** *Let $X_i \succeq 0$, $\|X_i\| \leq \delta$, $\operatorname{rank}(X_i) \leq r$, and $S = \sum_i X_i \preceq bI$, where $b, \delta > 0$. For every $0 < p < 1$ and integer $\nu \geq 6$, there is a subset $\mathcal{J}$ satisfying*

$$\|\sum_{i\in\mathcal{J}} X_i - pS\| \leq h(p)\mathcal{C}_{r,\nu}\sqrt{\delta b}. \tag{711}$$

In particular, $p = 1/2$ gives signs $\sigma_i \in \{-1, 1\}$ with $\|\sum_i \sigma_i X_i\| \leq \mathcal{C}_{r,\nu}\sqrt{\delta b}$.

Proof. Set $b_{\text{bias}} := 2p - 1$ and $\iota := \iota_\nu$. In the order-ν potential replace $M(y) := \sum_i y_i X_i$ by

$$M_{b_{\text{bias}}}(y) := \sum_i (y_i - b_{\text{bias}})X_i.$$

Keep the functional

$$f_i(X) := \|X_i^{1/2} X X_i^{1/2}\|_{S_\nu}, \qquad \mathcal{F}_y(X) := \iota \sum_i \beta(y_i) f_i(X) X_i,$$

and define

$$P_{\rho,b_{\text{bias}}}(y) := \min_{z\in\mathbb{R},\ U,V\succ 0} \{z + \rho \operatorname{tr}[U+V]\}$$

subject to

$$U^{-1} + M_{b_{\mathrm{bias}}}(y) + \mathcal{F}_y(V) \preceq zI,$$
$$V^{-1} - M_{b_{\mathrm{bias}}}(y) + \mathcal{F}_y(U) \preceq zI.$$

The fixed offset $-b_{\mathrm{bias}}S$ changes neither the endpoint increments nor any derivative of M with respect to the coordinates. The KKT equations for U, V, functional differential estimates, response covariance, and strict negative curvature at stationary light points are therefore unchanged. In particular, the order-ν certificate from Lemma 13.1 still applies. Existence and regularity also persist: $M_{b_{\mathrm{bias}}}$ is bounded on the cube, and the same inverse and trace bounds give compact optimizer sublevels.

More explicitly, changing y_i to 1 changes the first constraint by $(1 - y_i - \iota\beta_i f_i(V))X_i$ and the second by $(-(1 - y_i) - \iota\beta_i f_i(U))X_i$. These are the same endpoint tests as before, and the tests for -1 are unchanged as well. Thus a global minimum with the fewest active coordinates must be a vertex: an eligible endpoint removes an active coordinate without increasing the value; otherwise strict negative curvature excludes a stationary minimum on a positive-dimensional face. It applies to every starting point in the cube.

Let d be the matrix order. At $y = b_{\mathrm{bias}}\mathbf{1}$, we have $M_{b_{\mathrm{bias}}}(y) = 0$ and $\beta(y_i) = \sqrt{1 - {b_{\mathrm{bias}}}^2} f(b_{\mathrm{bias}})$. The subisotropic estimate

$$\sum_i f_i(I)X_i \preceq r^{1/\nu}\delta bI$$

and the trial $U = V = \ell I$ give

$$P_{\rho,b_{\mathrm{bias}}}(b_{\mathrm{bias}}\mathbf{1}) \le 2\sqrt{\iota\beta(b_{\mathrm{bias}})r^{1/\nu}\delta b + 2d\rho},$$

A minimizing vertex σ has no larger value and satisfies $\|\sum_i \sigma_i X_i - b_{\mathrm{bias}}S\| \le P_{\rho,b_{\mathrm{bias}}}(\sigma)$. Let $\rho \downarrow 0$ and pass to a constant-signing subsequence. For $\mathcal{J} := \{i : \sigma_i = 1\}$,

$$\sum_{i\in\mathcal{J}} X_i - pS = \frac{1}{2}(\sum_i \sigma_i X_i - b_{\mathrm{bias}}S).$$

Since $\sqrt{\beta(b_{\mathrm{bias}})}/2 = h(p)$ and $\mathcal{C}_{r,\nu} = 2\sqrt{\iota r^{1/\nu}}$, this proves Eq. (711). □

The upper and lower rounding errors can also be chosen separately, provided their product is fixed.

Lemma 13.3 (Asymmetric subisotropic rounding)**.** *Under the hypotheses of Lemma 13.2, fix* $0 < p < 1$*, an integer* $\nu \ge 6$*, and* $\ell, v > 0$ *such that*

$$\ell v = h(p)^2\mathcal{C}_{r,\nu}^2\delta b. \tag{712}$$

There is a subset $\mathcal{J}$ *satisfying*

$$-\ell I \preceq \sum_{i\in\mathcal{J}} X_i - pS \preceq vI.$$

Proof. Set $b_{\mathrm{bias}} := 2p - 1$, $\Delta := v - \ell$, and $M_{b_{\mathrm{bias}},\Delta}(y) := \sum_i (y_i - b_{\mathrm{bias}})X_i - \Delta I$. Use the potential from the proof of Lemma 13.2 with $M_{b_{\mathrm{bias}},\Delta}$ in place of $M_{b_{\mathrm{bias}}}$. The additional constant offset $-\Delta I$ changes no coordinate derivative or endpoint increment. The KKT equations for U, V, functional estimates, response covariance, and strict negative curvature at stationary light points are unchanged. The signed matrix is still bounded on the cube, so the same inverse and trace

bounds give compact optimizer sublevels. The endpoint-or-negative-curvature argument therefore gives a minimizing vertex for every $\rho > 0$.

At $y = b_{\mathrm{bias}}\mathbf{1}$, the signed matrix is $-\Delta I$. Set $g := \iota_\nu \beta(b_{\mathrm{bias}}) r^{1/\nu} \delta b = \ell v$ and use

$$U = v^{-1} I, \qquad V = \ell^{-1} I, \qquad z = \ell + v.$$

The subisotropic coupling estimate bounds the first constraint by $v - \Delta + g/\ell = \ell + v$, and the second by $\ell + \Delta + g/v = \ell + v$. If d is the matrix order, this gives

$$P_{\rho, b_{\mathrm{bias}}, \Delta}(b_{\mathrm{bias}}\mathbf{1}) \le \ell + v + d\rho(\ell^{-1} + v^{-1}).$$

At a minimizing vertex the coupling terms vanish. Let $\rho \downarrow 0$ along a subsequence with a fixed signing and put $\mathcal{J} := \{i : \sigma_i = 1\}$. The resulting bound is

$$\|2(\sum_{i \in \mathcal{J}} X_i - pS) - (v - \ell)I\| \le \ell + v.$$

Its two Loewner inequalities give the conclusion. □

13.2 Biased thinning with a positive lower bound

Fix ν, r and set $s_{\mathrm{thin}} := \mathcal{C}_{r,\nu}^2 \alpha$. Suppose a retained family of the original PSD matrices has sum S, all norms are at most α, all ranks are at most r, and

$$\lambda s_{\mathrm{thin}} I \preceq S \preceq \mu s_{\mathrm{thin}} I, \qquad 0 < \lambda \le \mu.$$

A regularized congruence balances the lower and upper spectral errors in one rounding step.

Lemma 13.4 (Biased regularized thinning)**.** *Let $0 < p \le 1/2$, $h \ge h(p)$, and*

$$\sqrt{\lambda} \le c \le \sqrt{\mu}, \qquad hc < p\lambda.$$

There is a retained subfamily with sum S' satisfying

$$(p\lambda - hc) s_{\mathrm{thin}} I \preceq S' \preceq (p\mu + h\mu/c) s_{\mathrm{thin}} I. \tag{713}$$

In particular $S' \succ 0$.

Proof. For $c < \sqrt{\mu}$, define

$$t := \frac{\mu(c^2 - \lambda)}{\mu - c^2}, \qquad C_i := (S + t s_{\mathrm{thin}} I)^{-1/2} X_i (S + t s_{\mathrm{thin}} I)^{-1/2}.$$

Then $t \ge 0$, $\|C_i\| \le \alpha/[s_{\mathrm{thin}}(\lambda + t)]$, and $\sum_i C_i \preceq \mu/(\mu + t) I$. Apply Lemma 13.2 and undo the congruence. Set $E := S' - pS$ and $d := \sqrt{\mu/((\lambda + t)(\mu + t))}$. The rounding error satisfies

$$-hd(S + t s_{\mathrm{thin}} I) \preceq E \preceq hd(S + t s_{\mathrm{thin}} I).$$

The identities $d(\lambda + t) = c$ and $d(\mu + t) = \mu/c$ give the desired endpoint bounds. The lower affine function is increasing because $hd = hc/(\lambda + t) < p\lambda/(\lambda + t) \le p$; the upper one is increasing as well. If $c = \sqrt{\mu}$, apply Lemma 13.2 directly to the retained matrices to obtain $\|E\| \le h s_{\mathrm{thin}} \sqrt{\mu}$. This proves the same bounds, including the case $\lambda = \mu$. □

Asymmetric rounding gives another step in which the lower loss can be chosen freely. The regularized step above will also be used with the constructive signing bound in Section 13.4.

Lemma 13.5 (Asymmetric biased thinning)**.** *Let $0 < p < 1$ and $0 < c < p\lambda$. There is a retained subfamily with sum S' satisfying*

$$(p\lambda - c)s_{\mathrm{thin}} I \preceq S' \preceq \mu\left(p + \frac{h(p)^2}{c}\right) s_{\mathrm{thin}} I. \tag{714}$$

In particular $S' \succ 0$.

Proof. Whiten by $S^{-1/2}$. The resulting inputs sum to identity, have norms at most $\alpha/(\lambda s_{\mathrm{thin}})$, and retain their ranks. Apply Lemma 13.3 with $\ell := c/\lambda$ and $v := h(p)^2/c$. Their product is $h(p)^2/\lambda$, as required by Eq. (712) for the normalized inputs. Undoing the congruence gives

$$(p - c/\lambda)S \preceq S' \preceq (p + h(p)^2/c)S.$$

Both coefficients are positive because $c < p\lambda$. The assumed bounds on S now prove Eq. (714). □

13.3 Exact verification of the scalar recurrence

Lemma 13.5 gives the scalar recurrence

$$\lambda' := p\lambda - c, \qquad \mu' := \mu\left(p + \frac{h(p)^2}{c}\right), \qquad 0 < p < 1, \quad 0 < c < p\lambda. \tag{715}$$

The curvature condition in Lemma 13.1 implies $-\beta'' - (\beta')^2/((\nu-1)\beta) > 0$. Thus β is concave and vanishes at both endpoints. Since $h(p)^2 = \beta(2p-1)/4$, the function

$$p \longmapsto h(p)^2/p \tag{716}$$

is nonincreasing on $(0, 1)$. This property lets one finite path control all smaller initial scales.

Lemma 13.6 (Transfer from one scalar path)**.** *Suppose a path of k legal steps of Eq. (715), allowing outward rounding, starts at $(\lambda_0, \mu_0) = (U, U)$ and ends with $\lambda_k > 0$ and $\mu_k \le C$. Then every initial pair (u, u) with $0 < u \le U$ has a path of at most k steps ending with positive lower bound and upper bound at most C.*

Proof. Write the saved controls as p_j, c_j, so that

$$0 < \lambda_{j+1} \le p_j\lambda_j - c_j < \lambda_j, \qquad \mu_{j+1} \ge \mu_j(p_j + h(p_j)^2/c_j).$$

Each state satisfies $\lambda_j \le \mu_j$. If $u \le \mu_k$, take no step. Otherwise $u > \mu_k \ge \lambda_k$, so choose j with $\lambda_{j+1} \le u \le \lambda_j$. If $u \le \mu_{j+1}$, the sandwich $(\lambda_{j+1}, \mu_{j+1})$ already contains uI; follow its saved suffix.

If $u > \mu_{j+1}$, set $m := p_j\lambda_j$ and $\widetilde{p} := m/u$. Since $\mu_{j+1} \ge p_j\mu_j \ge m$, we have $p_j \le \widetilde{p} < 1$. Apply one step with $\widetilde{p}$ and c_j. Its lower bound is $m - c_j \ge \lambda_{j+1} > 0$. Its upper bound satisfies

$$m + \frac{h(\widetilde{p})^2 u}{c_j} = m + \frac{m}{c_j}\frac{h(\widetilde{p})^2}{\widetilde{p}} \le m + \frac{h(p_j)^2\lambda_j}{c_j} \le \mu_{j+1}. \tag{717}$$

The first step substitutes $\widetilde{p} = m/u$. The second uses Eq. (716) and $m = p_j\lambda_j$. The third uses $\lambda_j \le \mu_j$ and the saved upper update.

Follow the saved suffix from this sandwich. Every step is legal by Lemma 13.5, and there are at most k steps in total. The adjusted probability may exceed $1/2$; the rounding lemma permits every $0 < p < 1$. □

An eight-step rational path controls bounded initial scales. Ordinary whitened halving reduces larger scales to this range.

Lemma 13.7 (The scalar thinning bound)**.** *For every $u > 0$, starting from $(\lambda, \mu) = (u, u)$, a finite sequence of legal steps of Eq. (715) gives $\lambda > 0$ and $\mu < 29/5$. The number of rounding selections is $O(\log(2 + u))$.*

Proof. Set $T_0 := 2^{28}$, $Q := 2^{48}$, and $D := 2^{24}$. Start the integer state at $A_0 = B_0 = T_0 Q$. A state (A, B) represents the conservative sandwich $(A/Q)s_{\text{thin}} I \preceq S \preceq (B/Q)s_{\text{thin}} I$. For each pair (P, C) in Table 3, put $p := P/D$ and $c := C/Q$. Let H be the least positive integer satisfying

$$\frac{H^2}{Q^2} \geq \frac{p(1-p)}{4} f(1-2p)^2. \tag{718}$$

Since $f > 0$, this gives $H/Q \geq h(p)^2$. Check

$$0 < P < D, \qquad 0 < C < PA/D, \tag{719}$$

and update

$$A' := \left\lfloor \frac{PA}{D} - C \right\rfloor, \qquad B' := \left\lceil B\left(\frac{P}{D} + \frac{H}{C}\right) \right\rceil. \tag{720}$$

These are outward-rounded versions of Eq. (715). Appendix B.6 gives the eight controls and their exact integer verification. Every lower bound is positive, and the terminal state is

$$A_8 = 2471088, \qquad B_8 = 1631855358146481. \tag{721}$$

In particular $\mu_8 = B_8/Q < 5.797515$. Lemma 13.6 therefore covers every $0 < u \leq T_0$ with at most eight selections.

For $u_0 > T_0$, first use $p = 1/2$ and $c = \sqrt{\lambda}/2$, until the first k with $u_k \leq T_0$. Set $u_j := \lambda_j$, $v_j := \mu_j$, and $x_j := u_j^{-1/2}$. Since $h(1/2)^2 = 1/4$, the recurrence becomes

$$u_{j+1} = (u_j - \sqrt{u_j})/2, \quad v_{j+1} = v_j(1 + u_j^{-1/2})/2, \quad \frac{v_k}{u_k} = \prod_{j<k} \frac{1+x_j}{1-x_j}.$$

Since $u_{j+1} < u_j/2$ and $u_{k-1} > T_0$,

$$\sum_{j<k} x_j < \frac{2+\sqrt{2}}{\sqrt{T_0}}, \qquad x_j^2 < 1/T_0.$$

Consequently

$$\begin{aligned} \log(v_k/u_k) &< \frac{2(2+\sqrt{2})}{\sqrt{T_0}(1-1/T_0)} \\ &< \frac{1}{2389}. \end{aligned} \tag{722}$$

The first step sums $\log((1+x_j)/(1-x_j)) \leq 2x_j/(1-x_j^2)$. The second uses $T_0 = 2^{28}$ and $\sqrt{2} < 10/7$. Since $e^t < 1/(1-t)$ for $0 < t < 1$, this gives $v_k/u_k < 2389/2388$.

Also $u_k > (T_0 - \sqrt{T_0})/2 > 1$. Every fixed control in Eq. (715) is independent of μ in its lower update and linear in μ in its upper update. Hence the path for (u_k, u_k) transfers to $(u_k, \vartheta u_k)$,

where $\vartheta := v_k/u_k$, by multiplying every upper bound by ϑ. The lower bounds and all legality tests remain unchanged. The final upper bound is strictly below

$$\frac{1631855358146481}{281474976710656}\frac{2389}{2388} < \frac{29}{5}. \tag{723}$$

This is an exact rational comparison. The finite path uses at most eight selections, and the preceding tail uses $O(\log(2+u))$ halvings. □

The scalar recurrence and asymmetric rounding produce a subfamily with positive definite sum. Positivity will ensure that its graph support is connected.

Lemma 13.8 (A small positive definite subfamily)**.** *Let $X_1, \dots, X_{m_{\mathrm{in}}} \succeq 0$ act on a space of dimension $d \geq 1$, with $\sum_i X_i = I_d$, $\|X_i\| \leq \alpha$, and $\operatorname{rank}(X_i) \leq r$. For each integer $\nu \geq 6$ there is a subfamily $\mathcal{J} \subseteq [m_{\mathrm{in}}]$ with*

$$0 \prec \sum_{i\in\mathcal{J}} X_i \preceq \min\{1, \tfrac{29}{5}\tfrac{137}{25}(\nu+1)r^{1/\nu}\alpha\}I. \tag{724}$$

It uses $O(\log(2m_{\mathrm{in}}))$ rounding selections on congruences of original subfamilies.

Proof. Set $s_{\mathrm{thin}} := \mathcal{C}_{r,\nu}^2\alpha$. Lemma 13.7 applies at $(\lambda,\mu) = (1/s_{\mathrm{thin}}, 1/s_{\mathrm{thin}})$. Apply Lemma 13.5 at each selected step. Strict positivity is preserved. The final upper bound is $(29/5)s_{\mathrm{thin}}$, and the subfamily sum is also at most I. Finally $1 \leq m_{\mathrm{in}}\alpha$ and $\mathcal{C}_{r,\nu} \geq 1$ imply $1/s_{\mathrm{thin}} \leq m_{\mathrm{in}}$. □

13.4 Deterministic construction and running time

The deterministic construction uses ordinary signings in the halving recurrence. Set $C_0 := 384/25$, $Q_0 := 2^{48}$, and $\eta := 10^{-8}$.

Constructive subisotropic signing. Lemma 14.1 supplies a deterministic signing with coefficient 4.171. Its proof also gives a finite-accuracy implementation when $S = \sum_i X_i \preceq bI$, $\|X_i\| \leq \delta$, and $\operatorname{rank}(X_i) \leq r$. To normalize the inputs, proceed as follows. If $S = 0$, all inputs vanish. Otherwise compute $\|S\| \leq b_0 \leq (1+\eta)\|S\|$, divide the inputs by b_0, and use $\delta_0 := \max\{\delta/b_0, 1/m_{\mathrm{in}}\}$ for m_{in} inputs. Since $\|S\| \leq m_{\mathrm{in}}\delta$, we have $\delta_0 \leq (1+\eta)\delta/b_0$. The normalized inputs X_i/b_0 have sum at most I, individual norm at most δ_0, and rank at most r. Since $\delta_0 \geq 1/m_{\mathrm{in}}$, the algorithmic clause of Lemma 14.1 applies directly. After rescaling, its signing has norm at most $4.171(1+\eta)\sqrt{\delta b\log(2r)}$, because

$$b_0\sqrt{\delta_0} \leq \sqrt{(1+\eta)\delta b_0} \leq (1+\eta)\sqrt{\delta b}.$$

The same proof retains the fast multiplication bound of Theorem 12.15.

In the applications below all required spectral scales have polynomial upper and inverse bounds. Approximate congruences and enclosing scalar scales can thus be computed to a sufficiently small relative error. Reserve the coefficient 4.171001 for these evaluations. To see that this reservation suffices, let R approximate $X^{-1/2}$ with $\|RX^{1/2} - I\| \leq \eta$, using a symmetric function of X. Apply the signing routine to the PSD matrices RX_iR. Their norms and sum norm increase by at most $(1+\eta)^2$ relative to the exact congruence; undoing the approximate congruence costs at most $(1-\eta)^{-2}$. Together with the preceding normalization, the coefficient is at most $4.171(1+\eta)^3/(1-\eta)^2 < 4.171001$. All products defining these inputs are performed in the real arithmetic model, so the common congruence preserves positivity and rank exactly.

Halving with the constructive signing constant. Set $\mathcal{C}_{\mathrm{alg}} := 4.171001\sqrt{\log(2k)}$ and $\alpha_{\mathrm{alg}} := \mathcal{C}_{\mathrm{alg}}^2\alpha_{\mathrm{edge}}$, where α_{edge} bounds the normalized edge matrices. For $\lambda\alpha_{\mathrm{alg}}I \preceq S \preceq \mu\alpha_{\mathrm{alg}}I$ and $\sqrt{\lambda} \le c \le \sqrt{\mu}$, $c < \lambda$, the proof of Lemma 13.4, with $p = h = 1/2$, uses only a signing bound $\mathcal{C}_{\mathrm{alg}}\sqrt{\delta\mu}$ and gives

$$\lambda' = (\lambda - c)/2, \qquad \mu' = (\mu + \mu/c)/2. \tag{725}$$

Thus this step is constructive. The choice $c = \sqrt{\lambda}$ is a whitened step. With $\eta_0 := 1/1024$, a final choice $c = \min\{\sqrt{\mu}, \lambda - \eta_0\}$, when $\lambda - \eta_0 > \sqrt{\lambda}$, gives $\lambda' \ge \eta_0/2$ and

$$\mu' \le \frac{1}{2}(\mu + \max\{\sqrt{\mu}, \mu/(\lambda - \eta_0)\}). \tag{726}$$

We use the halving certificate specified in Appendix B.4. The following scalar lemma gives the bound needed for the constructive constant.

Lemma 13.9 (The scalar halving bound)**.** *Starting from $(\lambda, \mu) = (u, u)$ for any $u > 0$, a sequence of legal whitened steps followed by at most four regularized steps gives $\lambda > 0$ and $\mu < 384/25$, unless the starting μ already satisfies this bound. The number of steps is $O(\log(2 + u))$.*

Proof. We first certify $1 \le u \le T_0 = 2^{28}$ using exact integer arithmetic, then give an analytic tail bound. Set $Q := 2^{48}$. Starting at $l = Q$, the adjoining closed initial intervals are

$$[l/Q, h/Q], \qquad h = \min\{\lceil 2049l/2048\rceil, T_0Q\};$$

replace l by h after each interval. There are exactly 39,758 intervals, and their union is $[1, T_0]$.

An integer state (A, B) certifies the sandwich $(A/Q)\alpha_{\mathrm{alg}}I \preceq S \preceq (B/Q)\alpha_{\mathrm{alg}}I$. At each step we may use the conservative parameters $\lambda = A/Q$ and $\mu = B/Q$ in the halving lemma, even when the true spectral bounds are tighter. Initialize the state at (l, h). For a whitened step put

$$\begin{aligned} S_- &= \lfloor\sqrt{AQ}\rfloor, & S_+ &= \lceil\sqrt{AQ}\rceil, \\ A' &= \lfloor(A - S_+)/2\rfloor, & E &= \lceil Q^2/S_-\rceil, \\ B' &= \lceil\frac{B(Q + E)}{2Q}\rceil. \end{aligned}$$

Only perform this step when $A > Q$ and $A' > 0$. The monotonicity of $(\lambda - \sqrt{\lambda})/2$ for $\lambda \ge 1$ and the displayed rounding directions prove that the state remains valid.

For a rational regularized control $c := C/Q$, check

$$AQ \le C^2 \le BQ, \qquad 0 < C < A,$$

and update

$$A' = \lfloor(A - C)/2\rfloor, \qquad B' = \lceil\frac{B + \lceil BQ/C\rceil}{2}\rceil, \tag{727}$$

requiring $A' > 0$. These are direct rational enclosures of Eq. (725). A final step of Eq. (726) is available when $A - Q/1024 > \lceil\sqrt{AQ}\rceil$ and has the integer upper bound

$$B_{\mathrm{final}} = \lceil\frac{1}{2}(B + \max\{\lceil\sqrt{BQ}\rceil, \lceil\frac{BQ}{A - Q/1024}\rceil\})\rceil. \tag{728}$$

Algorithm 6 and the interval traversal in Appendix B.4 generate the controls from the initial intervals. The integer updates in that appendix check coverage, legality of every step, all updates, and the strict terminal bound

$$20B_{\mathrm{terminal}} < 307Q. \tag{729}$$

All 39,758 intervals pass. The largest resulting upper bound is strictly below $307/20 = 15.350$. At most 22 whitened steps and four regularized steps occur in each finite-interval construction. The generation and all acceptance tests use integer square roots and rational inequalities.

Now let $u_0 > T_0$, and apply exact whitened steps until the first h with $u_h \le T_0$, where $u_j = \lambda_j$. With $v_j = \mu_j$ and $x_j = u_j^{-1/2}$,

$$\frac{v_h}{u_h} = \prod_{j<h} \frac{1+x_j}{1-x_j}, \qquad \sum_{j<h} x_j < \frac{2+\sqrt{2}}{\sqrt{T_0}}, \qquad x_j^2 < 1/T_0.$$

Here $u_{j+1} < u_j/2$ and $u_{h-1} > T_0$. Thus

$$\begin{aligned} \log(v_h/u_h) &< \frac{2(2+\sqrt{2})}{\sqrt{T_0}(1-1/T_0)} \\ &< \frac{1}{2389}. \end{aligned} \tag{730}$$

The first step sums $\log((1+x_j)/(1-x_j)) \le 2x_j/(1-x_j^2)$. The second uses $T_0 = 2^{28}$ and $\sqrt{2} < 10/7$. Since $e^t < 1/(1-t)$ for $0 < t < 1$, this gives $v_h/u_h < 2389/2388$.

Also $u_h > g(T_0) > 1$, where $g(u) = (u - \sqrt{u})/2$, so the finite certificate applies at u_h.

The certificate also transfers to $(\lambda, \mu) = (u_h, \vartheta u_h)$, $\vartheta \ge 1$, with at most a factor ϑ increase in the final upper bound. Whitened updates are linear in μ and independent of μ in their lower update. Each saved rational control remains legal: $c \le \sqrt{\mu}$ is preserved when μ increases, and $c \ge \sqrt{\lambda}$, $c < \lambda$ are unchanged. Its upper update is again linear in μ. Finally the right side of Eq. (726) is at most multiplied by ϑ, since $\sqrt{\vartheta\mu} \le \vartheta\sqrt{\mu}$. The same reasoning applies to rounded certificate enclosures by multiplying their upper bounds by ϑ and retaining their lower bounds. Consequently the global final bound is

$$\frac{307}{20}\frac{2389}{2388} = \frac{733423}{47760} < \frac{384}{25}. \tag{731}$$

The first step follows from simple algebra. The second is an exact rational comparison.

For $u < 1$, no step is needed. The halving count in the tail and the fixed finite-interval count prove the step bound. □

The scalar halving bound yields a deterministic common tree. A maximum-weight spanning-tree basis controls the numerical normalization for arbitrary positive weights.

Lemma 13.10 (Deterministic construction of a common thin tree)**.** *Let $G = (V, E)$ be a connected undirected loopless multigraph with $N := |V| \ge 2$ vertices and k positive weightings on the same edge set. Let $Q^{(\ell)}$ be the Laplacian of weighting ℓ, let u_e be a signed incidence vector, and suppose $\tau_e^{(\ell)} u_e^*(Q^{(\ell)})^\dagger u_e \le \alpha$ for every edge e and weighting ℓ. Set $m_{\mathrm{edge}} := |E|$ and $d_{\mathrm{blk}} := k(N-1)$. A deterministic algorithm constructs a spanning tree F with the original edge weights such that*

$$Q_F^{(\ell)} \preceq \min\{1, 268\alpha \log(2k)\} Q^{(\ell)} \qquad (\ell \in [k])$$

in $\widetilde{O}(m_{\mathrm{edge}} d_{\mathrm{blk}}^2 + d_{\mathrm{blk}}^{\omega+4})$ real arithmetic operations and comparisons for every admissible matrix multiplication exponent $\omega \in [2, 3]$.

Proof. For each weighting choose a maximum-weight spanning tree F_ℓ. Its weighted incidence columns form a basis on $\mathbf{1}^\perp$. In this basis the vector for an edge e has entries $c_{e,f}^{(\ell)} = \pm\sqrt{\tau_e^{(\ell)}/\tau_f^{(\ell)}}$ on the tree path of e, and zero elsewhere. Tree edges give the coordinate vectors. The maximum-tree

exchange property gives $\tau_e^{(\ell)} \le \tau_f^{(\ell)}$ for every edge f on that path. Consequently all entries have absolute value at most one and

$$I \preceq H_\ell := \sum_e c_e^{(\ell)} c_e^{(\ell)*} \preceq m_{\mathrm{edge}} N I. \tag{732}$$

The leverage bound is invariant under this change of basis: $c_e^{(\ell)*} H_\ell^{-1} c_e^{(\ell)} \le \alpha$. The implementation uses these path coordinates directly.

For an explicit finite-accuracy implementation, retain tree vectors exactly and approximate every other scalar square root on $[0,1]$ to absolute error $\xi := \eta/[100 Q_0^2 (m_{\mathrm{edge}} N)^3]$ by bisection. Denote the resulting vectors by $\widehat{c}_e^{(\ell)}$ and their sum of outer products by $\widehat{H}_\ell$. These matrices are PSD and $I \preceq \widehat{H}_\ell \preceq 2 m_{\mathrm{edge}} N I$. Uniformly for every edge set $\mathcal{E}$,

$$\| \sum_{e \in \mathcal{E}} (\widehat{c}_e^{(\ell)} \widehat{c}_e^{(\ell)*} - c_e^{(\ell)} c_e^{(\ell)*}) \| \le 3 m_{\mathrm{edge}} N \xi =: \zeta. \tag{733}$$

This follows by expanding each outer-product difference and using $\|c_e^{(\ell)}\| \le \sqrt{N}$ and $\|\widehat{c}_e^{(\ell)} - c_e^{(\ell)}\| \le \sqrt{N}\xi$.

Compute symmetric approximate inverse square roots R_ℓ of $\widehat{H}_\ell$, so that $(1-\eta) I \preceq R_\ell \widehat{H}_\ell R_\ell \preceq (1+\eta) I$, and form $\widehat{X}_e := \bigoplus_\ell R_\ell \widehat{c}_e^{(\ell)} \widehat{c}_e^{(\ell)*} R_\ell$. Residual-certified spectral computations on the interval $[1, 2 m_{\mathrm{edge}} N]$ have polylogarithmic accuracy cost. The matrices satisfy $\mathrm{rank}(\widehat{X}_e) \le k$, $(1-\eta) I \preceq \sum_e \widehat{X}_e \preceq (1+\eta) I$, and $\|\widehat{X}_e\| \le (1+10\eta)\alpha$. For the last inequality, apply the leverage bound after congruence and Eq. (733); use $\alpha \ge 1/m_{\mathrm{edge}}$ and $\|R_\ell\|^2 \le 1+\eta$.

Construction, error margins, and extraction. Use $\alpha_{\mathrm{edge}} := (1+10\eta)\alpha$ and start the scalar recurrence at $\lambda = (1-\eta)/\alpha_{\mathrm{alg}}$, $\mu = (1+\eta)/\alpha_{\mathrm{alg}}$. The proof of Lemma 13.9 transfers its certificate from (u,u) to $(u, \vartheta u)$ with $\vartheta = (1+\eta)/(1-\eta)$ by multiplying upper bounds by ϑ. Each selected signing is computed by the deterministic routine above. The terminal sum of approximate edge matrices obeys $0 \prec \widehat{S} \preceq C_0 \vartheta \alpha_{\mathrm{alg}} I$. Every nontrivial finite witness has lower bound at least $\alpha_{\mathrm{alg}}/Q_0$, since its terminal integer lower bound is positive. When no step is needed the initial positive lower bound suffices.

All matrix sums during this construction have polynomial spectral bounds: before a nontrivial step their least eigenvalue is at least $\alpha_{\mathrm{alg}}/Q_0$, and $\alpha_{\mathrm{alg}} \ge \mathcal{C}_{\mathrm{alg}}^2/m_{\mathrm{edge}}$. Regularization only improves the condition number. The fixed rational controls have fixed denominator bounds; the tail consists of well-conditioned whitened steps. Compute the scalar controls and matrix functions with outward enclosures below $\eta/(m_{\mathrm{edge}} d_{\mathrm{blk}})^C$ for a sufficiently large absolute constant C. The finite certificate has strict rational margins, and the tail has $O(\log(2m_{\mathrm{edge}}))$ steps, so these enclosures preserve legality and positivity. Their total relative loss can be bounded by another factor $1+\eta$ in the upper estimate. A logarithm used to define α_{alg} is likewise replaced by an upper enclosure of relative error at most η. These choices require only $O(\log(2m_{\mathrm{edge}} d_{\mathrm{blk}}))$ accuracy bits, with fixed multipliers.

Eq. (733) bounds the error of the selected subfamily independently of the sequence of signings. It is smaller than half its terminal positive lower bound. The original edge set is therefore connected in every weighting. The upper bound, after returning to the original graph coordinates, is at most

$$\{C_0 (4.171001)^2 (1+30\eta) + \eta\} \alpha \log(2k) < 268 \alpha \log(2k).$$

Here the factors cover the initial sandwich, the norm bound, congruence and scalar enclosures, and the uniform input error. An arbitrary spanning tree of this connected subfamily preserves the upper bounds because all weights are positive. It also obeys $Q_F^{(\ell)} \preceq Q^{(\ell)}$, giving the cap at one.

Operation count. There are $O(\log(2m_{\mathrm{edge}}))$ signing calls on at most m_{edge} matrices of order $d_{\mathrm{blk}} = k(N-1)$. Each costs $\widetilde{O}(m_{\mathrm{edge}} d_{\mathrm{blk}}^2 + d_{\mathrm{blk}}^{\omega+4})$ by Theorem 12.15 and the subisotropic reduction. Maximum spanning trees, path coordinates, and square-root bisections cost $\widetilde{O}(k m_{\mathrm{edge}} N)$. All congruences remain block diagonal, with one rank-one edge factor in each block. Thus input formation and each change of congruence cost $\widetilde{O}(m_{\mathrm{edge}} k N^2 + k N^3)$, which fit the stated bound. The final graph search costs $O(m_{\mathrm{edge}} + N)$. Logarithmic factors absorb the number of calls. This proves the stated arithmetic bound, including normalization and exact grouped elimination. □

13.5 A strongly Rayleigh bound for few weightings

A measure is strongly Rayleigh when its multi-affine generating polynomial is real stable: it is nonzero whenever every variable has positive imaginary part. Anari and Oveis Gharan [AO14] proved the following bound. If a homogeneous strongly Rayleigh measure has every marginal at most ϵ_1, and fixed real vectors z_i satisfy $\sum_i z_i z_i^* = I$ and $\|z_i\|^2 \le \epsilon_2$, then some sample in its support satisfies

$$\|\sum_{i\in S} z_i z_i^*\| \le 4(\epsilon_1+\epsilon_2) + 2(\epsilon_1+\epsilon_2)^2.$$

Cloning each edge reduces the quadratic term in this bound while retaining one common spanning tree.

Lemma 13.11 (A common tree from a strongly Rayleigh measure)**.** *Let $G=(V,E)$ be a connected undirected loopless multigraph with $N := |V| \ge 2$ vertices and k positive weightings on the same edge set. Let $Q^{(\ell)}$ be the Laplacian of weighting ℓ and u_e a signed incidence vector. If $\tau_e^{(\ell)} u_e^* (Q^{(\ell)})^\dagger u_e \le \alpha$ for every e and ℓ, then there is a spanning tree F with the original edge weights satisfying*

$$Q_F^{(\ell)} \preceq \min\{1, 4(k+1)\alpha\} Q^{(\ell)} \qquad (\ell \in [k]).$$

Proof. Take the weighted spanning-tree distribution for the first weighting. By the matrix-tree theorem its generating polynomial is a normalized determinant of a PSD linear pencil on $\mathbf{1}^\perp$. It is real stable, homogeneous, and multi-affine. Its edge marginals are $\tau_e^{(1)} u_e^* (Q^{(1)})^\dagger u_e \le \alpha$.

For an integer $R \ge 1$, make $M := R2^k$ clones of every edge, indexed by $\eta \in \{-1,1\}^k$ and $t \in [R]$. Associate to a clone the vector

$$z_{e,\eta,t} := M^{-1/2}(\eta_1 \psi_e^{(1)} \oplus \cdots \oplus \eta_k \psi_e^{(k)}),$$

where $\psi_e^{(\ell)} := \sqrt{\tau_e^{(\ell)}}(Q^{(\ell)})^{-1/2} u_e$, with inverse square roots restricted to $\mathbf{1}^\perp$. Summing over all sign patterns cancels the off-diagonal blocks, so

$$\sum_{e,\eta,t} z_{e,\eta,t} z_{e,\eta,t}^* = I_{k(N-1)}, \qquad \|z_{e,\eta,t}\|^2 \le k\alpha/M.$$

Sample a tree from the first weighting's distribution and independently choose one uniform clone of each selected edge. Its generating polynomial is obtained by replacing each original variable by the average of its clone variables. This preserves real stability, homogeneity, and multi-affinity. Each clone marginal is at most α/M.

Applying the cited theorem with $\epsilon_1 := \alpha/M$ and $\epsilon_2 := k\alpha/M$ gives a cloned sample whose matrix sum has norm at most

$$\frac{4(k+1)\alpha}{M} + \frac{2(k+1)^2\alpha^2}{M^2}.$$

Its ℓ-th principal block is $M^{-1}(Q^{(\ell)})^{-1/2}Q_F^{(\ell)}(Q^{(\ell)})^{-1/2}$, where F is its underlying tree. Thus the same tree satisfies, for every ℓ,

$$Q_F^{(\ell)} \preceq (4(k+1)\alpha + 2(k+1)^2\alpha^2/M)Q^{(\ell)}.$$

There are finitely many spanning trees. Letting R grow and choosing a tree occurring along an infinite subsequence therefore gives

$$Q_F^{(\ell)} \preceq 4(k+1)\alpha Q^{(\ell)} \qquad \text{for every } \ell \in [k].$$

Also $Q_F^{(\ell)} \preceq Q^{(\ell)}$ because $F \subseteq E$. This proves the stated simultaneous bound. □

13.6 Rounding with two fixed matrix targets

For PSD inputs $X_1, \dots, X_{m_{\text{in}}}$ and $X \succ 0$, write

$$f_i(X) := \|X_i^{1/2} X X_i^{1/2}\|_{S_\nu}, \qquad F_y(X) := g\sum_{i=1}^{m_{\text{in}}} \beta(y_i) f_i(X) X_i, \qquad M(y) := \sum_{i=1}^{m_{\text{in}}} y_i X_i.$$

We use the following two choices of Schatten order and budget:

$$\nu = 3: \quad g := \frac{17}{4}, \quad \beta(y) := \sqrt{1-y^2}, \tag{734}$$

$$\nu \in \mathbb{Z},\ \nu \geq 5: \quad g := \frac{119\nu}{100} + \frac{4}{5}, \quad \beta(y) := \sqrt{1-y^2} f(|y|), \tag{735}$$

$$f(t) := 1 + \frac{1}{25}t^2 - \frac{11}{100}t^4.$$

Every terminating decimal below is an exact rational number.

Lemma 13.12 (Fixed-target rounding)**.** *Fix either Eq. (734) or Eq. (735). Let $X_1, \dots, X_{m_{\text{in}}} \succeq 0$ be a finite family. For any $y^0 \in [-1,1]^{m_{\text{in}}}$ and $U_0, V_0 \succ 0$, there is one signing satisfying*

$$M(\sigma) \preceq M(y^0) + U_0^{-1} + F_{y^0}(V_0), \tag{736}$$

$$-M(\sigma) \preceq -M(y^0) + V_0^{-1} + F_{y^0}(U_0). \tag{737}$$

There is one sign per original input matrix.

Proof. Both choices of Schatten order and budget satisfy Lemma 11.12. The lemma applies to arbitrary PSD inputs and the two fixed targets, so it gives both inequalities for the same signing. □

Corollary 13.13 (Matrix interval for a selected class)**.** *Put $S = \sum_i X_i$. For $0 < p < 1$ and arbitrary $U, V \succ 0$, some subfamily $\mathcal{J}$ satisfies*

$$pS - \frac{1}{2}(V^{-1} + F_x(U)) \preceq \sum_{i\in\mathcal{J}} X_i \preceq pS + \frac{1}{2}(U^{-1} + F_x(V)), \qquad x = (2p-1)\mathbf{1}. \tag{738}$$

Proof. Use Lemma 13.12 at $y^0 = x$ and set $\mathcal{J} = \{i : \sigma_i = 1\}$. Since $M(x) = (2p-1)S$ and $2\sum_{i\in\mathcal{J}} X_i = S + M(\sigma)$, the conclusion follows. □

13.7 Thinning with a lower bound in one block

The first block will certify connectivity. The other blocks require upper bounds only. The two scalar bounds are measured in different units; retaining this distinction is essential.

Lemma 13.14 (Two-scale thinning)**.** *Let*

$$X_i = \bigoplus_{\ell=1}^{k} X_i^{(\ell)}, \qquad X_i^{(\ell)} \succeq 0, \qquad \operatorname{rank} X_i^{(\ell)} \le 1, \qquad \|X_i\| \le \alpha.$$

Choose a certified Schatten order ν and let $R \ge k^{1/\nu}$. Put

$$\alpha_0 = 4g\alpha, \qquad \alpha_{\mathrm{up}} = R\alpha_0. \tag{739}$$

Suppose the retained sum $S = \sum_i X_i$ has bounds

$$S^{(1)} \succeq \lambda\alpha_0 I, \qquad S^{(\ell)} \preceq \mu\alpha_{\mathrm{up}} I \quad (1 \le \ell \le k), \qquad \lambda, \mu > 0. \tag{740}$$

For $0 < p < 1$ and $0 < \lambda' < p\lambda$, define

$$d(p) = \frac{\beta(2p-1)}{4}, \qquad \mu' = p\mu + \frac{d(p)(\sqrt{\lambda} + \sqrt{\mu})^2}{4(p\lambda - \lambda')}. \tag{741}$$

There is a subfamily whose sum S' satisfies

$$S'^{(1)} \succeq \lambda'\alpha_0 I, \qquad S'^{(\ell)} \preceq \mu'\alpha_{\mathrm{up}} I \quad (1 \le \ell \le k).$$

In particular its first block is positive definite. No lower bound on the other blocks is required.

Proof. Abbreviate $d = d(p) > 0$, $L = p\lambda - \lambda' > 0$ and define

$$u = \frac{2L}{d(\lambda + \sqrt{\lambda\mu})}, \qquad v = \frac{\lambda + \sqrt{\lambda\mu}}{2L\sqrt{\lambda\mu}}. \tag{742}$$

Direct algebra gives

$$\frac{1}{v} + du\lambda = 2L, \qquad \frac{1}{u} + dv\mu = \frac{d(\sqrt{\lambda} + \sqrt{\mu})^2}{2L}, \qquad du < 2p. \tag{743}$$

For $h > 0$, choose the positive definite trial matrices

$$U = \frac{u}{\alpha_{\mathrm{up}}} I, \qquad V_h = \frac{R}{\alpha_{\mathrm{up}}}(vI_1 \oplus hI_2 \oplus \cdots \oplus hI_k). \tag{744}$$

Each block has rank at most one and norm at most α. Therefore

$$f_i(U) \le \frac{u\alpha}{\alpha_{\mathrm{up}}} k^{1/\nu} \le \frac{u\alpha R}{\alpha_{\mathrm{up}}}, \qquad f_i(V_h) \le \frac{\alpha R}{\alpha_{\mathrm{up}}}\xi_h, \quad \xi_h = (v^\nu + (k-1)h^\nu)^{1/\nu}.$$

At $x = (2p-1)\mathbf{1}$, the coupling maps obey

$$F_x(U) \preceq duS, \qquad F_x(V_h) \preceq d\xi_h S, \tag{745}$$

because $g\beta(2p-1)\alpha R/\alpha_{\mathrm{up}} = d$.

Apply Corollary 13.13. Its upper bound, in every block, is at most

$$(p\mu + \frac{1}{2}(\frac{1}{u} + d\xi_h\mu))\alpha_{\mathrm{up}} I.$$

In the first block its lower bound is at least

$$(p - \frac{du}{2})S^{(1)} - \frac{\alpha_{\mathrm{up}}}{2Rv} I \succeq \frac{\alpha_{\mathrm{up}}}{R}(p\lambda - \frac{du\lambda}{2} - \frac{1}{2v})I = \lambda'\alpha_0 I.$$

The Loewner comparison uses $p - du/2 > 0$, checked in (743). This lower bound is independent of h.

For each $h > 0$ a subfamily exists. Let $h \downarrow 0$ and choose a recurrent subfamily from the finite set of all subsets of the current family. Since $\xi_h \to v$, its upper bound converges to $\mu'\alpha_{\mathrm{up}} I$ by (743). The first-block lower bound remains $\lambda'\alpha_0 I \succ 0$. For $k = 1$, no limiting auxiliary blocks are needed. This proves the lemma. □

For an isotropic input family the initial state is $(\lambda, \mu) = (1/\alpha_0, 1/(R\alpha_0))$. Its upper-to-lower ratio is $1/R$, while the upper scale remains $4gR\alpha$. The saving uses the rank-one block structure and the lower bound in only the first block.

13.8 A prefix covering every initial scale

Fix a budget for the coupling, a rational bias $p_0 \in (0, 1)$, and a rational reference lower bound $\lambda_0 > 0$. Write $d_0 = d(p_0)$ and $h_0 = \sqrt{d_0}$. Define an increasing sequence backward by

$$\lambda_{k+1} = (\frac{h_0 + \sqrt{h_0^2 + 4p_0\lambda_k}}{2p_0})^2. \tag{746}$$

Then

$$\lambda_{k-1} = p_0\lambda_k - h_0\sqrt{\lambda_k} > 0, \qquad \lambda_k > \lambda_{k-1}/p_0. \tag{747}$$

Lemma 13.14 gives the corresponding upper update

$$\mu_{k-1} = p_0\mu_k + \frac{h_0}{4\sqrt{\lambda_k}}(\sqrt{\lambda_k} + \sqrt{\mu_k})^2. \tag{748}$$

All prefix selections are legal, since their lower gap is $h_0\sqrt{\lambda_k} > 0$.

Put $n_{\mathrm{pre}} = 16$ and $Q = 2^{40}$. The three plans in Section 13.9 satisfy

$$\frac{h_0}{\sqrt{p_0}(1 - \sqrt{p_0})} < 2, \qquad \frac{h_0}{p_0\sqrt{\lambda_{n_{\mathrm{pre}}}}} < \frac{1}{16}, \qquad \frac{1}{Q} + \frac{8}{\sqrt{\lambda_{n_{\mathrm{pre}}}}} < 1. \tag{749}$$

Lemma 13.15 (Uniform entry bound)**.** *For $K \geq n_{\mathrm{pre}}$, suppose the prefix begins with lower state λ_K and ratio $\mu_K/\lambda_K \leq 1/R + 1/Q$, where $R \geq 1$. After the prefix reaches lower state $\lambda_{n_{\mathrm{pre}}}$, it has*

$$\mu_{n_{\mathrm{pre}}} \leq \lambda_{n_{\mathrm{pre}}}(\frac{1}{R} + \frac{1}{Q} + \frac{8}{\sqrt{\lambda_{n_{\mathrm{pre}}}}}). \tag{750}$$

Proof. Let $r_k = \mu_k/\lambda_k$ and $x_k = h_0/(p_0\sqrt{\lambda_k})$. The exact ratio update is

$$r_{k-1} = \frac{r_k + x_k(1 + \sqrt{r_k})^2/4}{1 - x_k}.$$

If $r_k \le 2$ and $x_k \le 1/16$, then $\sqrt{r_k} < 3/2$ and

$$r_{k-1} - r_k = \frac{x_k}{1-x_k}\left(r_k + \frac{(1+\sqrt{r_k})^2}{4}\right) \le \frac{57}{15}x_k < 4x_k.$$

By (747),

$$\sum_{k>n_{\mathrm{pre}}} x_k < \frac{h_0}{\sqrt{p_0}(1-\sqrt{p_0})\sqrt{\lambda_{n_{\mathrm{pre}}}}} < \frac{2}{\sqrt{\lambda_{n_{\mathrm{pre}}}}}.$$

Starting from $r_K \le 1/R + 1/Q \le 1 + 1/Q$, a first-crossing argument and (749) show that the ratio cannot reach 2: the total increase before any such crossing would be less than $8/\sqrt{\lambda_{n_{\mathrm{pre}}}}$, and the resulting upper bound is still below 2. Thus the increment estimate holds for the entire prefix. Summing it proves (750). When $K = n_{\mathrm{pre}}$, the assertion is immediate. □

13.9 Finite plans and restoration of original inputs

Three finite plans are used. Each plan can use either choice above, with its budget determined by the Schatten order ν. All entries are exact rational numbers.

Parameter	C	L	H
p_0	0.1322257	0.1336908	0.1335
λ_0	5651.7592	4660.4663	6142
λ_1^{fin}	636.7130	560.8921	664.2
λ_2^{fin}	69.7255	63.2659	71.668
λ_3^{fin}	5.6824	5.2338	5.7817
λ_4^{fin}	10^{-6}	10^{-6}	10^{-6}
p_1	0.1202756	0.1280115	0.116057
p_2	0.1312033	0.1338242	0.130928
p_3	0.1377143	0.1368013	0.140214
p_4	0.1240438	0.1207605	0.131498

The backward prefix states λ_k in (746) are distinct from the four finishing targets λ_j^{fin} in this table. After the prefix reaches λ_0, use the four finishing transitions (741) at the displayed biases and targets. Each satisfies $0 < \lambda_j^{\mathrm{fin}} < p_j \lambda_{j-1}^{\mathrm{fin}}$, with $\lambda_0^{\mathrm{fin}} = \lambda_0$.

For a plan $P \in \{C, L, H\}$, an allowed order ν, and $R \ge 1$, define $C_{P,\nu}(R)$ by the following finite scalar evaluation. Initialize the upper state at the right side of (750); perform exactly $n_{\mathrm{pre}} = 16$ prefix upper updates (748) to reach lower state λ_0; then perform the four finishing transitions. The resulting final upper state is $C_{P,\nu}(R)$. This is an explicit expression involving only rational numbers and square roots. The final lower state is exactly 10^{-6}.

Lemma 13.16 (An original subfamily with a positive first block)**.** *Suppose X_i have the block-rank-one form in Lemma 13.14, satisfy $\sum_i X_i = I$, and obey $\|X_i\| \le \alpha$. Choose an allowed Schatten order ν with its associated budget, $R \ge k^{1/\nu}$, and a plan. There is an original subfamily $\mathcal{J}$ such that*

$$\sum_{i\in\mathcal{J}} X_i^{(1)} \succ 0, \qquad \sum_{i\in\mathcal{J}} X_i^{(\ell)} \preceq \min\{1, 4gRC_{P,\nu}(R)\alpha\} I \quad (1 \le \ell \le k). \tag{751}$$

Every retained original matrix has coefficient exactly one.

Proof. Put $\alpha_0 = 4g\alpha$ and choose $K \geq n_{\mathrm{pre}}$ so large that $\alpha_0\lambda_K \geq Q$. This is possible because $\lambda_K \to \infty$. Let

$$M = \lceil \alpha_0 \lambda_K \rceil.$$

Replace each input temporarily by M copies of X_i/M. The total sum remains I. Its norm cap is α/M, so the two scales are α_0/M and $R\alpha_0/M$. We may use initial lower state λ_K and upper state $M/(R\alpha_0)$. Their ratio satisfies

$$\frac{M}{R\alpha_0\lambda_K} \leq \frac{1}{R} + \frac{1}{RQ} \leq \frac{1}{R} + \frac{1}{Q}.$$

Apply the prefix down to $\lambda_{n_{\mathrm{pre}}}$, then use Lemma 13.15. Continue with the finite plan defining $C_{P,\nu}(R)$. The resulting copied sum S obeys

$$S^{(1)} \succeq 10^{-6}\frac{\alpha_0}{M} I \succ 0, \qquad S^{(\ell)} \preceq C_{P,\nu}(R)\frac{R\alpha_0}{M} I \quad (1 \leq \ell \leq k).$$

If k_i is the number of surviving copies of original input i, define $\mathcal{J} = \{i : k_i > 0\}$. Since $1 \leq k_i \leq M$ whenever $i \in \mathcal{J}$, the PSD inequalities

$$S = \sum_i \frac{k_i}{M} X_i \preceq \sum_{i \in \mathcal{J}} X_i \preceq \sum_i k_i X_i = MS \tag{752}$$

prove first-block positive definiteness and the claimed upper bound. The cap by one follows because $\mathcal{J}$ is a subfamily of the original isotropic family. The cloning count is finite, and the positive first-block lower bound is retained before removing repeated indices. □

13.10 The uniform coefficient

For $1 \leq k \leq 31$, use Lemma 13.11. The derivative of $(k+1)/\log(2k)$ has the sign of $\log(2k) - 1 - 1/k$, which is strictly increasing. Thus the ratio has no interior maximum. The endpoint comparisons are

$$\frac{8}{\log 2} < 11.541560327112, \qquad \frac{128}{\log 62} < 31.014255427160 < 31.6.$$

For the finite range, use order three with plan L on $32 \leq k \leq 33$, plan C on $34 \leq k \leq 82$, and plan H on $83 \leq k \leq 391$. Use order five with plan L on $392 \leq k \leq 441$, plan C on $442 \leq k \leq 2364$, and plan H on $2365 \leq k \leq 5644$. Use order six and plan H on $5645 \leq k \leq 5^6$. For each $7 \leq \nu \leq 13$, use order ν and plan H on $5^{\nu-1} < k \leq 5^\nu$.

We generate the partition of these fourteen integer ranges by midpoint subdivision, keeping the Schatten order and plan fixed within each range. For an interval $[l, h]$ with order ν and plan P, define

$$R_\nu(h) := 10^{-6} \min\{j \in \mathbb{Z}_{\geq 0} : j^\nu \geq 10^{6\nu} h\}.$$

Thus $R_\nu(h)$ is the least multiple of 10^{-6} whose νth power is at least h. Let $U_{P,\nu}(l, h)$ be the certified upper endpoint, computed by the rational interval rules below, of

$$\frac{4gR_\nu(h)C_{P,\nu}(R_\nu(h))}{\log(2l)}.$$

Retain $[l, h]$ when $U_{P,\nu}(l, h) < 31.6$. Otherwise, if $l < h$, put $m := \lfloor (l + h)/2 \rfloor$ and apply the same rule first to $[l, m]$ and then to $[m + 1, h]$. A singleton that fails the test terminates the construction with failure.

The arithmetic uses rational intervals rounded outward to multiples of 2^{-160}. Integer-square comparisons enclose every square root. Check the three prefix guards (749) for each of the three plans and each of the two budgets, and check all four finishing gaps using the rational plan parameters. Logarithms are reduced to $1 \le v \le 2$ and enclosed by

$$\log v = 2\sum_{j=0}^{89} \frac{z^{2j+1}}{2j+1} + R_{90}, \qquad z := \frac{v-1}{v+1}, \tag{753}$$

$$0 \le R_{90} \le \frac{2z^{181}}{181(1-z^2)}. \tag{754}$$

To evaluate $U_{P,\nu}(l,h)$, set $R := R_\nu(h)$ and generate $\lambda_0, \dots, \lambda_{16}$ by Eq. (746), initialize the upper state by Eq. (750), perform the 16 prefix updates in Eq. (748), and then perform the four finishing updates in Eq. (741). After each arithmetic operation round the lower endpoint down and the upper endpoint up to the grid $2^{-160}\mathbb{Z}$; divisions require a strictly positive denominator interval. For a nonnegative interval $[a,b]$, enclose its square root by $[j/2^{160}, J/2^{160}]$, where j is the largest integer with $j^2 \le 2^{320}a$ and J is the smallest integer with $J^2 \ge 2^{320}b$. Evaluate the displayed formulas in their written order, with powers formed by repeated multiplication and sums accumulated from left to right. For the logarithm write $2l = 2^q v$, $1 \le v < 2$, and bound $q\log 2 + \log v$ by Eqs. (753) and (754). Section 11.7 supplies the certificates for these choices of Schatten order and budget.

The following verification completes the finite range by extending each accepted comparison to every integer in its interval.

Lemma 13.17 (Recursive interval certificate)**.** *The midpoint construction terminates successfully after evaluating 134 intervals, with at most six subdivisions along any branch. Its 74 retained intervals partition the integers from 32 to 5^{13}. For every retained interval $[l,h]$, $U_{P,\nu}(l,h) < 31.599872189081$. In particular, with $R := R_\nu(h)$, every integer $k \in [l,h]$ satisfies $R \ge k^{1/\nu}$ and*

$$\frac{4gRC_{P,\nu}(R)}{\log(2k)} < 31.6. \tag{755}$$

Proof. Exact rational interval evaluation of the fourteen initial ranges and their midpoint descendants gives 60 subdivisions and 74 accepted intervals; the largest subdivision depth is six. Every required denominator is positive, all prefix guards and finishing gaps pass, and every singleton encountered is accepted. The largest upper endpoint occurs at $[l,h] = [34,35]$, with order three, plan C, and $R = 3271067/10^6$; it is strictly below 31.599872189081.

Each subdivision preserves the covered integers and produces disjoint children. The fourteen initial ranges partition $[32, 5^{13}] \cap \mathbb{Z}$, so their accepted descendants give the stated partition. The integer-power definition of R gives $R^\nu \ge h \ge k$. Finally, $\log(2k) \ge \log(2l) > 0$, so the ratio in Eq. (755) is at most $U_{P,\nu}(l,h) < 31.6$. □

The same calculation gives

$$C_{H,5}(5) < 1.950374193570 < 1.95038. \tag{756}$$

The budget is the same for every $\nu \ge 5$, so $C_{H,\nu}(5) = C_{H,5}(5)$. For $k > 5^{13}$, take the least integer $\nu \ge 14$ with $k \le 5^\nu$. Then $R = 5$ is valid and $k > 5^{\nu-1}$. Consequently,

$$\frac{4gRC_{H,\nu}(5)}{\log(2k)} < \frac{20(119\nu/100 + 4/5)(1.95038)}{(\nu-1)\log 5 + \log 2} \le \frac{20(873/50)(1.95038)}{13\log 5 + \log 2} < 31.508037377695 < 31.6.$$

The middle expression decreases with ν, because its derivative has the sign of $(119/100)\log 2 - (199/100)\log 5 < 0$. This covers every larger weighting count and proves

$$\min\{4(k+1),\ \inf_{\nu,P,R} 4gRC_{P,\nu}(R)\} \le 31.6\log(2k), \tag{757}$$

where the infimum ranges over the two choices of Schatten order and budget in Lemma 13.12, the three plans, and $R \ge k^{1/\nu}$.

13.11 One spanning tree for all weightings

The subfamily with a positive first block yields a common tree with logarithmic dependence on the number of weightings. The strongly Rayleigh estimate improves the bound for small k, and Lemma 13.10 gives the construction.

Theorem 13.18 (Formal version of Theorem 1.8)**.** *Let $G = (V, E)$ be a connected undirected loopless multigraph with $N := |V| \ge 2$ vertices. For each $\ell \in [k]$, assign positive weights $\tau_e^{(\ell)} > 0$ to every edge of the same edge set, and put*

$$Q^{(\ell)} := \sum_{e\in E} \tau_e^{(\ell)} u_e u_e^*, \qquad Q_{\mathcal{E}}^{(\ell)} := \sum_{e\in\mathcal{E}} \tau_e^{(\ell)} u_e u_e^*.$$

Here $\mathcal{E} \subseteq E$ is an arbitrary edge subset, u_e is a signed incidence vector, and $(Q^{(\ell)})^\dagger$ denotes the Moore–Penrose pseudoinverse. If every edge leverage satisfies $\tau_e^{(\ell)} u_e^* (Q^{(\ell)})^\dagger u_e \le \alpha$, there is one spanning tree $F \subseteq E$ such that, simultaneously for all $\ell \in [k]$,

$$Q_F^{(\ell)} \preceq \min\{1, 31.6\alpha\log(2k), 4(k+1)\alpha\} Q^{(\ell)}.$$

In particular, the same tree is $31.6\alpha\log(2k)$-spectrally thin for all k weightings, with every retained edge weight unchanged.

Put $m_{\mathrm{edge}} = |E|$ and $d_{\mathrm{blk}} = k(N-1)$. A deterministic algorithm constructs a possibly different common spanning tree satisfying $Q_F^{(\ell)} \preceq \min\{1, 268\alpha\log(2k)\} Q^{(\ell)}$ for every ℓ, with the original edge weights, in $\widetilde{O}(m_{\mathrm{edge}} d_{\mathrm{blk}}^2 + d_{\mathrm{blk}}^{\omega+4})$ real arithmetic operations and comparisons. The current bound $\omega < 2.371177$ of Dupont et al. [DEK+ 26] gives $\widetilde{O}(m_{\mathrm{edge}} d_{\mathrm{blk}}^2 + d_{\mathrm{blk}}^{6.371177})$. Ordinary dense multiplication gives $\widetilde{O}(m_{\mathrm{edge}} d_{\mathrm{blk}}^2 + d_{\mathrm{blk}}^7)$.

Proof. Orient the edges arbitrarily and let u_e be their signed incidence vectors. Work on $\mathbf{1}^\perp$, of dimension $N-1$, where each $Q^{(\ell)}$ is positive definite. Define

$$\psi_e^{(\ell)} := \sqrt{\tau_e^{(\ell)}} (Q^{(\ell)})^{-1/2} u_e, \qquad X_e := \bigoplus_{\ell=1}^{k} \psi_e^{(\ell)} \psi_e^{(\ell)*}.$$

The inverse square roots in this display are restricted to $\mathbf{1}^\perp$. The matrices satisfy

$$\sum_{e\in E} X_e = I_{k(N-1)}, \qquad \operatorname{rank}(X_e) \le k, \qquad \|X_e\| = \max_\ell \tau_e^{(\ell)} u_e^* (Q^{(\ell)})^\dagger u_e \le \alpha.$$

Each individual block has rank at most one. Apply Lemma 13.16 with one of the certified choices of (ν, P, R). It gives a common edge set $\mathcal{J}$ such that

$$(Q^{(1)})^{-1/2} Q_{\mathcal{J}}^{(1)} (Q^{(1)})^{-1/2} \succ 0, \qquad (Q^{(\ell)})^{-1/2} Q_{\mathcal{J}}^{(\ell)} (Q^{(\ell)})^{-1/2} \preceq \min\{1, 4gRC_{P,\nu}(R)\alpha\} I \quad (\ell \in [k]).$$

Since all weights in the first weighting are positive, its lower bound implies that $(V, \mathcal{J})$ is connected. Choose any spanning tree $F \subseteq \mathcal{J}$. Every weighting is positive on the same edge set, so $Q_F^{(\ell)} \preceq Q_{\mathcal{J}}^{(\ell)}$ for every ℓ. Undoing the congruences gives the asserted upper bounds on $\mathbf{1}^{\perp}$; all Laplacians vanish on the constant vector, so the inequalities hold on the full vertex space. The selected edges have their original weights by Eq. (752).

Lemma 13.11 supplies the complementary coefficient $4(k+1)$. The comparison in Eq. (757) depends only on k, so choose a tree from the smaller bound. Every tree also satisfies $Q_F^{(\ell)} \preceq Q^{(\ell)}$, giving the cap by one and the stated minimum. The deterministic guarantee and operation count follow from Lemma 13.10. The current bound $\omega < 2.371177$ [DEK+26] gives exponent 6.371177; ordinary dense multiplication gives exponent 7. □

13.12 A complementary bound by matrix pipage rounding

The matrix pipage-rounding theorem of Harvey and Olver [HO14] gives a second polynomial-time construction. It complements the logarithmic dependence on k with a bound in terms of $d_{\mathrm{blk}} = k(N-1)$.

Theorem 13.19 (Simultaneous thinness by matrix pipage rounding)**.** *Under the hypotheses of Theorem 13.18, with leverage bound α, put $d_{\mathrm{blk}} = k(N-1)$ and let $t_{d_{\mathrm{blk}}}$ be the least integer $t \geq 2$ satisfying*

$$t \log t - t + 1 \geq \log(2d_{\mathrm{blk}}). \tag{758}$$

There is a deterministic polynomial-time construction of a common spanning tree F such that, for every $\ell \in [k]$,

$$Q_F^{(\ell)} \preceq \min\{1, \alpha t_{d_{\mathrm{blk}}}\} Q^{(\ell)}. \tag{759}$$

Every retained edge keeps its original weight in each weighting. Moreover,

$$t_{d_{\mathrm{blk}}} = O\Big(\frac{\log(2d_{\mathrm{blk}})}{\log\log(4d_{\mathrm{blk}})}\Big).$$

Combining this construction with Theorem 13.18 gives a deterministic polynomial-time construction satisfying

$$Q_F^{(\ell)} \preceq \min\{1, 268\alpha \log(2k), \alpha t_{d_{\mathrm{blk}}}\} Q^{(\ell)} \qquad (\ell \in [k]). \tag{760}$$

In particular, its simultaneous thinness is

$$O\Big(\alpha \min\{\log(2k), \frac{\log(2Nk)}{\log\log(4Nk)}\}\Big).$$

The construction-time assertion uses the real-arithmetic model of matrix pipage rounding.

Proof. Work on $\mathbf{1}^{\perp}$, so each $Q^{(\ell)}$ is positive definite, and set

$$X_e^{(\ell)} := \tau_e^{(\ell)} (Q^{(\ell)})^{-1/2} u_e u_e^{*} (Q^{(\ell)})^{-1/2}, \qquad X_e := \bigoplus_{\ell=1}^{k} X_e^{(\ell)}. \tag{761}$$

The inverse square roots are restricted to $\mathbf{1}^{\perp}$. The common positive support and the leverage hypotheses imply

$$\sum_{e \in E} X_e = I_{d_{\mathrm{blk}}}, \qquad 0 \preceq X_e \preceq \alpha I_{d_{\mathrm{blk}}}. \tag{762}$$

Let x_e be the marginal probability of edge e in the weighted spanning-tree distribution for the first weighting. Thus a tree F has probability proportional to $\prod_{e\in F}\tau_e^{(1)}$. The vector x belongs to the graphic matroid base polytope, which is the convex hull of the incidence vectors of spanning trees. The matrix-tree theorem and differentiation of its determinant give

$$x_e = \tau_e^{(1)} u_e^* (Q^{(1)})^\dagger u_e \le \alpha. \tag{763}$$

If $Z := \sum_F \prod_{f\in F}\tau_f^{(1)}$, then $x_e = \tau_e^{(1)}\partial \log Z/\partial\tau_e^{(1)}$. The matrix-tree formula identifies Z with a constant multiple of the determinant of $Q^{(1)}$ restricted to $\mathbf{1}^\perp$; differentiating this determinant proves Eq. (763). Since all X_e are positive semidefinite,

$$\begin{aligned}\sum_e x_e X_e &\preceq \alpha \sum_e X_e \\ &= \alpha I_{d_{\mathrm{blk}}}.\end{aligned} \tag{764}$$

The first step uses Eq. (763) and $X_e \succeq 0$. The second uses Eq. (762).

In particular, the same fractional spanning tree controls every weighting simultaneously.

We use the following specialization of Harvey and Olver [HO14, Corollary 3.9]. If x lies in a matroid base polytope, $0 \preceq M_e \preceq I_{d_{\mathrm{blk}}}$, and $\sum_e x_e M_e \preceq I_{d_{\mathrm{blk}}}$, randomized pipage rounding to a base B satisfies, for $t > 1$,

$$\Pr[\|\sum_{e\in B} M_e\| \ge t] \le d_{\mathrm{blk}} \frac{e^{t-1}}{t^t}. \tag{765}$$

When this right-hand side is less than one, their deterministic pipage-rounding algorithm returns a base with norm strictly less than t. This is their theorem with norm bound $R = 1$, mean bound $\mu = 1$, and deviation parameter $t - 1$.

Apply it to $M_e := X_e/\alpha$ and the fractional point x above. Eqs. (762) and (764) verify its hypotheses. By Eq. (758), the right-hand side of Eq. (765) at $t = t_{d_{\mathrm{blk}}}$ is at most $1/2$. The resulting graphic-matroid base is a spanning tree F with $\|\sum_{e\in F} X_e\| < \alpha t_{d_{\mathrm{blk}}}$. Its ℓ-th block is $(Q^{(\ell)})^{-1/2} Q_F^{(\ell)} (Q^{(\ell)})^{-1/2}$. Undoing the congruence in each block gives $Q_F^{(\ell)} \preceq \alpha t_{d_{\mathrm{blk}}} Q^{(\ell)}$ on $\mathbf{1}^\perp$. Both Laplacians vanish on $\mathbf{1}$, so the inequality also holds on the full vertex space. Every spanning tree satisfies $Q_F^{(\ell)} \preceq Q^{(\ell)}$, proving Eq. (759). The marginal vector of the weighted spanning-tree distribution initializes the pipage-rounding procedure.

To bound $t_{d_{\mathrm{blk}}}$, set $a := \log(2d_{\mathrm{blk}})$. For $a \to \infty$, the integer $t = \lceil 2a/\log a\rceil$ satisfies $t\log t - t + 1 = (2 + o(1))a$, so it is feasible for all sufficiently large d_{blk}. The remaining bounded range is absorbed into an absolute constant. Since $\log\log(4d_{\mathrm{blk}}) > 0$ for $d_{\mathrm{blk}} \ge 1$, this proves $t_{d_{\mathrm{blk}}} = O(\log(2d_{\mathrm{blk}})/\log\log(4d_{\mathrm{blk}}))$ for every $d_{\mathrm{blk}} \ge 1$. Also $d_{\mathrm{blk}} = k(N-1)$ lies between $Nk/2$ and Nk, giving the stated bound in terms of N and k.

The marginals and normalized matrices are computable by polynomially many linear-algebra operations. The graphic matroid has a polynomial-time independence oracle, so the deterministic matrix pipage-rounding algorithm applies in polynomial time. Parallel edges are distinct matroid elements and cause no change to this argument. Finally, choose the construction with the smaller guaranteed bound between this one and the deterministic construction of Theorem 13.18. This choice depends only on d_{blk} and k and proves Eq. (760). □

14 Row-constrained discrepancy and fractional-coloring walks

14.1 Subisotropic inputs

After normalization by the largest row sum, the diagonal matrices associated with the columns of a nonnegative matrix have a subisotropic sum. The following extension supplies the signing bound for this normalization.

Lemma 14.1 (Subisotropic signing via Schatten norms)**.** *Let* $A_1, \dots, A_{n_{\mathrm{col}}} \succeq 0$ *satisfy* $S := \sum_{i=1}^{n_{\mathrm{col}}} A_i \preceq I_d$, $\|A_i\| \le \alpha$, *and* $\operatorname{rank}(A_i) \le r$, *where* $r \ge 1$. *There are signs with*

$$\|\sum_{i=1}^{n_{\mathrm{col}}} \sigma_i A_i\| \le \min\{\|S\|, 3.48745\sqrt{\alpha \log(2r)}\}. \tag{766}$$

If $\alpha \ge 1/n_{\mathrm{col}}$, *the deterministic construction of Theorem 12.15 gives Eq. (766) with coefficient* 4.171 *in place of* 3.48745, *using* $\widetilde{O}(n_{\mathrm{col}} d^2 + d^7)$ *real arithmetic operations.*

Proof. Corollary 11.22 proves the existence bound, including the cap by $\|S\|$. For the algorithmic argument, define $\sigma_{\mathrm{tr}}^2 := \|\sum_{i=1}^{n_{\mathrm{col}}} \operatorname{tr}[A_i] A_i\|$. Since $\operatorname{tr}[A_i] \le r\alpha$ and $S \preceq I_d$,

$$\sigma_{\mathrm{tr}}^2 \le r\alpha \|S\| \le r\alpha. \tag{767}$$

The initial coupling estimate also remains valid:

$$\sum_{i=1}^{n_{\mathrm{col}}} \|A_i\|_{S_p} A_i \preceq r^{1/p} \alpha S \preceq r^{1/p} \alpha I_d. \tag{768}$$

For the algorithm with coefficient 4.171, assume $\alpha \ge 1/n_{\mathrm{col}}$. This condition holds whenever $\|S\| = 1$, since $1 \le \sum_{i=1}^{n_{\mathrm{col}}} \|A_i\| \le n_{\mathrm{col}}\alpha$. If $4.171\sqrt{\alpha \log(2r)} \ge 1$, every signing satisfies the claimed bound. Otherwise $\alpha < 1$. Every $A_i \preceq S \preceq I_d$ and $\|\sum_{i=1}^{n_{\mathrm{col}}} y_i A_i\| \le \|S\| \le 1$ throughout the cube. The normalized inequalities in Eq. (625) therefore hold with $\sum_{i=1}^{n_{\mathrm{col}}} A_i \preceq I_d$. Exact grouped reduction still preserves $\sum_{i=1}^{n_{\mathrm{col}}} y_i A_i = 0$ and leaves at most d^2 fractional inputs. The entropy-smoothed potential, weighted response estimates, and deferred-rounding variance bound use only these upper bounds, so the analysis in Section 12 applies unchanged. The quadratic-growth proof uses the same dual stationarity equations, and its smoothness bound uses only the normalized coupling upper bound. Thus the faster inner solve also applies. Use the parameter choices, reservoir choice, and quantitative curvature margins in Eq. (624), Eq. (649), and Lemma 12.2. The perturbed projection and its quadratic outer-gradient error use the same individual-input identities and dual lower bounds, so they also hold under this upper normalization. The final error allowance in the proof of Theorem 12.15 is unchanged. For $1 \le r \le 2$, use Theorem 9.1, which applies to arbitrary PSD inputs with the variance bound in Eq. (767). Its operation count is within the stated bound. The universal bound $\|\sum_{i=1}^{n_{\mathrm{col}}} \sigma_i A_i\| \le \|S\|$ supplies the cap. □

14.2 The coloring bound

Representing each column by a diagonal matrix turns its support bound into a rank bound. Normalization by the largest row sum allows us to apply Lemma 14.1.

Theorem 14.2 (Formal version of Theorem 1.10)**.** *Let* $W \in [0, \theta]^{m_{\mathrm{row}} \times n_{\mathrm{col}}}$, *where* $\theta, \Lambda > 0$, *have row sums at most* Λ *and at most* $d_{\mathrm{col}} \ge 1$ *nonzero entries in each column. There are signs* $\sigma \in \{-1, 1\}^{n_{\mathrm{col}}}$ *satisfying*

$$\|W\sigma\|_\infty \le \min\{\Lambda, 3.48745\sqrt{\theta\Lambda \log(2d_{\mathrm{col}})}\}.$$

With coefficient 4.171 *in place of* 3.48745*, signs can be found deterministically by the signing construction below, using* $\widetilde{O}(n_{\mathrm{col}}m_{\mathrm{row}}+m_{\mathrm{row}}^{3.38})$ *real arithmetic operations. More generally, for every admissible square matrix multiplication exponent* $\omega \in [2,3]$*, the operation count is* $\widetilde{O}(n_{\mathrm{col}}m_{\mathrm{row}}+m_{\mathrm{row}}^{\omega+1})$*. The current bound* $\omega < 2.371177$ *of Dupont et al. [DEK+26] gives* $\widetilde{O}(n_{\mathrm{col}}m_{\mathrm{row}}+m_{\mathrm{row}}^{3.371177})$*, implying the rounded bound above. Ordinary matrix multiplication gives* $\widetilde{O}(n_{\mathrm{col}}m_{\mathrm{row}}+m_{\mathrm{row}}^{4})$*. The same discrepancy guarantee holds after restriction to any subset of columns. In particular, a zero-one incidence matrix with at most* $d_0 \geq 1$ *ones in every row and every column admits discrepancy at most* $\min\{d_0, 3.48745\sqrt{d_0 \log(2d_0)}\}$*.*

Proof. If $W = 0$, every signing has zero discrepancy. Otherwise, define $\Lambda_{\mathrm{act}} := \max_j \sum_{i=1}^{n_{\mathrm{col}}} W_{ji} > 0$ and $\theta_0 := \max_{j,i} W_{ji} > 0$. For each column, set

$$A_i := \Lambda_{\mathrm{act}}^{-1} \operatorname{diag}(W_{1i}, \ldots, W_{m_{\mathrm{row}}i}). \tag{769}$$

The diagonal entries of $S := \sum_{i=1}^{n_{\mathrm{col}}} A_i$ are the normalized row sums. Hence $S \preceq I_{m_{\mathrm{row}}}$ and $\|S\| = 1$. Each matrix satisfies $\|A_i\| \leq \theta_0/\Lambda_{\mathrm{act}}$ and $\operatorname{rank}(A_i) \leq d_{\mathrm{col}}$ by the entry and column support bounds. Lemma 14.1 therefore applies with $\alpha = \theta_0/\Lambda_{\mathrm{act}}$. The diagonal representation gives

$$\|\sum_{i=1}^{n_{\mathrm{col}}} \sigma_i A_i\| = \Lambda_{\mathrm{act}}^{-1} \|W\sigma\|_\infty. \tag{770}$$

We have

$$\begin{aligned}
\|W\sigma\|_\infty &= \Lambda_{\mathrm{act}} \|\sum_{i=1}^{n_{\mathrm{col}}} \sigma_i A_i\| \\
&\leq \Lambda_{\mathrm{act}} \min\{1, 3.48745\sqrt{(\theta_0/\Lambda_{\mathrm{act}}) \log(2d_{\mathrm{col}})}\} \\
&= \min\{\Lambda_{\mathrm{act}}, 3.48745\sqrt{\theta_0 \Lambda_{\mathrm{act}} \log(2d_{\mathrm{col}})}\} \\
&\leq \min\{\Lambda, 3.48745\sqrt{\theta \Lambda \log(2d_{\mathrm{col}})}\}.
\end{aligned} \tag{771}$$

The first step uses Eq. (770). The second applies Lemma 14.1. The third distributes the positive factor Λ_{act}. The last uses $\Lambda_{\mathrm{act}} \leq \Lambda$ and $\theta_0 \leq \theta$.

For the deterministic bound with coefficient 4.171, $\theta_0 \leq \Lambda_{\mathrm{act}} \leq n_{\mathrm{col}}\theta_0$, so the normalized instance satisfies all the algorithmic normalization bounds. Computing $\Lambda_{\mathrm{act}}, \theta_0$ and the normalized diagonal entries takes $O(n_{\mathrm{col}}m_{\mathrm{row}})$ operations. If $d_{\mathrm{col}} \geq m_{\mathrm{row}}^{1/4}$, Lemma 14.3 gives the claimed signing in $\widetilde{O}(n_{\mathrm{col}}m_{\mathrm{row}})$ operations. We henceforth assume $d_{\mathrm{col}} < m_{\mathrm{row}}^{1/4}$ and implement the grouped reduction of Section 12.1 in the m_{row}-dimensional space of diagonal matrices.

Grouped reduction. Pad the columns by zeros to a power of two and form a balanced binary tree. For each node G, store the column sum $g_G := \sum_{i \in G} W_{:i}$. There are $O(n_{\mathrm{col}})$ nodes, so forming all these sums costs $O(n_{\mathrm{col}}m_{\mathrm{row}})$. A group coefficient is assigned to every column in that group. Initially the root has coefficient zero. At each level, split the at most m_{row} fractional groups, giving both children the coefficient of their parent. This preserves the signed sum and leaves at most $2m_{\mathrm{row}}$ fractional groups. If more than m_{row} remain, find a nonzero dependence among their vectors g_G and move along it until a coefficient reaches an endpoint of $[-1, 1]$. The signed sum is unchanged, and at least one group is fixed. Repeat until at most m_{row} groups remain fractional. A group fixed at a sign is not split again; all its original columns receive that sign.

Each dependence can be found by exact rank-revealing elimination and a triangular solve in $O(m_{\mathrm{row}}^{\omega})$ operations [JPS11]. At most m_{row} moves are needed at a level. Thus recomputing the

dependence after every move costs $O(m_{\mathrm{row}}^{\omega+1})$ per level. After $O(\log(2n_{\mathrm{col}}))$ levels, the remaining fractional groups are original columns. The reduction costs $\widetilde{O}(n_{\mathrm{col}}m_{\mathrm{row}} + m_{\mathrm{row}}^{\omega+1})$ and returns

$$y \in [-1,1]^{n_{\mathrm{col}}}, \qquad Wy = 0, \qquad F := \{i : |y_i| < 1\}, \qquad |F| \le m_{\mathrm{row}}. \tag{772}$$

When $n_{\mathrm{col}} \le m_{\mathrm{row}}$, take $y = 0$ directly. Zero columns and simultaneous endpoint hits do not affect the argument.

Completion from the reduced state. Retain the coefficients y_i for $i \in F$ and the fixed contribution $M_{\mathrm{fix}} := \sum_{i \notin F} y_i A_i$. Eq. (772) gives $M_{\mathrm{fix}} + \sum_{i \in F} y_i A_i = 0$. The completion algorithm in Section 12 already starts from inherited fractional coefficients and keeps the fixed contribution in its signed sum. Its active-count requirement is satisfied since $|F| \le m_{\mathrm{row}} \le m_{\mathrm{row}}^2$. For each Schatten order p, put $f_i(X) := \|A_i^{1/2} X A_i^{1/2}\|_{S_p}$. Every budget used by the construction satisfies $0 \le \beta \le 1$ and $\beta(\pm 1) = 0$, so

$$\sum_{i \in F} \beta(y_i) f_i(I_{m_{\mathrm{row}}}) A_i \preceq \sum_{i=1}^{n_{\mathrm{col}}} f_i(I_{m_{\mathrm{row}}}) A_i. \tag{773}$$

Consequently the reduced state has zero signed sum and the same initial coupling upper bound. The retained matrices are original inputs, so their rank bound remains d_{col}; the group sums are used only during the reduction. Lemma 14.4 completes this state with coefficient 4.171 using $\widetilde{O}(m_{\mathrm{row}}^{\omega+1} + m_{\mathrm{row}}^{11/4} d_{\mathrm{col}})$ operations. It uses the scalar form of the inner problem and a reservoir chosen for at most m_{row} active columns. For support bounds below three, it uses a direct completion that preserves the rows with many active entries.

Restoring the signs of all fixed groups costs $O(n_{\mathrm{col}})$, and computing $W\sigma$ costs $O(n_{\mathrm{col}}m_{\mathrm{row}})$. Hence the total operation count is

$$\widetilde{O}(n_{\mathrm{col}}m_{\mathrm{row}} + m_{\mathrm{row}}^{\omega+1}) + \widetilde{O}(m_{\mathrm{row}}^{\omega+1} + m_{\mathrm{row}}^{11/4} d_{\mathrm{col}}) + O(n_{\mathrm{col}}m_{\mathrm{row}}) = \widetilde{O}(n_{\mathrm{col}}m_{\mathrm{row}} + m_{\mathrm{row}}^{\omega+1}), \tag{774}$$

where $d_{\mathrm{col}} < m_{\mathrm{row}}^{1/4}$ and $\omega \ge 2$ imply $m_{\mathrm{row}}^{11/4} d_{\mathrm{col}} \le m_{\mathrm{row}}^3 \le m_{\mathrm{row}}^{\omega+1}$. Ordinary matrix multiplication gives $\widetilde{O}(n_{\mathrm{col}}m_{\mathrm{row}} + m_{\mathrm{row}}^4)$. The current bound $\omega < 2.371177$ of Dupont et al. [DEK+26] gives $\widetilde{O}(n_{\mathrm{col}}m_{\mathrm{row}} + m_{\mathrm{row}}^{3.371177})$. This proves the deterministic operation bound.

Restricting to any subset of columns preserves the entry bound, row-sum bound, and column support bound. Repeating the same argument therefore proves the restriction assertion. For a zero-one matrix with at most d_0 ones in each row and each column, take $\theta = 1$, $\Lambda = d_0$, and $d_{\mathrm{col}} = d_0$. □

14.3 The scalar potential and the walk

A fractional coloring is a vector $z \in [-1,1]^{n_{\mathrm{col}}}$; a completed coloring has $z \in \{-1,1\}^{n_{\mathrm{col}}}$. The algorithm starts at the zero coloring and fixes original elements one by one or moves several fractional coordinates together, using the gradient and Hessian of the optimized potential.

Assume $W \ne 0$ and define $\Lambda_{\mathrm{act}} := \max_j \sum_{i=1}^{n_{\mathrm{col}}} W_{ji}$. Set $\overline{W}_{ji} := W_{ji}/\Lambda_{\mathrm{act}}$ and $\zeta_j(z) := \sum_{i=1}^{n_{\mathrm{col}}} \overline{W}_{ji} z_i$. Let I be the active set of fractional coordinates; $\zeta_j(z)$ includes the contributions from all fixed coordinates. Choose the Schatten order p as in Eq. (601), with rank bound $r = d_{\mathrm{col}}$, and set $\beta(t) := \sqrt{1-t^2}$ and $\iota := 16(p+3)$. For $u \in (0,\infty)^{m_{\mathrm{row}}}$, define

$$f_i(u) := \Big(\sum_{j=1}^{m_{\mathrm{row}}} (\overline{W}_{ji} u_j)^p\Big)^{1/p}. \tag{775}$$

For $\rho > 0$, the baseline matrix potential has the scalar representation

$$P_\rho(z) = \min_{q\in\mathbb{R},\ u,v\in(0,\infty)^{m_{\mathrm{row}}}} \{q + \rho \sum_{j=1}^{m_{\mathrm{row}}} (u_j + v_j)\},$$
$$\text{subject to } u_j^{-1} + \zeta_j(z) + \iota \sum_{i\in I} \beta(z_i)\overline{W}_{ji} f_i(v) \le q,$$
$$v_j^{-1} - \zeta_j(z) + \iota \sum_{i\in I} \beta(z_i)\overline{W}_{ji} f_i(u) \le q \quad (1 \le j \le m_{\mathrm{row}}). \tag{776}$$

To justify equality with the matrix program, average a feasible primal pair (U, V) over conjugation by diagonal sign matrices. All A_i and $M(z)$ commute with these conjugations, so feasibility and objective are invariant. The feasible set is convex, since the inverse and every Schatten-norm functional have the convexity used in Section 8.2. Averaging therefore gives a diagonal feasible pair with the same objective. Conversely every feasible scalar pair is feasible for the matrix program. On a diagonal matrix the functional is the ℓ_p norm in Eq. (775). This proves Eq. (776).

For coefficient 4.171 when $d_{\mathrm{col}} \ge 3$, use the retuned Schatten orders and budgets in Eq. (624), with the rank bound $r = d_{\mathrm{col}}$, and the final error allowance in the proof of Theorem 12.15. For $1 \le d_{\mathrm{col}} < 3$, use the direct completion in Lemma 14.4. For the faster Schatten construction, use the normalized entropy-smoothed potential in Eq. (627), with the reservoir chosen in Eq. (784), and the grouped reduction and deferred rounding of Section 12. The same diagonal conjugation averaging applies to that convex minimization; its log-trace-exponential becomes the log-sum-exponential of the $2m_{\mathrm{row}}$ diagonal constraint entries. Thus the faster construction also specializes to this fractional-coloring setting.

Every occurrence of column i uses the same variable z_i and budget $\beta(z_i)$; a move changes all incident rows together. Consequently the endpoint and gradient-or-curvature steps specialize to a walk of fractional colorings of the original ground set. The row-sum bound supplies $S \preceq I_{m_{\mathrm{row}}}$ in the diagonal normalization used in the proof of Theorem 14.2.

14.4 Moment signing for larger column support

For larger column support, a moment bound gives the desired discrepancy guarantee. We fix the signs one at a time using conditional expectations to obtain a deterministic construction.

Lemma 14.3 (Deterministic moment signing)**.** *Let $W \in [0,\theta]^{m_{\mathrm{row}}\times n_{\mathrm{col}}}$ have row sums at most Λ, and suppose each column has at most $d_{\mathrm{col}} \ge m_{\mathrm{row}}^{1/4}$ nonzero entries. A deterministic algorithm finds signs satisfying $\|W\sigma\|_\infty \le \min\{\Lambda, 4.171\sqrt{\theta\Lambda \log(2d_{\mathrm{col}})}\}$ using $\widetilde{O}(n_{\mathrm{col}} m_{\mathrm{row}})$ real arithmetic operations.*

Proof. Every signing satisfies $\|W\sigma\|_\infty \le \Lambda$. Thus arbitrary signs suffice if $4.171\sqrt{\theta\Lambda\log(2d_{\mathrm{col}})} \ge \Lambda$. Otherwise, put $h := \max\{1, \lceil \log_2 m_{\mathrm{row}} \rceil\}$ and let the η_i be independent uniform signs. For each row j, write $S_j := \sum_i \eta_i W_{ji}$. The entry and row-sum bounds imply $\sum_i W_{ji}^2 \le \theta\Lambda$. Comparing the even-moment expansion with that of a centered Gaussian of the same variance gives

$$\mathbb{E}[S_j^{2h}] \le (2h-1)!!(\theta\Lambda)^h. \tag{777}$$

Fix the signs one at a time, always choosing the sign with smaller conditional expectation of $\sum_j S_j^{2h}$. The resulting signing satisfies

$$\|W\sigma\|_\infty^2 \le [m_{\mathrm{row}}(2h-1)!!]^{1/h}\theta\Lambda \le \frac{8h}{e}\theta\Lambda. \tag{778}$$

The last inequality uses $(2h-1)!! \le 2(2h/e)^h$ and $m_{\mathrm{row}}^{1/h} \le 2$. Since $h \le 1 + \log_2 m_{\mathrm{row}}$ and $d_{\mathrm{col}} \ge m_{\mathrm{row}}^{1/4}$,

$$\frac{8h}{e} \le \frac{8}{e}(1 + \log_2 m_{\mathrm{row}}) \le (4.171)^2(\log 2 + \tfrac{1}{4}\log m_{\mathrm{row}}) \le (4.171)^2 \log(2d_{\mathrm{col}}). \tag{779}$$

The middle inequality follows from $8/e < (4.171)^2 \log 2$ and $8/(e \log 2) < (4.171)^2/4$.

For the operation count, precompute for every row and every suffix the moments through degree $2h$ of the random signed sum on that suffix. The binomial recurrence for adjoining one uniform sign uses $O(h^2)$ arithmetic operations per row and column. Given the current partial row sum, either conditional value in the greedy choice is a degree-$2h$ polynomial evaluated from the stored suffix moments in $O(h)$ operations per row. Hence the total work is $O(n_{\mathrm{col}} m_{\mathrm{row}} h^2) = \widetilde{O}(n_{\mathrm{col}} m_{\mathrm{row}})$. □

14.5 Faster completion for diagonal inputs

The reduction in Eq. (772) leaves at most m_{row} fractional columns. We use this bound both in the scalar implementation of the potential and in the choice of its reservoir.

Lemma 14.4 (Deterministic completion of a reduced coloring)**.** *Use the normalization in Eq. (769), put $\alpha := \theta_0/\Lambda_{\mathrm{act}}$, and let y satisfy Eq. (772). Given the remaining fractional columns, their coefficients, and the fixed signed contribution, a deterministic algorithm computes signs for the remaining columns such that the complete signing satisfies*

$$\|\sum_{i=1}^{n_{\mathrm{col}}} \sigma_i A_i\| \le \min\{1, 4.171\sqrt{\alpha \log(2d_{\mathrm{col}})}\}. \tag{780}$$

For every admissible square matrix multiplication exponent $\omega \in [2,3]$, its operation count is

$$\widetilde{O}(m_{\mathrm{row}}^{\omega+1} + m_{\mathrm{row}}^{11/4} d_{\mathrm{col}}). \tag{781}$$

Proof. The case with no remaining fractional column is immediate. Let $q \le m_{\mathrm{row}}$ denote the number of active columns; this number decreases during the construction. Every completed signing has norm at most one, since $\sum_i A_i \preceq I_{m_{\mathrm{row}}}$. If the other bound in Eq. (780) is at least one, arbitrary signs therefore suffice.

Columns of support at most two. For $1 \le d_{\mathrm{col}} < 3$, put $r := \lfloor d_{\mathrm{col}} \rfloor \in \{1,2\}$. Maintain exactly the row sums in rows with more than r active nonzero entries. There are at most $rq/(r+1) < q$ such rows, since each active column has at most r nonzero entries. Find a nonzero kernel direction for these row constraints and move along it until a coefficient reaches an endpoint of $[-1,1]$. Repeat until all columns are fixed. There are at most m_{row} moves. A row has signed sum zero until it is released, by Eq. (772); if it is never maintained, its initial signed sum is zero. At release at most r active entries remain in that row, each at most α. Each remaining coefficient changes by at most two, so its final discrepancy is at most $2r\alpha$. Since $0 < \alpha \le 1$,

$$2r\alpha \le 4.171\sqrt{\alpha \log(2d_{\mathrm{col}})} \qquad (r = 1, 2). \tag{782}$$

For $r = 1$, this follows from $4 < (4.171)^2 \log 2$; for $r = 2$, it follows from $16 < (4.171)^2 \log 4$. Exact rank-revealing elimination and a triangular solve find a kernel direction in $O(m_{\mathrm{row}}^{\omega})$ operations per move [JPS11]. Updating the active row counts costs $O(m_{\mathrm{row}}^2)$ per move. Thus this branch costs $O(m_{\mathrm{row}}^{\omega+1})$.

The scalar inner problem. Assume henceforth that $d_{\mathrm{col}} \geq 3$. Use the Schatten-order and budget parameters of Eq. (624) and the rescaling $C_i := A_i/s$ from Section 12.1. Write $C_i = \mathrm{diag}(c_{1i}, \ldots, c_{m_{\mathrm{row}}i})$. The entropy-smoothed inner problem in Eq. (627) is invariant under simultaneous conjugation by diagonal sign matrices. Convexity and averaging show that its optimum is attained at diagonal U, V. Uniqueness makes the optimizer diagonal for every fractional state. Hence the scalar restriction agrees with the full optimized function, including its outer gradient and Hessian. For $U = \mathrm{diag}(u)$ the functional becomes

$$f_i(u) = (\sum_{j=1}^{m_{\mathrm{row}}} c_{ji}^p u_j^p)^{1/p}. \tag{783}$$

There are $2m_{\mathrm{row}}$ positive primal variables and $2m_{\mathrm{row}}$ scalar constraints. Keep $\rho = c/m_{\mathrm{row}}$, $\tau = c/\log(4m_{\mathrm{row}})$, and the endpoint strip and small-input cutoff from Section 12. The constant $c > 0$ is decreased as needed within the same final discrepancy allowance.

A reservoir for at most m_{row} active columns. Let $A_{m_{\mathrm{row}}}, B_{m_{\mathrm{row}}}$ be the fixed polylogarithmic factors used in the proof of Lemma 12.2. Choose

$$\lambda := \frac{c}{m_{\mathrm{row}}(1 + 16A_{m_{\mathrm{row}}}B_{m_{\mathrm{row}}})}. \tag{784}$$

Then $q\lambda \leq c$, so the initial reservoir contributes at most c. Eq. (773) gives the same initial coupling bound, and the initial optimized potential is $2 + O(c)$.

We check the absorption estimate in the weighted alternative. Use its notation $b_i = -\beta_i'' \geq b_* > 0$, $\ell_i = g_i - \lambda\beta_i'$, and $\omega_i = \iota b_i S_i + \lambda b_i$. Eq. (656) gives a nonzero $K \succeq 0$ with

$$\mathrm{tr}[K\nabla^2 F] \leq \sum_i [-d_* S_i/2 - \lambda b_i + m_{\mathrm{row}} A_{m_{\mathrm{row}}} \ell_i^2] K_{ii}. \tag{785}$$

The coefficients in this estimate depend on the bounded inner sublevel and the retained strip. The reservoir is constant in the inner variables and nonnegative in the sublevel bound. Thus the proof is uniform when $q\lambda \leq c$; the condition $m_{\mathrm{row}}^2\lambda \leq c$ in the general construction ensures this same bound for up to m_{row}^2 active coordinates. Moreover $(\beta_i')^2 \leq B_{m_{\mathrm{row}}} b_i$ gives

$$2m_{\mathrm{row}} A_{m_{\mathrm{row}}} \lambda^2 (\beta_i')^2 \leq 2m_{\mathrm{row}} A_{m_{\mathrm{row}}} B_{m_{\mathrm{row}}} \lambda^2 b_i \leq \lambda b_i/8. \tag{786}$$

The last inequality follows from $2m_{\mathrm{row}} A_{m_{\mathrm{row}}} B_{m_{\mathrm{row}}} \lambda = 2cA_{m_{\mathrm{row}}} B_{m_{\mathrm{row}}}/(1 + 16A_{m_{\mathrm{row}}} B_{m_{\mathrm{row}}}) \leq 1/8$. The remainder of the weighted-alternative proof therefore applies. Writing P for a sufficiently large fixed polynomial in $\log(4m_{\mathrm{row}})$, we obtain

$$\Omega \succeq (Pm_{\mathrm{row}})^{-1} I, \qquad D_i \leq P(\omega_i + m_{\mathrm{row}} g_i^2). \tag{787}$$

Either some $|g_i| > \sqrt{\omega_i}/(P\sqrt{m_{\mathrm{row}}})$, or $\Omega^{-1/2}\nabla^2 F \Omega^{-1/2}$ has an eigenvalue at most $-1/P$. The polynomial P may be enlarged in the estimates below.

The number of outer steps. In the curvature branch take $h^\top \Omega h = 1$ and $h^\top \nabla^2 F h \leq -1/P$, reserving a fixed part of the margin for numerical error. Eq. (787) gives $\|h\|_\infty \leq P\sqrt{m_{\mathrm{row}}}$, and Eq. (666) gives $R \leq P\sqrt{m_{\mathrm{row}}}$. Choose $t = 1/(P\sqrt{m_{\mathrm{row}}})$ with P large enough to preserve the endpoint strip. Apply Eq. (669) to both orientations. The linear terms cancel in their average, and its remainder terms are absorbed by the negative quadratic term. One orientation decreases the potential by at least $1/(Pm_{\mathrm{row}})$.

For a coordinate gradient step put $\vartheta = \omega_i$ and $T = P(\vartheta + m_{\mathrm{row}} g_i^2)$. The response estimate and step rule in the proof of Lemma 12.5 give a decrease bounded below by a fixed multiple of the minimum of

$$\delta|g_i|, \qquad \frac{g_i^2}{PT}, \qquad \frac{|g_i|}{P(1+\sqrt{m_{\mathrm{row}}T})}. \tag{788}$$

Failure of the gradient test, together with Eq. (787), implies $|g_i| \geq 1/(Pm_{\mathrm{row}})$ and $g_i^2/T \geq 1/(Pm_{\mathrm{row}})$ after enlarging P. The first two quantities in Eq. (788) are therefore at least $1/(Pm_{\mathrm{row}})$. For the third, use the first bound if $\sqrt{m_{\mathrm{row}}T} \leq 1$ and the second otherwise. It has the same lower bound after another enlargement of P.

There are at most m_{row} boundary events. Keep the original endpoint accuracy $\eta = c/m_{\mathrm{row}}^2$; their total increase is $O(c)$. The initial potential is $O(1)$ and it remains nonnegative, so there are $\widetilde{O}(m_{\mathrm{row}})$ outer iterations and boundary events. Inverse-polynomial evaluation errors and separated comparisons preserve a fixed fraction of each decrease.

Scalar inner solves. At each current or trial state, recompute the inner optimum by the restarted projected method of Section 12.3. No warm correction is needed for this operation bound. Lemma 12.6 restricts to diagonal pairs with quadratic-growth constant of order $m_{\mathrm{row}}^{-3/2}$ and polylogarithmic Lipschitz constant. Thus each solve uses $\widetilde{O}(m_{\mathrm{row}}^{3/4})$ value and gradient queries. A query evaluates all $q \leq m_{\mathrm{row}}$ scalar functionals and their gradients, forms the $2m_{\mathrm{row}}$ constraints and their Gibbs weights, and assembles the primal gradient. Each functional depends on at most d_{col} variables, and its gradient has at most d_{col} nonzero entries. Scattering their contributions to the scalar constraints and primal gradient therefore costs $\widetilde{O}(qd_{\mathrm{col}})$; the Gibbs weights and remaining scalar work cost $\widetilde{O}(m_{\mathrm{row}})$. Thus a query costs $\widetilde{O}(m_{\mathrm{row}} + qd_{\mathrm{col}})$ operations. Projection clips $2m_{\mathrm{row}}$ scalar coordinates to $[a_{\min}, C_0\sqrt{m_{\mathrm{row}}}]$ and costs $O(m_{\mathrm{row}})$. One complete inner solve therefore costs

$$\widetilde{O}(m_{\mathrm{row}}^{3/4}(m_{\mathrm{row}} + qd_{\mathrm{col}})). \tag{789}$$

The small-input cutoff bounds every retained functional value away from zero by an inverse polynomial. The Schatten order is $O(\log m_{\mathrm{row}})$, so scalar powers, roots, and exponentials can be evaluated to the required inverse-polynomial accuracies with polylogarithmic overhead. The growth estimate converts objective accuracy to optimizer accuracy; the root bounds then control the errors in the derivatives below.

Derivative assembly and direction selection. Let $g_i^u := \nabla f_i(u)$. Differentiating Eq. (783) gives

$$(g_i^u)_j = \frac{c_{ji}^p u_j^{p-1}}{f_i(u)^{p-1}}, \tag{790}$$

$$\nabla^2 f_i(u) = (p-1)\,\mathrm{diag}\Big(\frac{c_{ji}^p u_j^{p-2}}{f_i(u)^{p-1}}\Big) - \frac{p-1}{f_i(u)} g_i^u (g_i^u)^\top. \tag{791}$$

Every dual-weighted sum of functional Hessians is therefore a diagonal matrix minus $G\,\mathrm{diag}(a)G^\top$, where G has size $m_{\mathrm{row}} \times q$. Forming its diagonal part costs $O(m_{\mathrm{row}}q)$; the remaining product costs $O(m_{\mathrm{row}}^\omega)$. The mixed primal and outer blocks use the functional gradients and cost $O(m_{\mathrm{row}}q)$ to assemble.

If w is the vector of $2m_{\mathrm{row}}$ Gibbs weights, the Hessian of scalar log-sum-exponential is

$$\tau^{-1}(\mathrm{diag}(w) - ww^\top).$$

The constraint Jacobian has $2m_{\mathrm{row}}$ rows and $2m_{\mathrm{row}} + q$ columns. Composing with it requires a constant number of products of order $O(m_{\mathrm{row}})$. The other chain-rule term is the weighted sum of constraint Hessians just described. Thus all joint Hessian blocks cost $\widetilde{O}(m_{\mathrm{row}}^{\omega})$ to form.

Write $x = (u, v)$ and let $\phi(x, y)$ be the inner objective without the reservoir. At its optimizer the outer Hessian is

$$\nabla^2 F = \phi_{yy} - \phi_{yx}\phi_{xx}^{-1}\phi_{xy} + \lambda \operatorname{diag}(\beta_i''). \tag{792}$$

The positive definite primal block has order $2m_{\mathrm{row}}$ and polynomially bounded inverse by the root and inverse-Hessian estimates of Section 12. Its inverse action and the Schur complement in Eq. (792) can be computed in $\widetilde{O}(m_{\mathrm{row}}^{\omega})$ operations. The matrix-power direction selection in Lemma 12.7 acts on an at most $m_{\mathrm{row}} \times m_{\mathrm{row}}$ normalized Hessian and has the same operation bound. All inverse and derivative evaluations use the inverse-polynomial accuracies required by the gradient and curvature margins.

Rounding and total work. The small-input and deferred-rounding variance bounds of Section 12.1 are unchanged. There are at most m_{row} deferred coordinates. Since their matrices are diagonal, the two-sided trace-exponential conditional estimator is a sum of scalar row estimators. Precompute its scalar moment factors in $\widetilde{O}(m_{\mathrm{row}}^2)$ operations and maintain their products as each sign is fixed. Each choice costs $\widetilde{O}(m_{\mathrm{row}})$, for total $\widetilde{O}(m_{\mathrm{row}}^2)$ work. The strict estimator margin absorbs the prescribed numerical errors.

The enlarged reservoir still contributes only $O(c)$ to the initial potential. The same Schatten-order and budget parameters and final error allowance therefore give coefficient 4.171. There are $\widetilde{O}(m_{\mathrm{row}})$ current and trial states, each requiring an inner solve and at most $\widetilde{O}(m_{\mathrm{row}}^{\omega})$ derivative and direction work. Including the small-support branch, the total cost is

$$\widetilde{O}(m_{\mathrm{row}})(m_{\mathrm{row}}^{3/4}(m_{\mathrm{row}} + m_{\mathrm{row}} d_{\mathrm{col}}) + m_{\mathrm{row}}^{\omega}) + \widetilde{O}(m_{\mathrm{row}}^2) = \widetilde{O}(m_{\mathrm{row}}^{\omega+1} + m_{\mathrm{row}}^{11/4} d_{\mathrm{col}}). \tag{793}$$

This proves the claimed operation count and discrepancy guarantee. □

15 Lower bounds

All logarithms in this section are natural.

15.1 A finite diagonal construction

With the convention $0 \log 0 := 0$, define

$$\psi(t) := \frac{(1+t)\log(1+t) + (1-t)\log(1-t)}{2}, \qquad 0 \le t \le 1, \tag{794}$$

Expanding $\psi'(t)$ and integrating its geometric series gives

$$\psi(t) = \sum_{j=1}^{\infty} \frac{t^{2j}}{(2j)(2j-1)} \qquad (0 \le t < 1).$$

The series also converges at one, with sum $\log 2$. In particular,

$$\psi(t) \le (\log 2) t^2 \le t^2 \qquad (0 \le t \le 1). \tag{795}$$

Every coefficient after the quadratic term is at most $1/12$. Summing a geometric series therefore gives the sharper estimate

$$\frac{t^2}{2} \le \psi(t) \le \frac{t^2}{2} + \frac{t^4}{12(1-t^2)} \qquad (0 \le t < 1). \tag{796}$$

We encode random subsets as diagonal supports. A lower bound on the probability of a large row sum then allows a union bound over all signings.

Lemma 15.1 (A finite diagonal lower bound)**.** *Let $k \ge 2$ be even, let $h \in \{2, 4, \dots, k\}$, and let n be a positive integer. Suppose*

$$n > 4k^2(k+1)\log 2 \exp(k\psi(h/k)). \tag{797}$$

There are $m := 2k^2$ diagonal positive semidefinite matrices $A_1, \dots, A_m \in \mathbb{R}^{n\times n}$ satisfying

$$\sum_{i=1}^{m} A_i = I_n, \qquad \|A_i\| \le 1/k, \qquad \operatorname{rank}(A_i) \le n,$$

such that every $\sigma \in \{-1, 1\}^m$ satisfies

$$\|\sum_{i=1}^{m} \sigma_i A_i\| \ge h/k. \tag{798}$$

Proof. Independently choose n uniformly random k-element subsets $S_1, \dots, S_n$ of $\{1, \dots, m\}$, allowing repeated subsets. Define

$$A_i := \frac{1}{k} \operatorname{diag}(\mathbf{1}_{\{i\in S_1\}}, \dots, \mathbf{1}_{\{i\in S_n\}}) \qquad (1 \le i \le m).$$

Every matrix is diagonal and positive semidefinite. Each diagonal coordinate belongs to exactly k of the supports, so the isotropy, norm, and rank conditions hold for every outcome.

Fix a signing σ. After negating all its signs if necessary, it has at least $m/2$ positive entries. Select exactly $m/2$ of these positions, and let σ' be positive there and negative elsewhere. Then $\sigma'_i \le \sigma_i$ for every i. A row with sum at least h under σ' therefore has sum at least h under σ. Negating back, if needed, preserves the absolute row sum.

Put $M := m/2 = k^2$ and $a := (k+h)/2$. For the balanced signing σ', the probability that a random row has sum exactly h is

$$q := \frac{\binom{M}{a}\binom{M}{k-a}}{\binom{2M}{k}}.$$

Define $(M)_b := \prod_{j=0}^{b-1}(M-j)$ for $b \ge 1$, with $(M)_0 := 1$. Then

$$q = \binom{k}{a} \frac{(M)_a (M)_{k-a}}{(2M)_k}.$$

The denominator is at most $(2M)^k$. We have

$$\begin{aligned} \frac{(M)_a (M)_{k-a}}{M^k} &\ge 1 - \frac{a(a-1) + (k-a)(k-a-1)}{2M} \\ &\ge 1 - \frac{k(k-1)}{2M} \\ &= \frac{1}{2} + \frac{1}{2k} > \frac{1}{2}. \end{aligned} \tag{799}$$

The first step applies $\prod_j (1-x_j) \ge 1 - \sum_j x_j$ for $0 \le x_j \le 1$ to the two falling factorials. The second uses $a(a-1) + (k-a)(k-a-1) \le k(k-1)$. The last substitutes $M = k^2$ and simplifies; $1/(2k) > 0$.

Substituting Eq. (799) in the formula for q gives

$$q \geq \frac{1}{2}\binom{k}{a}2^{-k}. \tag{800}$$

Set $x := a/k$ and $H(x) := -x\log x - (1-x)\log(1-x)$. When $a < k$, the outcome a is a mode of a binomial distribution with parameters (k, x). Its probability is at least $1/(k+1)$, since the distribution has $k+1$ outcomes. Rearranging gives

$$\binom{k}{a} \geq \frac{\exp(kH(x))}{k+1}. \tag{801}$$

The same inequality is immediate for $a = k$. Since $\log 2 - H((1+t)/2) = \psi(t)$, Eq. (800) implies

$$q \geq \frac{\exp(-k\psi(h/k))}{2(k+1)}. \tag{802}$$

This is a lower bound on the probability that an arbitrary fixed signing has an absolute row sum at least h.

Let E be the event that some signing has discrepancy below h/k. We have

$$\begin{aligned}\Pr[E] &\leq 2^m(1-q)^n \\ &\leq \exp(m\log 2 - nq) \\ &\leq \exp(m\log 2 - n\exp(-k\psi(h/k))/(2(k+1))) \\ &< 1.\end{aligned} \tag{803}$$

The first step uses independence of the rows and the union bound over all 2^m signings. The second uses $1 - q \leq \exp(-q)$. The third substitutes Eq. (802). The last uses Eq. (797), with $m = 2k^2$.

Some choice of the rows therefore satisfies Eq. (798) for every signing. □

Using both tails of the row distribution reduces the number of matrices from $2k^2$ to $16k$. A comparison with binomial probabilities gives the stronger finite estimate needed for the uniform parameter range.

Lemma 15.2 (A sharper finite diagonal lower bound)**.** *Let $k \geq 2$ and $r \geq 1$ be integers, and let $1 \leq h \leq k$ have the same parity as k. Put $a := (k+h)/2$ and*

$$p_{k,h} := \frac{\binom{8k}{a}\binom{8k}{k-a}}{\binom{16k}{k}}. \tag{804}$$

If $rp_{k,h} > 8k\log 2$, there are $m := 16k$ diagonal positive semidefinite matrices $A_1, \dots, A_m \in \mathbb{R}^{r\times r}$ summing to I_r, with norms at most $1/k$, such that every signing has norm at least h/k. Moreover,

$$\begin{aligned}p_{k,h} &\geq \frac{\exp(-k\psi(h/k) - h^2/(28k))}{2\sqrt{k}} \\ &\geq \frac{\exp(-73h^2/(100k))}{2\sqrt{k}}.\end{aligned} \tag{805}$$

Proof. Choose r independent uniformly random k-element subsets $S_1, \dots, S_r$ of $[16k]$, and define

$$A_i := \frac{1}{k}\operatorname{diag}(\mathbf{1}_{\{i\in S_1\}}, \dots, \mathbf{1}_{\{i\in S_r\}}) \qquad (1 \leq i \leq 16k).$$

The positivity, isotropy, and norm conditions hold for every outcome.

We first show that the probability that one row has absolute sign sum at least h is minimized by a balanced signing. Let K be the number of positive signs and let $F(K)$ denote this probability. By symmetry, $F(K) = F(16k - K)$. When a negative sign is changed to positive, conditional on selecting that position the remaining $k - 1$ positions form a uniform subset of a population with K positives and $16k - 1 - K$ negatives. Let Y denote their sign sum. The selected position changes the absolute-threshold event only when $Y = h - 1$ or $Y = 1 - h$, so

$$F(K+1) - F(K) = \frac{1}{16}(\Pr[Y = h-1] - \Pr[Y = 1-h]). \tag{806}$$

For $8k \le K < 16k$, the first probability is at least the second. Indeed, put $u := (k + h - 2)/2$ and $v := (k - h)/2$. The two numerators are respectively $\binom{K}{u}\binom{16k-1-K}{v}$ and $\binom{K}{v}\binom{16k-1-K}{u}$. Since $u \ge v$, the ratio $\binom{N}{u}/\binom{N}{v}$ is increasing in $N \ge u$; cases with zero binomial coefficients satisfy the same comparison. Hence F is nondecreasing for $K \ge 8k$, proving the claim. At balance, the row sums h and $-h$ each have probability $p_{k,h}$. Thus every signing has absolute row sum at least h with probability at least $2p_{k,h}$. Let E be the event that some signing has norm below h/k. We have

$$\Pr[E] \le 2^{16k}(1 - 2p_{k,h})^r \le \exp(16k \log 2 - 2rp_{k,h}) < 1. \tag{807}$$

The first step uses independent rows and the union bound over all signings. The second uses $1 - z \le \exp(-z)$. The last uses the hypothesis $rp_{k,h} > 8k \log 2$.

For the probability estimate, set $M := 8k$ and define $(M)_b := \prod_{j=0}^{b-1}(M - j)$, with $(M)_0 := 1$. Let q_a denote the probability that a row of a balanced signing contains a positive positions. Define the ratio

$$f(a) := \frac{q_a}{\binom{k}{a}2^{-k}} = \frac{2^k (M)_a (M)_{k-a}}{(2M)_k}. \tag{808}$$

For $0 \le j < k$, cancellation of the falling factorials gives

$$\frac{f(j)}{f(j+1)} = 1 + \frac{2j + 1 - k}{M - j}. \tag{809}$$

Thus f is maximal at $a_0 := \lceil k/2 \rceil$ and its symmetric counterpart. Since both distributions have total mass one, $f(a_0) \ge 1$. We have

$$\begin{aligned}
\log f(a) &\ge -\sum_{j=a_0}^{a-1} \log(1 + (2j + 1 - k)/(M - j)) \\
&\ge -\sum_{j=a_0}^{a-1} \frac{2j + 1 - k}{M - j} \\
&\ge -\frac{\sum_{j=a_0}^{a-1}(2j + 1 - k)}{M - k} \\
&= -\frac{(a - k/2)^2 - (a_0 - k/2)^2}{M - k} \\
&\ge -\frac{h^2}{28k}.
\end{aligned} \tag{810}$$

The first step telescopes Eq. (809) from $f(a_0) \ge 1$. The second uses $\log(1 + z) \le z$. The third uses $M - j \ge M - k$ and $2j + 1 - k \ge 0$. The fourth sums the arithmetic progression. The last uses $a - k/2 = h/2$, $M - k = 7k$, and $(a_0 - k/2)^2 \ge 0$.

For $1 \le a \le k-1$, set $E := \exp(-k\psi(h/k))/\sqrt{k}$. We have

$$\begin{aligned}\binom{k}{a}2^{-k} &\ge E\sqrt{\frac{k^2}{2\pi a(k-a)}}\exp(-1/(12a)-1/(12(k-a)))\\ &\ge E\sqrt{2/\pi}\exp(-1/(12a)-1/(12(k-a)))\\ &\ge E\sqrt{2/\pi}\exp(-1/6)\\ &> E/2. \end{aligned} \tag{811}$$

The first step uses the Stirling inequalities and the definition of E. The second uses $a(k-a) \le k^2/4$. The third uses $1/a + 1/(k-a) \le 2$. The last uses $\sqrt{2/\pi}\exp(-1/6) > 1/2$.

The endpoint $a = k$ satisfies the same lower bound directly. Since $p_{k,h} = q_a$, Eqs. (808), (810), and (811) give the first inequality in Eq. (805). For the second, Eq. (795) gives $k\psi(h/k)+h^2/(28k) \le (\log 2 + 1/28)h^2/k$. Now use $\log 2 + 1/28 < 73/100$. □

15.2 A lower bound over a parameter range

The sharper construction matches the dependence on α and r in Theorem 1.4, up to an absolute constant, when $r \ge 256\alpha^{-6}$. The all-positive signing sums to the identity, so the cap by one is optimal whenever the lower bound reaches one.

Theorem 15.3 (A matching-order lower bound)**.** *For every $0 < \alpha \le 1/256$ and all integers $n \ge r \ge 256\alpha^{-6}$, there are diagonal positive semidefinite matrices $A_1, \dots, A_m \in \mathbb{R}^{n\times n}$ satisfying*

$$\sum_{i=1}^{m} A_i = I_n, \qquad \|A_i\| \le \alpha, \qquad \operatorname{rank}(A_i) \le r,$$

such that every $\sigma \in \{-1,1\}^m$ satisfies

$$\|\sum_{i=1}^{m} \sigma_i A_i\| \ge \min\{1, \sqrt{\alpha \log(2r)}\}. \tag{812}$$

Proof. We first prove the statement when $n = r$. Set $k := \lceil 1/\alpha \rceil \ge 256$ and $L := \log(2r)$. Then $1/k \le \alpha < 1/(k-1)$, and the rank hypothesis gives

$$r \ge 256(k-1)^6. \tag{813}$$

Define $L_0 := 9\log 2 + 6\log(k-1)$. Taking logarithms in Eq. (813) gives $L \ge L_0$. Set $y := \min\{1, \sqrt{L/(k-1)}\}$, and let h be the smallest positive integer of the same parity as k with $h \ge ky$. Since $ky \le k$, we have $h \le k$. We have

$$\frac{h}{k} \ge y \ge \min\{1, \sqrt{\alpha L}\}. \tag{814}$$

The first step is the choice of h. The second uses $\alpha < 1/(k-1)$ in the definition of y.

Rounding to the prescribed parity increases ky by less than two, so

$$h < k\sqrt{L/(k-1)} + 2. \tag{815}$$

It remains to verify the criterion of Lemma 15.2. Set $a_* := 73/100$ and $b_* := 71/126$. Assume $0 \le t \le 1/2$. We have

$$\begin{aligned}\psi(t) + t^2/28 &\le t^2/2 + t^4/(12(1-t^2)) + t^2/28 \\ &\le (1/2 + 1/36 + 1/28)t^2 \\ &= b_* t^2. \end{aligned} \tag{816}$$

The first step uses Eq. (796). The second uses $t^2/(1-t^2) \le 1/3$. The last follows from simple algebra.

Lemma 15.2 gives

$$p_{k,h} \ge \frac{\exp(-a_* h^2/k)}{2\sqrt{k}}. \tag{817}$$

Combining its first probability estimate with Eq. (816) gives the stronger bound

$$p_{k,h} \ge \frac{\exp(-b_* h^2/k)}{2\sqrt{k}} \qquad \text{if } h/k \le 1/2. \tag{818}$$

For $c \in \{a_*, b_*\}$, define

$$\begin{aligned}F_c(L) := L - \log 2 - \tfrac{3}{2}\log k - \log(16 \log 2) \\ - c[kL/(k-1) + 4\sqrt{L/(k-1)} + 4/k]. \end{aligned} \tag{819}$$

The square of Eq. (815) bounds h^2/k by the bracketed expression. Since $r = \exp(L)/2$, the corresponding probability bound gives $rp_{k,h}/(8k \log 2) > \exp(F_c(L))$. Thus $F_c(L) > 0$ suffices whenever that probability bound applies. Assume $L \ge 1$. We have

$$\begin{aligned}F_c'(L) &= 1 - c\frac{k}{k-1} - \frac{2c}{\sqrt{(k-1)L}} \\ &\ge 1 - c\frac{256}{255} - \frac{2c}{\sqrt{(k-1)L}} \\ &> 1 - c\frac{256}{255} - \frac{2c}{15} \\ &> 0. \end{aligned} \tag{820}$$

The first step differentiates Eq. (819). The second uses $k \ge 256$. The third uses $(k-1)L \ge 255 > 15^2$. The last holds for both $c = 73/100$ and $c = 71/126$.

We use the elementary estimates

$$2/3 < \log 2 < 7/10, \tag{821}$$

$$\log(16 \log 2) < 5/2. \tag{822}$$

The series $\log 2 = 2\sum_{j=0}^{\infty}(1/3)^{2j+1}/(2j+1)$ gives $2/3 < \log 2 \le 25/36 < 7/10$. Also $\exp(5/2) > \sum_{j=0}^{6}(5/2)^j/j! > 56/5 > 16 \log 2$, which proves Eq. (822). Set $L_s := (k-1)(1/2 - 2/k)^2$ and $C_0 := \log 2 + \frac{3}{2}\log k + \log(16 \log 2)$.

The range $L \le L_s$. Eq. (815) gives $h/k \le 1/2$, so Eq. (818) applies. Differentiation shows that $L_0/(k-1)$ decreases for $k \ge 256$. We have

$$\frac{L_0}{k-1} \le \frac{9 \log 2 + 6 \log 255}{255} < \frac{57 \log 2}{255} < \frac{4}{25}. \tag{823}$$

The first step evaluates the decreasing ratio at $k = 256$. The second uses $\log 255 < 8 \log 2$. The last uses Eq. (821).

Set $a_{\text{aux}} := 256b_*/255$, $A := 1 - a_{\text{aux}}$, and $\delta := 4\sqrt{L_0/(k-1)} + 4/k$. Eq. (823) and $k \geq 256$ give $\delta < 4(2/5) + 4/256 = 517/320$. Define $C := 57A - 13 > 0$ and $\eta := 6A/255 + b_*517/320$. We have

$$\begin{aligned} F_{b_*}(L_0) &= (1 - b_*k/(k-1))L_0 - C_0 - b_*\delta \\ &\geq AL_0 - C_0 - b_*\delta \\ &\geq AL_0 - C_0 - b_*517/320 \\ &\geq (9A-1)\log 2 + (6A - 3/2)\log k - \eta - \log(16\log 2) \\ &\geq C\log 2 - \eta - \log(16\log 2). \end{aligned} \tag{824}$$

The first step substitutes L_0 in Eq. (819). The second uses $k/(k-1) \leq 256/255$. The third uses $\delta \leq 517/320$. The fourth substitutes $\log(k-1) \geq \log k - 1/255$ into L_0. The last uses $\log k \geq 8\log 2$, since $6A - 3/2 > 0$.

Monotonicity now transfers this estimate from L_0 to L. We have

$$\begin{aligned} F_{b_*}(L) &\geq F_{b_*}(L_0) \\ &\geq C\log 2 - \eta - \log(16\log 2) \\ &> 2C/3 - \eta - \log(16\log 2) \\ &> 2C/3 - \eta - 5/2 \\ &> 4. \end{aligned} \tag{825}$$

The first step uses Eq. (820) and $L \geq L_0$. The second uses Eq. (824). The third uses Eq. (821) and $C > 0$. The fourth uses Eq. (822). The last substitutes the rational values of C and η.

The range $L > L_s$. Eq. (817) applies for every h. The definition of L_s gives $L_s \geq k/4 - 9/4 > 1$ and $k\sqrt{L_s/(k-1)} + 2 = k/2$. Define $G(k) := 27k/400 - \frac{3}{2}\log k - 9/4 - \log 2 - \log(16\log 2)$. We have

$$\begin{aligned} F_{a_*}(L) &\geq F_{a_*}(L_s) \\ &= L_s - C_0 - a_*k/4 \\ &\geq k/4 - 9/4 - C_0 - a_*k/4 \\ &= G(k). \end{aligned} \tag{826}$$

The first step uses Eq. (820). The second substitutes L_s into Eq. (819). The third uses $L_s \geq k/4 - 9/4$. The last uses $a_* = 73/100$ and the definition of C_0.

Since $G'(k) = 27/400 - 3/(2k) > 0$ for $k \geq 256$, G is increasing. We have

$$\begin{aligned} G(k) &\geq G(256) \\ &= 27\cdot 256/400 - 13\log 2 - 9/4 - \log(16\log 2) \\ &> 27\cdot 256/400 - 13(7/10) - 9/4 - \log(16\log 2) \\ &> 27\cdot 256/400 - 13(7/10) - 9/4 - 5/2 \\ &> 3. \end{aligned} \tag{827}$$

The first step uses the monotonicity of G. The second substitutes $k = 256$. The third uses Eq. (821). The fourth uses Eq. (822). The last follows from simple arithmetic.

Eqs. (825) and (827) verify $rp_{k,h} > 8k\log 2$ in both ranges. Lemma 15.2 supplies matrices with norms at most $1/k \leq \alpha$ and discrepancy at least h/k. Their dimension is r, so their ranks are at most r. Eq. (814) proves Eq. (812) when $n = r$.

For $n > r$, let $B_1, \dots, B_M \in \mathbb{R}^{r\times r}$ be the family just constructed. Define

$$\widetilde{A}_i := B_i \oplus 0_{n-r} \quad (1 \le i \le M), \qquad C_{j,t} := \frac{1}{k} e_j e_j^\top \quad (r+1 \le j \le n,\ 1 \le t \le k),$$

where e_j is the jth standard basis vector of $\mathbb{R}^n$. These matrices are diagonal and positive semidefinite, have norms at most α, and have ranks at most r. Moreover,

$$\sum_{i=1}^{M} \widetilde{A}_i + \sum_{j=r+1}^{n} \sum_{t=1}^{k} C_{j,t} = I_n.$$

For every choice of signs σ_i and $\tau_{j,t}$, compression to the first r coordinates gives

$$\|\sum_{i=1}^{M} \sigma_i \widetilde{A}_i + \sum_{j=r+1}^{n} \sum_{t=1}^{k} \tau_{j,t} C_{j,t}\| \ge \|\sum_{i=1}^{M} \sigma_i B_i\| \ge \min\{1, \sqrt{\alpha \log(2r)}\}.$$

Relabeling the combined family as $A_1, \dots, A_m$, where $m := M + (n-r)k$, completes the proof. □

15.3 The asymptotic leading constant

In the regime $\alpha \log(2r) \to 0$, the quadratic term of ψ determines the leading coefficient. The first diagonal construction then gives a lower bound of $\sqrt{2}$ on any universal coefficient.

Theorem 15.4 (A lower bound on the leading coefficient)**.** *For every $0 < c < \sqrt{2}$, there are infinitely many pairs (α, r) and diagonal positive semidefinite families $A_1, \dots, A_m \in \mathbb{R}^{r\times r}$ satisfying*

$$\sum_{i=1}^{m} A_i = I_r, \qquad \|A_i\| \le \alpha, \qquad \operatorname{rank}(A_i) \le r,$$

such that every $\sigma \in \{-1, 1\}^m$ satisfies

$$\|\sum_{i=1}^{m} \sigma_i A_i\| \ge c\sqrt{\alpha \log(2r)}, \tag{828}$$

while $\alpha \log(2r) \to 0$ along these pairs. Consequently any universal coefficient in the signing bound of Theorem 1.4 is at least $\sqrt{2}$.

Proof. Let s run through sufficiently large even integers. Set $k := s^4$, $\alpha := 1/k$, $r := n := 2^{s^2}$, and $m := 2k^2$. Define $h := 2\lceil c\sqrt{k \log(2r)}/2 \rceil$ and $t := h/k$. For large s, the integer h is positive, even, and smaller than k. Also $t = O_c(1/s)$, so $kt^4 = O_c(1)$. Rounding h upward by less than two changes $h^2/(2k)$ by $O_c(1/s)$. We have

$$\begin{aligned} k\psi(t) &= \frac{h^2}{2k} + O_c(kt^4) \\ &= \frac{c^2}{2} \log(2r) + O_c(1). \end{aligned} \tag{829}$$

The first step applies Eq. (796) with $t = h/k$. The second uses the rounding estimate and $kt^4 = O_c(1)$.

Substituting $r = 2^{s^2}$ in Eq. (829) gives

$$\log r - k\psi(t) = (1 - c^2/2)s^2 \log 2 + O_c(1). \tag{830}$$

Since $1 - c^2/2 > 0$, this exceeds $\log(4k^2(k+1)\log 2) = 12\log s + O(1)$ for all sufficiently large s. Thus Eq. (797) holds, and Lemma 15.1 supplies matrices with discrepancy at least h/k. We have

$$\begin{aligned} h/k &\geq c\sqrt{\log(2r)/k} \\ &= c\sqrt{\alpha \log(2r)}. \end{aligned} \tag{831}$$

The first step uses the definition of h. The second uses $\alpha = 1/k$.

The parameter choices give

$$\alpha \log(2r) = \frac{(s^2+1)\log 2}{s^4} \longrightarrow 0. \tag{832}$$

Each of the n diagonal coordinates belongs to k supports, so $\sum_{i=1}^{m} \operatorname{rank}(A_i) = nk$. For $R := \max_{1\leq i\leq m} \operatorname{rank}(A_i)$, averaging over $m = 2k^2$ matrices gives $n/(2k) \leq R \leq n$. Since $\log k = o(\log n)$ and $r = n$,

$$\frac{\log(2R)}{\log(2r)} \longrightarrow 1. \tag{833}$$

Thus the asymptotic comparison also applies to the actual maximum rank.

If a universal coefficient $C < \sqrt{2}$ were valid, choose $\max\{0, C\} < c < \sqrt{2}$ and apply these examples. The upper bound would eventually be $C\sqrt{\alpha \log(2r)} < 1$, contradicting Eq. (828). Thus $C \geq \sqrt{2}$. □

16 Lower bounds for simultaneous spectral thinness

All logarithms in this section are natural. Every retained tree edge keeps its original weight in each weighting.

The construction assigns a weighting to each balanced cut and each selected set of crossing edges. Every spanning tree then has a cut whose weight is concentrated on its retained edges in one weighting.

Theorem 16.1 (A finite simultaneous thin-tree lower bound)**.** *For every pair of integers $h \geq 2$ and $M \geq 1$, define*

$$k_h := \binom{4h-1}{2h-1}\binom{4h^2}{h}, \tag{834}$$

$$\alpha_{h,M} := \frac{1}{hM}, \tag{835}$$

$$a_{h,M} := \frac{h}{M(h+1-1/(4h))}. \tag{836}$$

There is a connected undirected loopless multigraph $G = (V, E)$ on $4h$ vertices with k_h positive weightings of the same edge set. Let $Q^{(\ell)}$ be the graph Laplacian under weighting ℓ. Every edge satisfies

$$\tau_e^{(\ell)} u_e^* (Q^{(\ell)})^\dagger u_e \leq \alpha_{h,M} \qquad (e \in E,\ \ell \in [k_h]). \tag{837}$$

For every spanning tree $F \subseteq E$, there are a weighting $\ell \in [k_h]$ and a nonempty proper subset $S \subset V$ with

$$\frac{\mathbf{1}_S^* Q_F^{(\ell)} \mathbf{1}_S}{\mathbf{1}_S^* Q^{(\ell)} \mathbf{1}_S} \geq a_{h,M}. \tag{838}$$

In particular, simultaneous θ_{thin}-spectral thinness requires $\theta_{\text{thin}} \geq a_{h,M}$, and

$$\frac{a_{h,M}}{\alpha_{h,M}} = \frac{h^2}{h+1-1/(4h)} \geq h-1. \tag{839}$$

For $M=1$, the graph is simple.

Proof. We first construct the instance for $M=1$, on the complete graph $G_0 = K_{4h}$ with vertex set $[4h]$. For $S \subseteq [4h]$, let $\delta(S)$ denote the edges with exactly one endpoint in S. Index the weightings by pairs $(S,\mathcal{E})$, where $|S| = 2h$, $1 \in S$, and $\mathcal{E} \subseteq \delta(S)$ has size h. Thus each unordered balanced partition is counted once. Each cut has $4h^2$ crossing edges, so the number of weightings is exactly k_h from Eq. (834). For this pair set

$$\tau_e^{(S,\mathcal{E})} := \begin{cases} 4h, & \text{if both endpoints of } e \text{ lie in } S \text{ or both in } S^c, \\ 1, & \text{if } e \in \mathcal{E}, \\ 1/(4h^2), & \text{if } e \in \delta(S) \setminus \mathcal{E}. \end{cases} \tag{840}$$

All these weights are positive. The entire family is fixed before a spanning tree is chosen.

We verify the leverage bound for an arbitrary fixed pair $(S,\mathcal{E})$. For an edge inside either side, retain only the weight-$4h$ clique on its $2h$ vertices. Its effective resistance between distinct vertices is $2/(8h^2)$. Adding the remaining edges cannot increase effective resistance. Multiplying by the edge weight $4h$ shows that the leverage of every internal edge is at most $1/h$.

Define the total crossing weight by $B := \sum_{e\in\delta(S)} \tau_e^{(S,\mathcal{E})}$. We have

$$B = h + \frac{4h^2-h}{4h^2} = h+1-\frac{1}{4h}. \tag{841}$$

The first step counts the h edges of weight one and the remaining $4h^2-h$ crossing edges in Eq. (840). The second follows from simple algebra.

For a crossing pair $u \in S$, $v \in S^c$, orient every crossing edge from S to S^c and send $\tau_e^{(S,\mathcal{E})}/B$ units of flow along edge e. Let $a \in \mathbb{R}^S$ and $b \in \mathbb{R}^{S^c}$ be the distributions of the two endpoints of a crossing edge sampled with probability $\tau_e^{(S,\mathcal{E})}/B$. They are probability vectors. Route divergence $e_u - a$ inside the first clique and divergence $b - e_v$ inside the second clique, where e_u, e_v are the corresponding coordinate vectors. Together these flows form a unit flow from u to v.

The Laplacian of each internal clique equals $8h^2 I$ on the subspace perpendicular to the all-ones vector. Thus the minimum energy of an internal flow with zero-sum divergence d is $\|d\|_2^2/(8h^2)$. Since a probability vector has squared Euclidean norm at most one, $\|e_u - a\|_2^2 \leq 2$ and $\|b - e_v\|_2^2 \leq 2$. The crossing flow has energy $\sum_{e\in\delta(S)} (\tau_e^{(S,\mathcal{E})}/B)^2/\tau_e^{(S,\mathcal{E})} = 1/B$. Before bounding the total energy, we record the available margin. Eq. (841) and $h \geq 2$ give $B - h \geq 7/8$ and $B \leq h+1 \leq 3h/2$. We have

$$\frac{1}{h} - \frac{1}{B} = \frac{B-h}{hB} \geq \frac{7}{8hB} \geq \frac{7}{12h^2} > \frac{1}{2h^2}. \tag{842}$$

The first step follows from simple algebra. The second uses $B - h \geq 7/8$. The third uses $B \leq 3h/2$. The last uses $7/12 > 1/2$.

Let En be the energy of the constructed unit flow, with minimum-energy routing inside each clique. We have

$$\mathsf{En} = \frac{1}{B} + \frac{\|e_u - a\|_2^2}{8h^2} + \frac{\|b - e_v\|_2^2}{8h^2} \leq \frac{1}{B} + \frac{1}{2h^2} < \frac{1}{h}. \tag{843}$$

The first step adds the energies on the crossing edges and the two cliques. The second uses $\|e_u - a\|_2^2, \|b - e_v\|_2^2 \leq 2$. The last uses Eq. (842).

The minimum-energy characterization of effective resistance gives $R_{\mathrm{eff}}(u,v) < 1/h$. Every crossing edge has weight at most one, so its leverage is at most $1/h$. This proves Eq. (837) when $M = 1$.

Now fix any spanning tree F_0 of K_{4h}. For a uniformly random $2h$-element subset S of $[4h]$, a fixed edge crosses the partition with probability $2h/(4h-1)$. Since F_0 has $4h-1$ edges,

$$\mathbb{E}[|F_0 \cap \delta(S)|] = 2h. \tag{844}$$

Some balanced cut therefore contains at least $2h$ tree edges. Replace S by S^c if needed to ensure $1 \in S$, and choose any h-element set $\mathcal{E} \subseteq F_0 \cap \delta(S)$. The weighting $(S, \mathcal{E})$ belongs to the already fixed family. The h edges of $\mathcal{E}$ are retained by F_0 and each has weight one, so

$$\tau_{F_0}(\delta(S)) \geq h. \tag{845}$$

Eq. (841) gives the full cut weight:

$$\tau_{G_0}(\delta(S)) = B = h + 1 - \frac{1}{4h}. \tag{846}$$

For every weighted graph H, $\mathbf{1}_S^* Q_H \mathbf{1}_S = \tau_H(\delta(S))$. Eqs. (845) and (846) therefore give a cut ratio at least $h/B = a_{h,1}$, proving Eq. (838) for $M = 1$.

For general M, replace every edge of G_0 by M parallel copies, each with the same weight as the original edge in each weighting. The full graph Laplacian in every weighting is multiplied by M, and its pseudoinverse is divided by M. Consequently each copy's leverage is at most $1/(hM)$. A spanning tree of the resulting multigraph cannot contain two parallel copies of one edge, because they form a cycle. Its projection to G_0 is therefore a spanning tree F_0. Apply Eqs. (845) and (846) to this projection. The tree cut weight is unchanged, while the full graph cut weight is multiplied by M, proving Eq. (838). The cut ratio is a lower bound on every simultaneous spectral-thinness factor, since the spectral inequality must hold on $\mathbf{1}_S$. Finally, $h^2 - (h-1)(h+1-1/(4h)) = 1 + (h-1)/(4h) > 0$ proves Eq. (839). □

The number of weightings satisfies $\log k_h \sim h \log h$. Expressing the finite lower bound in terms of k_h gives the following asymptotic loss relative to the leverage bound.

Corollary 16.2 (Formal version of Theorem 1.9)**.** *There are sequences $k \to \infty$ and $\alpha \to 0$ of instances satisfying the hypotheses of Theorem 1.8 for which every simultaneously θ_{thin}-spectrally thin spanning tree obeys*

$$\theta_{\mathrm{thin}} \geq (1 - o(1))\alpha \frac{\log k}{\log \log k}. \tag{847}$$

Here $o(1)$ tends to zero along the sequence. These instances can additionally satisfy $\alpha \log k \to 0$. Thus an $O(\alpha)$ guarantee with an absolute constant independent of k is impossible, even without a running-time restriction.

Proof. The elementary binomial estimates $(N/r)^r \leq \binom{N}{r} \leq (eN/r)^r$ give

$$(4h)^h \leq k_h \leq 2^{4h}(4eh)^h. \tag{848}$$

Taking logarithms in Eq. (848) gives

$$\log k_h = h \log h + O(h). \tag{849}$$

A second logarithm yields

$$\log\log k_h = \log h + \log\log h + o(1). \tag{850}$$

Dividing Eqs. (849) and (850) gives

$$\frac{\log k_h}{\log\log k_h} = (1 + o(1))h. \tag{851}$$

The explicit ratio in Eq. (839) also gives

$$\frac{a_{h,M}}{\alpha_{h,M}} = (1 + o(1))h. \tag{852}$$

The ratio of the quantities in Eqs. (851) and (852) tends to one, uniformly in M. Theorem 16.1 therefore proves Eq. (847). Take $M = h^2$, so $\alpha_{h,h^2} = 1/h^3 \to 0$. We have

$$\alpha_{h,h^2} \log k_h = \frac{h \log h + O(h)}{h^3} = O(\log h/h^2) \longrightarrow 0. \tag{853}$$

The first step uses Eq. (849) and $\alpha_{h,h^2} = 1/h^3$. The second simplifies the quotient; $\log h/h^2 \to 0$.

Theorem 16.1 and Eq. (852) show that $\theta_{\text{thin}}/\alpha_{h,h^2}$ is unbounded along this sequence. Thus no absolute constant can give an $O(\alpha)$ bound independently of k. □

Remark 16.3 (Comparison with a single weighting)**.** The $O(\alpha)$ bound of Anari, Oveis Gharan, Saberi, and Srivastava [AOSS18] in Table 2 concerns a single weighting. Applying that result separately gives, for every ℓ, a tree F_ℓ with $Q^{(\ell)}_{F_\ell} \preceq C\alpha Q^{(\ell)}$, where C is an absolute constant. For a common tree F, Theorem 16.1 gives a weighting ℓ whose cut ratio is at least $a_{h,M}$. The single-weighting guarantee allows the tree to depend on ℓ; simultaneous thinness requires one tree for all ℓ. A factor of $\log\log k$ remains between Corollary 16.2 and the upper bound in Theorem 1.8.

References

[ADW+25] Josh Alman, Ran Duan, Virginia Vassilevska Williams, Yinzhan Xu, Zixuan Xu, and Renfei Zhou. More asymmetry yields faster matrix multiplication. In *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2005–2039. SIAM, 2025.

[AKK+20] Thomas D. Ahle, Michael Kapralov, Jakob B. T. Knudsen, Rasmus Pagh, Ameya Velingker, David P. Woodruff, and Amir Zandieh. Oblivious sketching of high-degree polynomial kernels. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 141–160. SIAM, 2020.

[AO14] Nima Anari and Shayan Oveis Gharan. The Kadison–Singer problem for strongly Rayleigh measures and applications to asymmetric TSP. *arXiv preprint arXiv:1412.1143*, 2014.

[AO15] Nima Anari and Shayan Oveis Gharan. Effective-resistance-reducing flows, spectrally thin trees, and asymmetric TSP. In *Proceedings of the 56th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 20–39. IEEE, 2015.

[AOSS18] Nima Anari, Shayan Oveis Gharan, Amin Saberi, and Nikhil Srivastava. Approximating the largest root and applications to interlacing families. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1015–1028. SIAM, 2018.

[Ban98] Wojciech Banaszczyk. Balancing vectors and Gaussian measures of n-dimensional convex bodies. *Random Structures & Algorithms*, 12(4):351–360, 1998.

[BBL09] Julius Borcea, Petter Brändén, and Thomas M. Liggett. Negative dependence and the geometry of polynomials. *Journal of the American Mathematical Society*, 22(2):521–567, 2009.

[BCMS19] Marcin Bownik, Pete G. Casazza, Adam W. Marcus, and Darrin Speegle. Improved bounds in Weaver and Feichtinger conjectures. *Journal für die reine und angewandte Mathematik*, 749:267–293, 2019.

[BDGL18] Nikhil Bansal, Daniel Dadush, Shashwat Garg, and Shachar Lovett. The Gram–Schmidt walk: A cure for the Banaszczyk blues. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 587–597. ACM, 2018.

[BF81] József Beck and Tibor Fiala. Integer-making theorems. *Discrete Applied Mathematics*, 3(1):1–8, 1981.

[BJ26] Nikhil Bansal and Haotian Jiang. Decoupling via affine spectral-independence: Beck–Fiala and Komlós bounds beyond Banaszczyk. In *Proceedings of the 58th Annual ACM Symposium on Theory of Computing (STOC)*, pages 432–442. ACM, 2026.

[Bow26] Marcin Bownik. On Akemann–Weaver conjecture. *Advances in Mathematics*, 487:110772, 2026.

[Brä18] Petter Brändén. Hyperbolic polynomials and the Kadison–Singer problem. arXiv preprint arXiv:1809.03255, 2018.

[CLS21] Michael B Cohen, Yin Tat Lee, and Zhao Song. Solving linear programs in the current matrix multiplication time. *Journal of the ACM (JACM)*, 68(1):1–39, 2021.

[Coh16] Michael B. Cohen. Improved spectral sparsification and Kadison–Singer for sums of higher-rank matrices. BIRS Workshop Lecture Videos, Banff International Research Station, 2016.

[CW90] Don Coppersmith and Shmuel Winograd. Matrix multiplication via arithmetic progressions. *Journal of Symbolic Computation*, 9(3):251–280, 1990.

[DDHK07] James Demmel, Ioana Dumitriu, Olga Holtz, and Robert Kleinberg. Fast matrix multiplication is stable. *Numerische Mathematik*, 106(2):199–224, 2007.

[DEK+26] Emilien Dupont, Marvin Eisenberger, Borislav Kozlovskii, Abbas Mehrabian, Francisco J. R. Ruiz, Abigail See, Renfei Zhou, Josh Alman, Virginia Vassilevska Williams, and Matej Balog. Improving the matrix multiplication exponent with modern optimization and AlphaEvolve. arXiv preprint arXiv:2608.16884, 2026.

[EJ26] Ekene Ezeunala and Haotian Jiang. Rank-one matrix discrepancy and algorithmic Kadison–Singer. *arXiv preprint arXiv:2609.17266*, 2026.

[FTU23] Zachary Frangella, Joel A. Tropp, and Madeleine Udell. Randomized Nyström preconditioning. *SIAM Journal on Matrix Analysis and Applications*, 44(2):718–752, 2023.

[Goe09] Michel X. Goemans. Lecture notes on the ellipsoid algorithm. Massachusetts Institute of Technology, 18.433: Combinatorial Optimization, May 2009. Handout 19.

[Har15] Nicholas J. A. Harvey. A note on the discrepancy of matrices with bounded row and column sums. *Discrete Mathematics*, 338(4):517–521, 2015.

[HO14] Nicholas J. A. Harvey and Neil Olver. Pipage rounding, pessimistic estimators and matrix concentration. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 926–945. SIAM, 2014.

[JPS11] Claude-Pierre Jeannerod, Clément Pernet, and Arne Storjohann. Rank-profile revealing Gaussian elimination and the CUP matrix decomposition, 2011.

[Kat26] Tarun Kathuria. A walk from free probability to matrix discrepancy II: Weaver's problem and the Kadison–Singer conjecture. *arXiv preprint arXiv:2609.18913*, 2026.

[KKS22] Tali Kaufman, Rasmus Kyng, and Federico Soldá. Scalar and matrix Chernoff bounds from ℓ_∞-independence. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 3732–3753. SIAM, 2022.

[KLS20] Rasmus Kyng, Kyle Luh, and Zhao Song. Four deviations suffice for rank 1 matrices. *Advances in Mathematics*, 375:107366, 2020.

[KS59] Richard V. Kadison and I. M. Singer. Extensions of pure states. *American Journal of Mathematics*, 81(2):383–400, 1959.

[KS18] Rasmus Kyng and Zhao Song. A matrix Chernoff bound for strongly Rayleigh distributions and spectral sparsifiers from a few random spanning trees. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, 2018.

[LS19] Yin Tat Lee and Aaron Sidford. Solving linear programs with $\sqrt{\text{rank}}$ linear system solves. arXiv preprint arXiv:1910.08033, 2019.

[MJF19] Alaa Maalouf, Ibrahim Jubran, and Dan Feldman. Fast and accurate least-mean-squares solvers. In *Advances in Neural Information Processing Systems*, volume 32, pages 8305–8316. Curran Associates, Inc., 2019.

[MSS15] Adam W. Marcus, Daniel A. Spielman, and Nikhil Srivastava. Interlacing families II: Mixed characteristic polynomials and the Kadison–Singer problem. *Annals of Mathematics*, 182(1):327–350, 2015.

[Son26] Zhao Song. Faster linear programming with $\sqrt{\text{rank}}$ linear system solves. manuscript, 2026.

[SWYZ21] Zhao Song, David P. Woodruff, Zheng Yu, and Lichen Zhang. Fast sketching of polynomial kernels of polynomial degree. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 9812–9823. PMLR, 2021.

[SY26a] Zhao Song and Song Yue. Polynomial time algorithms for the Kadison–Singer problem. *arXiv preprint arXiv:2609.19794*, 2026.

[SY26b] Zhao Song and Song Yue. Two standard deviations are necessary for the Kadison–Singer problem. Cryptology ePrint Archive, Paper 2026/2123, 2026.

[SZ26] Zhao Song and Lichen Zhang. Matrix spencer: Eight standard deviations suffice and an almost-linear time algorithm for dense input. *arXiv preprint arXiv:2609.15025*, 2026.

[Tro12] Joel A. Tropp. User-friendly tail bounds for sums of random matrices. *Foundations of Computational Mathematics*, 12:389–434, 2012.

[Wea04] Nik Weaver. The Kadison–Singer problem in discrepancy theory. *Discrete Mathematics*, 278(1–3):227–239, 2004.

[XXZ22] Jiaxin Xie, Zhiqiang Xu, and Ziheng Zhu. Upper and lower bounds for matrix discrepancy. *Journal of Fourier Analysis and Applications*, 28(6):81, 2022.

AI Usage Disclosure

We used Claude Code Fable 5, ChatGPT Pro 5.6/6, and Codex 5.6/6 Sol in preparing this draft. To obtain the square-root-log discrepancy, the authors came up with the idea of modifying the rank-one potential function by using a Schatten norm. This idea is more or less analogous to the relationship between the log-barrier function and the Lee–Sidford barrier function in convex optimization, for example in solving linear programs [CLS21, LS19, Son26]. The author sent the new potential function to AI models and asked them to figure out the details of the proofs. All of the proofs have been substantially rewritten by the authors. All of the proofs have been verified by different AI models, as well as by the human authors. The authors take full responsibility for the correctness of the proofs.

A Bernstein certificate for Schatten order three

This appendix gives the partition certificate used in Section 8.8. The polynomials P_3 and G_3 are defined in Eqs. (383) and (382), with the budget in Eq. (384), $\iota = 123/20$, and $\mu(t) = 499(1-t)^2/1000$.

For $g(t) = \sum_{j=0}^{d} g_j t^j$, its degree-$d$ Bernstein coefficient numbered i is $\sum_{j=0}^{i} g_j \binom{i}{j} / \binom{d}{j}$. On another interval first substitute its affine parametrization; apply this formula in each variable for a tensor polynomial. The Bernstein basis functions are nonnegative and sum to one. The polynomials $P_3(t, u/2)$ and $G_3(t, u/2)$ have bidegrees $(64, 6)$ and $(28, 4)$. The partition is generated by the following deterministic rule. For a rectangle $Q = [a, b] \times [c, d]$ in (t, u) coordinates, substitute $t = a + (b-a)v$ and $u = c + (d-c)w$. Let $b_P(Q)$ and $b_G(Q)$ be the smallest tensor Bernstein coefficients of the resulting polynomials, using the respective bidegrees $(64, 6)$ and $(28, 4)$. All coefficients and comparisons are computed exactly over the rationals.

Start with $\textsc{Partition}([0,1]^2, 0)$. Exact evaluation visits 311 rectangles and terminates with 86 leaves of type P and 70 leaves of type G; the maximum depth is 18. Direct affine substitution and the Bernstein formula above independently verify the stated strict coefficient bounds on every terminal rectangle. Each bisection preserves coverage and disjoint interiors, so the leaves partition the unit square. A G leaf contains no admissible point because admissibility requires $G_3 \le 0$.

Algorithm 4 Construction of the order-three Bernstein partition

procedure PARTITION(Q, k)
  if $b_P(Q) > 1/50000$ **then**
    return the leaf (Q, P)
  end if
  if $b_G(Q) > 1/10000$ **then**
    return the leaf (Q, G)
  end if
  Require $k < 18$
  Bisect Q at the midpoint in t if k is even, and in u if k is odd
  Let Q_- and Q_+ be the lower and upper children
  $\mathcal{L}_- \leftarrow$ PARTITION($Q_-, k+1$)
  $\mathcal{L}_+ \leftarrow$ PARTITION($Q_+, k+1$)
  return $\mathcal{L}_-$ followed by $\mathcal{L}_+$
end procedure

Every remaining leaf has $P_3 > 1/50000$. This proves the bound throughout the admissible domain, including its boundary.

B Finite scalar certificates

This appendix specifies the finite data and procedures used in the scalar certificates. These finite verifications were performed computationally, using exact rational arithmetic for polynomial and integer checks and rigorous outward enclosures for interval checks. The displayed formulas, parameters, partitions, and acceptance rules specify the inputs to each verification.

B.1 Exact coefficient verification

All terminating decimals in a certificate denote exact rational numbers. Construct its polynomials from the displayed formulas by rational addition, coefficient convolution for multiplication, and formal differentiation. Cancel any displayed polynomial quotient exactly before evaluation at an endpoint. A polynomial identity is checked by bringing its two sides to a common denominator and checking that every numerator coefficient is zero.

Here is a direct rule for checking any of the printed Bernstein certificates. Write $P(x) = \sum_{\mathbf{a}} M_{\mathbf{a}} x^{\mathbf{a}}$, with coordinate degrees $n_1, \dots, n_d$. On a box $Q = \prod_{j=1}^{d} [a_j, b_j]$, substitute $x_j = a_j + (b_j - a_j)u_j$. Its monomial coefficients and Bernstein coefficients are, respectively,

$$C_\gamma(Q) = \sum_{\mathbf{a} \geq \gamma} M_{\mathbf{a}} \prod_{j=1}^{d} \binom{\mathbf{a}_j}{\gamma_j} a_j^{\mathbf{a}_j - \gamma_j} (b_j - a_j)^{\gamma_j}, \tag{854}$$

$$B_\eta(Q) = \sum_{\gamma \leq \eta} C_\gamma(Q) \prod_{j=1}^{d} \frac{\binom{\eta_j}{\gamma_j}}{\binom{n_j}{\gamma_j}}, \qquad 0 \leq \eta_j \leq n_j. \tag{855}$$

Zero powers, including 0^0, equal one in these formulas. Use the degrees specified by the certificate, or the coordinate degrees of the displayed polynomial when no degree is specified. All sums are finite and all comparisons are rational.

Read each printed tree in preorder from its stated root box. A split letter bisects the indicated coordinate and is followed by the lower and upper subtrees, in that order. A leaf applies its stated polynomial and coefficient threshold to the entire current box. Accept a tree only after every leaf passes and every letter is consumed exactly. For generated partitions, use the split order, depth cap, thresholds, and stopping rules specified below. Midpoint subdivision preserves the whole parent box, so this also verifies coverage. When zero coefficients are allowed, the separate positive row margin stated in the corresponding proof supplies strictness at actual light states. For a parameter family expanded in powers of $s-s_0 \geq 0$, apply the same test to each displayed coefficient polynomial.

B.2 Partitions for the rank-dependent bounds

The four partitions for the coefficients 4.307 and 4.225 are generated by the same alternating subdivision rule as Algorithm 4. Use the polynomials P and G, with their exact parameters, from Sections 11.3 and 11.4. The rectangle coordinates are $(t,u) \in [0,1]^2$, with $x = u/2$.

For a rectangle $Q = [a,b] \times [c,d]$, substitute $t = a + (b-a)v$ and $x = (c + (d-c)w)/2$ into each polynomial. If the resulting polynomial is $\sum_{i,j} q_{ij} v^i w^j$ and the original polynomial has bidegree (m,n), compute its tensor Bernstein coefficients exactly as

$$b_{kl} = \sum_{i=0}^{k} \sum_{j=0}^{l} q_{ij} \frac{\binom{k}{i}\binom{l}{j}}{\binom{m}{i}\binom{n}{j}} \qquad (0 \leq k \leq m,\ 0 \leq l \leq n).$$

Let $b_P(Q)$ and $b_G(Q)$ be their respective minima. Start from $Q = [0,1]^2$ at depth zero. Accept a P leaf if $b_P(Q) > 10^{-30}$. Otherwise accept a G leaf if $b_G(Q) > 10^{-30}$. If neither test passes, bisect t at even depths and u at odd depths, and process the lower child before the upper child. Require success by depth 17. Equivalently, use Algorithm 4 with both acceptance thresholds replaced by 10^{-30} and the depth limit replaced by 17.

Exact execution gives the following partitions.

Coefficient	Ranks	P leaves	G leaves	Total	Maximum depth
4.307	3	65	52	117	17
4.307	4	12	7	19	10
4.225	3	27	18	45	16
4.225	4–18	57	40	97	17

Direct affine substitution independently verifies the strict coefficient bound at every generated leaf. Each bisection preserves coverage and disjoint interiors, so the terminal rectangles partition the unit square.

For each branch also compute the univariate Bernstein coefficients of

$$f, \quad -N/t, \quad (f^2 - z)/t^2, \quad f + zf', \quad A, \quad \eta, \quad H, \quad G(t, 1/2), \qquad z = 1 - t^2.$$

The divisions by t and t^2 here are polynomial divisions: the corresponding low-degree coefficients vanish exactly. Here $A = f - zf'$ is denoted by A_0 in Section 11.4. All these Bernstein coefficients are strictly positive on $[0,1]$. For the coefficient 4.225, also apply the tensor formula on the whole square to V_0, K_D, H_J, defined in Section 11.4; all their coefficients are positive. These checks supply the budget, denominator, and row-sign conditions used in the reductions. On a G leaf the admissibility inequality $G \leq 0$ is impossible; on every other leaf $P > 0$. Thus the finite checks imply the scalar inequalities throughout the admissible domain.

B.3 Interval subdivision for the trace-variance bounds

Use the rescaled formulas of Section 3.7, with the parameters of Lemma 3.4 or Lemma 3.6. The following arithmetic and subdivision rules generate the terminal boxes from the unit cube.

Put $M = 2^{40}$. Before an interval operation, replace $[a, b]$ by $[\lfloor Ma \rfloor / M, \lceil Mb \rceil / M]$. An unbounded operand, an endpoint of magnitude at least 2^{22}, or division by an interval containing zero gives the unbounded interval. For bounded operands, addition and subtraction use their usual endpoint formulas. Multiplication and division take the minimum and maximum of the four endpoint products or quotients and round outward to $M^{-1}\mathbb{Z}$. Squaring uses lower endpoint zero when the operand contains zero. For a square root, use zero as lower endpoint if the input lower endpoint is negative, and otherwise use the floor square root of its scaled integer; the upper endpoint uses the ceiling square root. Explicitly, for input $[A/M, B/M]$ with $B \geq 0$, these endpoints are

$$\frac{\lfloor \sqrt{M \max\{0, A\}} \rfloor}{M}, \qquad \frac{\lceil \sqrt{MB} \rceil}{M}.$$

A negative upper endpoint gives the unbounded interval. Integer square comparisons determine both square roots exactly.

To specify the stored enclosures, convert each resulting finite dyadic endpoint to binary64 by rounding to nearest, ties to even, and then move it one representable value outward. An exact zero lower endpoint of a square or square root remains zero. Negation and intersection are exact endpoint operations. Rational constants are initialized by outward division on the same grid. Integers and box endpoints are initialized exactly. No fused operations are used.

Evaluate polynomials by Horner's rule, starting with the leading coefficient. Differentiate their coefficient lists first. Evaluate integer powers by repeated squaring, multiplying into an accumulator initialized at one for each set binary digit, from the least significant digit upwards. Other displayed products and sums are evaluated from left to right, respecting parentheses. Propagate interval derivatives with the product rule and the quotient rule

$$\partial_a(v/w) = (\partial_a v - (v/w)\partial_a w)/w.$$

Use $\partial_a \sqrt{v} = (\partial_a v)/(2\sqrt{v})$. Initialize the three coordinate derivatives as the corresponding unit vectors. The boundary enclosures for $S_0 = 0$ and the derivative conventions are those in the two certificate proofs.

For a box Q, let $L(Q)$ be the largest available lower bound for $G = 1 - \mathcal{E}_0 - \Phi(R_U^*, d_U) - \Phi(R_V^*, d_V)$ obtained by direct evaluation, by the five separate centered mean-value enclosures, and, for the improved parameters, by the centered enclosure of the complete G. Intersect each separate mean-value enclosure with its direct enclosure. Sum the three derivative contributions in coordinate order (t, h, k). Round a radius one binary64 value upwards before using it. If a required denominator fails its strict positivity test, set $L(Q) = -\infty$.

For the repair parameters choose a coordinate using the scores

$$|Q_a| \sum_{g \in \{1-\mathcal{E}_0, R_U^*, R_V^*, d_U, d_V\}} \max\{|\underline{\partial_a g}|, |\overline{\partial_a g}|\}.$$

For the improved parameters use $|Q_a| \max\{|\underline{\partial_a G}|, |\overline{\partial_a G}|\}$. When the center cannot be evaluated, all scores are zero. Choose a maximal score, breaking ties in coordinate order. If that score is zero or nonfinite, choose a longest coordinate, with the same tie rule.

The following procedure specifies traversal and acceptance. The volume sums can be taken as exact dyadic rationals. The accepted boxes give the counts in the two certificate proofs. For the improved parameters their minimum lower bound also exceeds 2^{-30}, as needed in the algorithmic extension. The budget tests use the univariate Bernstein formula above on $[0, 1/2]$ and $[1/2, 1]$.

Algorithm 5 Verification of a trace-variance scalar certificate

procedure VERIFYTRACE(parameters)
  for $o = 0, \dots, 7$ **do**
    $\mathcal{S} \leftarrow [(Q_o, 0)]$, where bit $a-1$ of o selects $[0, 1/2]$ or $[1/2, 1]$ in coordinate a
    $v \leftarrow 0$; $n \leftarrow 0$
    while $\mathcal{S}$ is nonempty **do**
      Pop (Q, d); $n \leftarrow n + 1$; compute $L(Q)$
      if $L(Q) \geq 0$ **then**
        $v \leftarrow v + \mathrm{vol}(Q)$
      else
        Fail if $d \geq 65$, or $n > 10^7$ for repair, or $n > 5 \cdot 10^7$ for improvement
        Bisect the selected coordinate; push the upper child, then the lower child, at depth $d + 1$
      end if
    end while
    Require $v = 1/8$
  end for
  return success
end procedure

B.4 Generation of the halving controls

For the constructive tree bound, put $Q = 2^{48}$, $T = 2^{28}$, $e = Q/1024$, and $c_0 = 307/20$. Starting with $l = Q$, set $h = \min\{\lceil 2049l/2048 \rceil, TQ\}$, process $[l, h]$, and replace l by h until $l = TQ$. Use the integer updates in Lemma 13.9. All square roots in this subsection mean integer floor or ceiling square roots.

The recursive procedure below tries a regularized suffix; the first successful branch is retained. Duplicate grid controls are discarded.

For each interval initialize $(A, B) = (l, h)$. Try FINISH$(A, B, 4)$, then perform one legal whitened step and try again. Continue until success or until no whitened step is legal. In the latter case bisect $[l, h]$ at $\lfloor (l + h)/2 \rfloor$ and process its lower half before its upper half, with a maximum of 15 bisections. Reinitialize the state on each child interval. Verify every chosen control and update using the integer inequalities in the lemma. Require $20B < 307Q$ at termination. This computation succeeds on all 39,758 initial intervals without bisection, uses at most 22 whitened steps and at most four regularized steps per interval, and gives the upper bound reported in the lemma.

B.5 Partition for the thin-tree budget

Use f, N, A, H_6, D_6, P_6 from Lemma 13.1. The partition is generated by the alternating subdivision rule in Algorithm 4, with the following parameters. Start with $Q = [0, 1] \times [0, 2/5]$ in coordinates (t, x) at depth zero. For each rectangle $Q = [a, b] \times [c, d]$, substitute $t = a + (b - a)v$ and $x = c + (d - c)w$. Compute the tensor Bernstein coefficients of P_6 and D_6 at their respective bidegrees $(68, 6)$ and $(32, 4)$, using the formula in Appendix B.2. Denote their minima by $b_P(Q)$ and $b_D(Q)$.

Accept a P leaf if $b_P(Q) > 0$. Otherwise accept a D leaf if $b_D(Q) > 0$. If neither test passes, bisect t at even depths and x at odd depths, and process the lower child before the upper child. Require success by depth 12. Thus the acceptance thresholds in Algorithm 4 are both zero, the second polynomial is D_6, and the depth limit is 12. All computations and comparisons use exact rational arithmetic.

Algorithm 6 A regularized suffix for the halving certificate

procedure FINISH(A, B, d)
  if $20B < 307Q$ **then**
    return the empty suffix
  end if
  if $A \le Q$ **then**
    return failure
  end if
  if $A - e > \lceil \sqrt{AQ} \rceil$ **then**
    Compute B_f from Eq. (728)
    if $20B_f < 307Q$ **then**
      return the final step
    end if
  end if
  if $d \le 1$ or $B \ge 2^d c_0 Q$ **then**
    return failure
  end if
  $L \leftarrow \lceil \sqrt{AQ} \rceil$; $H \leftarrow \min\{\lfloor \sqrt{BQ} \rfloor, A - 2Q - 2e\}$
  if $H < L$ **then**
    return failure
  end if
  for distinct $C = L + \lfloor (H - L)j/12 \rfloor$, in increasing order, $j = 0, \dots, 12$ **do**
    Compute (A', B') by Eq. (727)
    $S \leftarrow$ FINISH$(A', B', d - 1)$
    if S succeeds **then**
      return the control C followed by S
    end if
  end for
  return failure
end procedure

Exact execution visits 103 rectangles and terminates with 33 leaves of type P and 19 of type D; the maximum depth is 12. Direct affine substitution independently verifies that every Bernstein coefficient of the named polynomial is strictly positive on each terminal rectangle. The univariate formula on $[0, 1]$ also gives strictly positive coefficients for all six polynomials in Eq. (706). Each bisection preserves coverage and disjoint interiors. A D leaf has no point with $D_6 \le 0$, and every other leaf has $P_6 > 0$. These rational checks establish the scalar inequality on its entire admissible domain.

B.6 The biased thinning controls

Put $Q := 2^{48}$, $D := 2^{24}$, and $T_0 := 2^{28}$. The initial integer state is $A_0 = B_0 = T_0 Q$. Table 3 specifies all controls; in row j use $p = P/D$ and $c = C/Q$.

For each row compute H by Eq. (718), check Eq. (719), and perform Eq. (720). The eight

Step	P	C
1	1514298	1691346677476098048
2	1716154	524833238967530429
3	1786728	169468856139666568
4	1805162	55320068483034227
5	1794431	17989217307058159
6	1738288	5704923317240544
7	1545519	1645362096142320
8	722119	246819608938732

Table 3: Rational controls for the biased thinning recurrence.

values of H are

$$37865342171149,\quad 40415454071498,\quad 41258457034566,$$
$$41474775786323,\quad 41349041813857,\quad 40682413169660,$$
$$38274027155813,\quad 25102884608282.$$

For each value, exact squaring checks

$$\frac{(H-1)^2}{Q^2} < \frac{p(1-p)}{4} f (1-2p)^2 \le \frac{H^2}{Q^2}.$$

Every updated lower bound is positive. The final state is

$$A_8 = 2471088, \qquad B_8 = 1631855358146481.$$

Thus the terminal lower bound is positive and $B_8/Q < 5.797515$. Lemma 13.6 covers every $0 < u \le T_0$, and the tail comparison in Eq. (723) gives the bound $29/5$ for all positive starting scales. The certificate uses only the displayed rational controls and the integer updates above; its verification is independent of any numerical procedure used to choose those controls.

B.7 Scalar comparisons for the numerical constants

All remaining logarithms, roots, and exponentials in the numerical comparisons admit rational bounds. For $y \ge 1$ put $z = (y-1)/(y+1)$ and use

$$L_N(y) = 2 \sum_{j=0}^{N-1} \frac{z^{2j+1}}{2j+1}, \qquad L_N(y) \le \log y \le L_N(y) + \frac{2z^{2N+1}}{(2N+1)(1-z^2)}.$$

For the rank-signing comparisons take $N = 200$. For the tree comparisons write $y = 2^k v$ with $1 \le v < 2$, use $\log y = k \log 2 + \log v$, and take $N = 40$ for each logarithm. To bound $r^{1/q}$ from above, take the smallest integer J with $J^q \ge r 10^{12q}$ and use $J/10^{12}$, or verify the displayed decimal upper bound by the same integer-power comparison. For $0 \le x \le 1$, use

$$\sum_{j=0}^{24} \frac{x^j}{j!} \le e^x \le \sum_{j=0}^{24} \frac{x^j}{j!} + \frac{x^{25}}{25!} \frac{1}{1-x/26}.$$

Substitute these bounds, with the appropriate outward direction, into the endpoint expressions in the rank-signing and tree proofs. They give the strict comparisons stated there. All coefficients,

subdivision rules, integer updates, and scalar bounds required for these finite computations are specified in the paper.

For the comparisons in Corollary 11.22, use binary range reduction and $N = 40$ for every logarithm. A lower root enclosure is $(J-1)/10^{12}$ for the same integer J defined above. Put $\lambda = \log 2$ and $x = 559\lambda/400$, and apply the degree-24 exponential bounds to $C_{\mathrm{new}}^2 = 16e^x/(5\lambda)$. Substitution of the lower and upper rational endpoints gives

$$12.1622502785 < C_{\mathrm{new}}^2 < 12.1622502786 < (3.48745)^2.$$

The same rules give $2.176^2/\log 2 < 6.832$ and the four rational upper bounds in the corollary's table. In its last row the two expressions are $206/(25\log 2)$ and $309/(5\log 162)$. Thus each finite-rank bound is strictly below the displayed lower bound on C_{new}^2.