Equicorrelation is the exact maximizer of the permanental cofactor ratio in large dimensions

Scott Schubr
Luminary Oddities

November 2026

Abstract

For a complex correlation matrix B let $C(B)$ be its permanental cofactor matrix, let $\Lambda(B) = \lambda_{\max}(C(B))/\operatorname{per} B$, and let γ_n be the maximum of Λ in order n. An isolated-row induction and a positive equicorrelation family give $(n+1)^2/(4n) < \gamma_n \le (n^2+2n+3)/(4n)$ for $n \ge 2$. We prove that for all sufficiently large n the optimal positive equicorrelation matrix is the exact maximizer of Λ over all complex correlation matrices, unique up to diagonal unitary conjugation, so that $\gamma_n = (n+1)^2/(4n) + O(ne^{-cn})$; maximality and uniqueness also hold in orders three and four. The proof iterates a drift law for cofactor diagonals and closes with local rigidity. In order five it is the maximizer, unique up to diagonal unitary conjugation, among correlation matrices whose monomial unitary symmetries act transitively, with maximum $1.8727665\ldots$, an algebraic number of degree eight; over all correlation matrices its orbit is a strict local maximum at Frobenius radius $1/100$, and $\gamma_5 < 1.876674$. Cyclic dephasing drives the permanent of an order-five circulant correlation matrix to one at the sharp exponential rate $1.2000466507488\ldots$. For all large n the article proves exact maximality under spectral floors and among one-factor matrices, the Bapat–Sunder inequality near weak-entry correlation matrices, and, near the optimizer, permanent-on-top and hence Lieb's permanental dominance inequality for every subgroup and character. It also gives cutoff profiles for the equicorrelation subset shuffle, including observation through colour tables.

Keywords. Permanent; permanental cofactor matrix; correlation matrix; equicorrelation; local rigidity; circulant matrix; cutoff.
MSC 2020. Primary 15A15; Secondary 15A42, 15B05, 15B48, 60J10.

Overview

Sections 36–41 prove the main result: the sharp third term (Theorem 11) and exact unrestricted maximality with uniqueness in all large orders (Theorem 12). They use the exact equicorrelation optimum and local Hessian of Sections 1–5 (Proposition 2 and Theorem 1) and the uniform local rigidity of Section 20 (Theorem 7). Section 20 also proves exact maximization under the spectral floor, Sections 25–30 extend it to block noise (Theorem 8), and Sections 31–35 give the vanishing-floor result (Theorem 9). Section 6 treats small orders; Section 7 proves global maximality among one-factor matrices (Theorem 3); Sections 11–15 and 16–19 give the fixed-rank and arbitrary-rank localization (Theorems 4 and 6). Sections 42–46 treat Walsh families (Theorems 16–19), and Sections 47–50 give the matching lower comparisons (Theorem 20), whose exponential scale in every real rank is determined in Sections 51–53 (Theorem 22), with improved rank-three witnesses in Sections 54–55. Sections 56–61 settle structured cases at orders three and four, and Section 62 determines the unrestricted order-four maximum. Section 63 turns to Schur products:

in all sufficiently large orders it proves the Bapat–Sunder inequality on a dimension-free Frobenius neighbourhood of every weak-entry correlation matrix, which contains all sharp-scale cofactor near-maximizers, together with a permanent-on-top spectral gap, which gives Lieb's permanental dominance inequality for every subgroup and character on the same band (Section 63.3), and a strong self-conjugate inequality.

Section 64 gives the sharp fixed-order comparison between a correlation minor and its normalized cofactor matrix in near-null directions, and Section 65 turns the Schur power of positive equicorrelation into a permutation shuffle and determines its cutoff and Rényi profiles at every order for the full permutation; under partial observation it determines the cutoff at every order above one, with exact profiles on the saturated branch $\zeta > 1 - \rho$, at orders $1 < \alpha \leq 2$, and in the Poisson regime; Sections 65.16–65.17 observe shuffles through colour tables. Section 66 treats order five: the exact maximum among transitive correlation matrices (Theorem 52), local rigidity at an explicit radius (Theorem 53), a window of width 0.004 for γ_5 (Theorem 54) and the sharp cyclic dephasing rate (Theorem 67). Section 67 proves the bordering identities, the recursive bound and the order-three theorem used throughout.

This is the seventeenth edition of this article; it adds Section 66 on order five and cyclic dephasing, Sections 65.16–65.17 on shuffles observed through colour tables, and Section 67, which proves inside the article the bordering identities and the order-three theorem used throughout; it also records Lieb's permanental dominance inequality on the band of Section 63.3.

1 Introduction and main result

For a matrix $B = (B_{ij})$ of order n, write $\operatorname{per} B = \sum_\sigma \prod_i B_{i,\sigma(i)}$ and $C(B)_{ij} = \operatorname{per} B(i,j)$, where $B(i,j)$ deletes row i and column j. A correlation matrix is Hermitian positive semidefinite with diagonal entries one. On this space $C(B)$ is Hermitian and $\operatorname{per} B \geq 1$ by Lieb's permanent inequality [1]. Define

$$\Lambda(B) = \frac{\lambda_{\max}(C(B))}{\operatorname{per} B}.$$

The raw cofactor convention matters for phase conjugation and differentiation. It also differs from the entry-weighted cofactor matrix that arises in a conjecture of Bapat and Sunder [2].

Lemma 71 puts the maximum of Λ between $(n+1)^2/(4n)$ and $(n^2+2n+3)/(4n)$ for $n \geq 2$. The lower construction is a positive equicorrelation matrix with off-diagonal entry $2/(n+1)$. Its exact scalar quotient and the exact order-three classification of Theorem 73 motivate the study of perturbations in every complex correlation direction. Lavi [3, Section 4] bounds the same normalized raw cofactor spectrum at order four. Marcus and Merris [4] prove $C(B)^{\mathsf{T}} \preceq n(\operatorname{per} B)B^{-1}$ for positive definite B. The scalar permanent derivative, expressed through its cofactors, also enters the fixed-spectrum extremal analysis of Grone, Johnson, Sá and Wolkowicz [5]. Bhatia, Grover and Jain [6, Section 4] give higher derivatives of the scalar permanent. Here the simple-eigenvalue derivative and the combinatorics of two directed edges give the second variation of the normalized raw spectrum.

Let I be the identity, $J = \mathbf{1}\mathbf{1}^*$, and $B(b) = (1-b)I + bJ$ for $-1/(n-1) \leq b \leq 1$. Choose b_n to maximize $\Lambda(B(b))$ on this interval, and put $B_n = B(b_n)$. The space of affine perturbations consists of Hermitian matrices X with diagonal zero. We use the real Frobenius inner product and write $\|X\|_F^2 = \sum_{i,j} |X_{ij}|^2$.

Theorem 1. *For all sufficiently large n, b_n is interior and positive. The matrix B_n is a strict local maximizer of Λ modulo diagonal-unitary conjugation. The Hessian kernel is precisely the space of imaginary gradients $X_{ij} = i(\theta_i - \theta_j)$, where θ_i are real. On its Frobenius orthogonal complement*

the Hessian has the following four orthogonal sectors. On each sector its quadratic form equals the indicated coefficient times $\|X\|_F^2$.

constant real off-diagonals	$-n/8 + o(n)$
real stars $X_{ij} = (v_i + v_j)/2$ $(i \neq j)$, $\sum_i v_i = 0$	$-n/16 + o(n)$
real symmetric, zero row sums	$-1/8 + o(1)$
purely imaginary Hermitian, zero row sums	$-1/8 + o(1)$.

Proposition 2. *There is an absolute $c > 0$ such that every maximizing b_n satisfies*

$$t_n = \frac{1 - b_n}{b_n} = \frac{n-1}{2} + O(ne^{-cn/2}), \qquad \Lambda(B_n) = \frac{(n+1)^2}{4n} + O(ne^{-cn})$$

for all sufficiently large n.

The proposition optimizes the previously known scalar formula over the full real family, and the theorem treats perturbations in the full complex correlation space near the optimizing orbit. The exact optimum of the scalar formula (Section 2) improves the known lower bound at every finite order; Table 1 lists representative values.

n	t_n	$b_n = 1/(1+t_n)$	$\Lambda(B_n)$	$(n+1)^2/(4n)$
3	1.79549	0.35772	1.42285	1.33333
4	2.19358	0.31313	1.64523	1.56250
5	2.60559	0.27735	1.87277	1.80000
6	3.02927	0.24818	2.10437	2.04167
8	3.90485	0.20388	2.57665	2.53125
12	5.73673	0.14844	3.54389	3.52083
20	9.57853	0.09453	5.51829	5.51250
50	24.50076	0.03921	13.00503	13.00500

Table 1: The optimal real equicorrelation value $\Lambda(B_n)$, computed from the exact formula (1) at 40 digits, compared with the earlier lower bound. The values agree with direct Ryser evaluation of the cofactor spectrum for $n \leq 8$.

For $4 \leq n \leq 8$ we also evaluated the second difference of Λ at B_n along 60 random complex directions per order, after removing the phase-orbit component. Every value was strictly negative (the largest was -0.27, at $n = 8$), in agreement with Theorem 1 already at these small orders.

2 Optimizing the real equicorrelation family

Write $a = 1 - b$, and for $b > 0$ set $t = a/b$. Expansion in the diagonal part gives $\operatorname{per}(aI + bJ) = n! b^n \sum_{j=0}^{n} t^j/j!$. The raw cofactor matrix has positive off-diagonal entries, so its largest eigenvalue is its common row sum. This row sum is $n^{-1}\partial_b \operatorname{per}(aI + bJ)$ with a fixed, since the derivative adds every raw cofactor. Accordingly, for N_t Poisson with mean t, the quotient is

$$\Lambda(B(b)) = (1+t)\left(1 - \frac{t}{n} + \frac{t}{n}\frac{\mathbb{P}(N_t = n)}{\mathbb{P}(N_t \leq n)}\right) = \frac{1+t}{n}\mathbb{E}[n - N_t \mid N_t \leq n]. \tag{1}$$

It gives the lower bound of Lemma 71.

For $0 \le t \le 3n/4$, Chernoff's inequality gives $\mathbb{P}(N_t \ge n) \le e^{-cn}$ with $c = \log(4/3) - 1/4 > 0$. Thus the Poisson correction in (1) is $O(ne^{-cn})$ uniformly on this range. The quadratic part is maximized at $t_* = (n-1)/2$, with maximum $(n+1)^2/(4n)$ and exact defect $(t-t_*)^2/n$.

For $3n/4 \le t \le n$, the quadratic is at most $3n/16 + O(1)$, and $\mathbb{P}(N_t = n) \le \mathbb{P}(N_n = n) = O(n^{-1/2})$. Also $\mathbb{P}(N_t \le n) \ge \mathbb{P}(N_n \le n) \ge c_0 > 0$: Stirling's formula gives an atom of order $n^{-1/2}$ at n for N_n, and the $\lfloor \sqrt{n}/2 \rfloor$ atoms immediately below it are each comparable to it by the exact ratios of Poisson masses. The correction is consequently $O(\sqrt{n})$. For $n \le t \le 3n/2$ the conditional law $N_t \mid N_t \le n$ is stochastically increasing by monotone likelihood ratio. At $t = n$ its expected deficiency is $O(\sqrt{n})$ by Cauchy–Schwarz and the same conditioning bound. For $t \ge 3n/2$ successive deficiency masses have ratio at most $q = n/t$. If $p_j = \mathbb{P}(n - N_t = j \mid N_t \le n)$, then $\sum_j p_{j+k} \le q^k$, so the conditional mean deficiency is at most $n/(t-n)$. The quotient on this last range is $O(1)$. These bounds place the positive-branch maximum inside $0 \le t < 3n/4$ for large n. Comparison to t_* there proves the two exponential estimates.

For the negative branch, set $b = -r$ and $a = 1 + r$ with $0 \le r \le 1/(n-1)$. The identity $\operatorname{per}(aI + bJ) = \int_0^\infty (a + bu)^n e^{-u}\, du$ follows by expanding the polynomial and hence is valid for negative b as well. The cofactor diagonal d and row sum w have the integral expressions

$$d = \int_0^\infty (a - ru)^{n-1} e^{-u}\, du, \qquad w = \int_0^\infty u(a - ru)^{n-1} e^{-u}\, du.$$

The other eigenvalue is $(nd - w)/(n-1)$. On $0 \le u \le a/r$, one uses $a^{n-1} \le e$. Translating the two tails for $r > 0$ gives absolute bounds $e^{-a/r} r^{n-1}(n-1)!$ and that quantity times $a/r + n$, respectively. Since $r^{n-1}(n-1)! \le 1$ and $a/r \ge n$, both eigenvalues are uniformly bounded. The case $r = 0$ is the identity, and $\operatorname{per} B(b) \ge 1$ on this branch. Thus its quotient is $O(1)$, proving Proposition 2.

3 Spectral variation and directed edges

For the remaining proof write $B = aI + bJ$, $P = \operatorname{per} B$, $Q = C(B)/P$, and $e = n^{-1/2}\mathbf{1}$. We evaluate at $b = b_n, a = 1 - b_n$ after taking partial derivatives at fixed a where specified. The top eigenvalue of Q is simple and its eigenvector is e. Denote the gap to the other, common eigenvalue by $\Delta > 0$. Simplicity gives smoothness near B_n. Permutation invariance and invariance under complex conjugation imply that the first variation in a zero-diagonal Hermitian direction is a scalar multiple of the sum of the real off-diagonal entries. Since b_n is an interior branch maximizer, this scalar is zero.

For a Hermitian zero-diagonal X with total sum zero, write $P'' = D_X^2 \operatorname{per}(B)$. The first variation $D_X \operatorname{per}(B)$ is zero by the same invariance of the permanent. The identity $e^* C(B) e = n^{-1} \partial_b P$ holds also along $B + sX$ if a, X are held fixed. The second variation of the simple eigenvalue is therefore

$$D_X^2 \Lambda = \frac{1}{n} \partial_b (P''/P) + \frac{2}{\Delta} \| (I - ee^*) D_X Q\, e \|^2. \tag{2}$$

Write $r_i = \sum_j X_{ij}$ for its row sums. Equivariance makes $D_X Q e$ zero when $r = 0$. Indeed, for a real symmetric input the coefficient of each unordered off-diagonal edge at a given output coordinate depends only on incidence with that coordinate. When the total sum is zero, the image is a scalar multiple of r. For a purely imaginary Hermitian input, exchanging the ends of an edge disjoint from the output coordinate fixes the coordinate and negates the edge. Such edges contribute zero. Every incident edge contributes a constant times the entry oriented from the output coordinate. These statements also identify the mixing term in each pure sector of total sum zero as a scalar multiple of $\|r\|^2$.

Expand the permanent by the diagonal part:

$$\operatorname{per}(aI + bJ + sX) = \sum_{S \subseteq [n]} a^{n-|S|} \operatorname{per}((bJ + sX)[S]).$$

A second derivative chooses two ordered directed X-edges with distinct rows and distinct columns. When their union contains $m = 2, 3, 4$ vertices, the coefficient of that ordered product in P''/P is

$$D_m = \frac{\mathbb{E}[(K-2)_{m-2}\, \mathbf{1}_{\{K \geq 2\}}]}{b^2 (n)_m}, \qquad K = n - (N_t \mid N_t \leq n). \tag{3}$$

Here $(x)_j$ is a falling factorial. A principal set of size k in the expansion has aggregate weight proportional to $\binom{n}{k} a^{n-k} k! b^k$, so its size has law K. For $k \geq m$ the chance it contains the designated vertices is $(k)_m/(n)_m$, and the two specified edges leave $(k-2)!$ permutations instead of $k!$. The coefficient also contains b^{-2}. Events with $2 \leq k < m$ give zero through the falling factorial in (3).

Write $F_2 = \|X\|_F^2$ and $s_0 = \sum_i r_i$. The sums of ordered edge products of each type are

$$\begin{aligned} &\Sigma_2 = F_2, && \Sigma_3 = 2(\|r\|^2 - F_2), \\ &\Sigma_4 = s_0^2 - 2 \operatorname{Re} \sum_i r_i^2 + \sum_{i,j} X_{ij}^2 - \Sigma_2 - \Sigma_3. \end{aligned} \tag{4}$$

The two-vertex pairs are precisely the reciprocal edges. The three-vertex pairs have one common endpoint, outgoing for one edge and incoming for the other. The oriented sum for each order is $\|r\|^2$ before removing reciprocal edges. For the total of all admissible pairs, subtract common-source and common-target pairs from s_0^2, then add identical edges once. This yields (4), and $P''/P = \sum_{m=2}^{4} D_m \Sigma_m$. At zero row sums the relative coefficients Σ_m/F_2 are $(1, -2, 2)$ for real X and $(1, -2, 0)$ for imaginary X.

At $t = t_n$, the conditioning in (3) and the truncation at $K \geq 2$ alter the polynomial moments by exponentially small errors times powers of n. This follows from Poisson exponential tail estimates for each fixed polynomial degree. The conclusion holds after one t-derivative by the identity $\partial_t \mathbb{E} H(N_t) = \mathbb{E}[H(N_t + 1) - H(N_t)]$, including for the truncated polynomials and the conditioning probability. The three unconditioned polynomial numerators are

$$1, \qquad n - t - 2, \qquad (n - t - 2)(n - t - 3) + t.$$

Since $\partial_b t = -t/b$ at fixed a, differentiation gives

$$\frac{\partial_b D_2}{n} \longrightarrow -\frac{1}{4}, \qquad \frac{\partial_b D_3}{n} \longrightarrow -\frac{1}{16}, \qquad \frac{\partial_b D_4}{n} \longrightarrow 0.$$

For D_4, the leading terms from differentiating the numerator and b^{-2} cancel, using $t_n \sim n/2$, $b_n \sim 2/n$. When $r = 0$, the mixing term in (2) vanishes. The two triples in (4) then give the two $-1/8 + o(1)$ coefficients.

4 Real star directions

Let v be a real vector of sum zero. Set $w_i = b_n + \eta v_i$ and

$$B(\eta) = \operatorname{diag}(1 - w_i) + xx^*, \qquad x_i = \sqrt{w_i}.$$

For η near zero this is a correlation matrix. Its tangent is the real star $X_{ij} = (v_i + v_j)/2$ for $i \neq j$, with $\|X\|_F^2 = (n-2)\|v\|^2/2$. Stationarity at B_n means the second derivative along this family equals the affine Hessian on X.

For $T \geq 0$ put $f_i = 1 + (T-1)w_i$, and use the probability density proportional to $e^{-T}\prod_i f_i$. The normalized raw cofactor matrix is the expectation of

$$R(T) = \operatorname{diag}(h(w_i)) + Tgg^*, \qquad h(w) = \frac{1-w}{[1+(T-1)w]^2}, \qquad g_i = g(w_i) = \frac{\sqrt{w_i}}{1+(T-1)w_i}. \quad (5)$$

To see this, expand the permanent along the diagonal part and represent the factorials of the rank-one principal submatrices by gamma integrals. A diagonal cofactor omits one f_i. An off-diagonal cofactor has an unmatched row and column, producing the factor $\sqrt{w_i w_j}T$ and omitting f_i, f_j. The diagonal of (5) equals $1/f_i$, as required.

At $\eta = 0$ write $b = b_n$, $a = 1-b$, and $f = a + bT$. Primes on h, g denote w-derivatives evaluated at b. The logarithm of the unnormalized density has first derivative zero pointwise and second derivative $-\|v\|^2(T-1)^2/f^2$. The row sum of $R(T)$ at this point is $\lambda_T = h + nTg^2$, and $\Delta = n\mathbb{E}(Tg^2)$. Differentiating (5) shows that the first derivative of $R(T)e$ is $(h' + nTgg')v/\sqrt{n}$. The projection of the second derivative on e gives $\|v\|^2(h''/n + 2Tgg'')$. Differentiating the expectation and applying the simple-eigenvalue formula thus give

$$\frac{D_X^2\Lambda}{\|v\|^2} = \mathbb{E}[h''/n + 2Tgg''] - \operatorname{Cov}\big(\lambda_T, (T-1)^2/f^2\big) + \frac{2}{n\Delta}\big(\mathbb{E}[h' + nTgg']\big)^2. \quad (6)$$

The random variable $T + t_n$ has the $\Gamma(n+1, 1)$ law conditioned to be at least t_n, since $f = b(T + t_n)$. Proposition 2 and gamma moments imply $T/n \to 1/2$ in probability and $\mathbb{E}|T - t_n|^k = O_k(n^{k/2})$ for $k = 2, 4$. Also

$$\frac{1}{4a} - Tg^2 = \frac{(a-bT)^2}{4a(a+bT)^2}.$$

This identity and the fourth moment give $\operatorname{Var}\lambda_T = O(1)$, while $\Delta \asymp n$. The function $(T-1)^2/f^2$ has derivative $2(T-1)/f^3$, uniformly $O(n)$, giving variance $O(n^3)$ using the second moment about t_n. The covariance in (6) is thus $O(n^{3/2})$. Direct differentiation gives $h''/n = O(n)$ and $h' = O(n)$ uniformly. Moreover

$$g' = \frac{1-(T-1)b}{2\sqrt{b}f^2}.$$

The numerator has second moment $O(n^{-1})$, and $T/f^3 = O(n)$, so $\mathbb{E}|Tgg'| = O(\sqrt{n})$. The last term in (6) is $O(n)$.

The remaining contribution is determined by

$$gg'' = -\frac{1}{4bf^2} - \frac{T-1}{f^3} + \frac{2b(T-1)^2}{f^4}.$$

Set $q = (T-1)b$; then $q \to 1$ in probability. At $q = 1$ the right side equals $-1/(16b)$. The rational expression for $2Tgg''/n^2$ is uniformly bounded for $q \geq -b$, so $T/n \to 1/2$ and $nb \to 2$ give convergence in mean to $-1/32$. Equation (6) is therefore $-n^2/32 + o(n^2)$. Dividing by $(n-2)/2$ in the norm identity yields the $-n/16 + o(n)$ coefficient.

5 The constant direction and the phase orbit

On a fixed relative neighborhood of $t_n \sim n/2$, (1) is $(1+t)(1-t/n)$ plus an exponentially small correction times a polynomial in n, also after up to two t-derivatives. In fact the atom $p_n(t) =$

$\mathbb{P}(N_t = n)$ satisfies $p_n' = (n/t - 1)p_n$, and the conditioning probability has derivative $-p_n(t)$. The scalar derivative at b_n is zero, and $dt/db = -b^{-2}$ along the correlation branch. Therefore its second derivative in b is $-2/(nb_n^4) + o(n^3)$. The constant off-diagonal direction with entries one has $\|X\|_F^2 = n(n-1)$, giving $-n/8 + o(n)$.

For $U = \mathrm{diag}(u_i)$ diagonal unitary, the retained row and column factors give

$$C(UBU^*)_{ij} = \overline{u}_i u_j C(B)_{ij}.$$

The permanent is unchanged and $C(UBU^*) = U^* C(B) U$. The orbit at B_n has tangent the imaginary gradients. Differentiating the invariance at the stationary point shows that the Hessian vanishes when either input is tangent to this orbit. The imaginary gradients are Frobenius orthogonal to the purely imaginary Hermitian matrices with zero row sums. The constant real direction has zero mixed terms with its orthogonal complement by permutation averaging. Complex conjugation separates the real and imaginary sectors. In the total-sum-zero real sector, (2) and (4) express the Hessian as a linear combination of Frobenius and row-sum squares. These forms have zero mixed terms between real stars and real zero-row-sum directions. These subspaces exhaust the tangent space, giving the claimed Hessian and its kernel.

An explicit local slice to the phase orbit is obtained by making the nonzero first-row off-diagonal entries real and positive by diagonal-unitary conjugation. Its tangent intersects the imaginary gradients only at zero. Restricting the Hessian to the slice therefore gives a negative definite form. The matrix B_n is positive definite, so the slice lies locally in the interior of the correlation domain. Taylor's theorem on the slice proves strict local maximality modulo the phase orbit and completes the proof of Theorem 1.

6 Exact and numerical maximality at small orders

The order-three constant is known exactly: by Theorem 73, every maximizer of Λ among 3×3 complex correlation matrices is a diagonal-unitary conjugate of a positive equicorrelation matrix, so γ_3 is the maximum of formula (1) at $n = 3$,

$$\gamma_3 = \Lambda(B_3) = 1.4228496379\ldots, \qquad t_3 = 1.7954900\ldots, \quad b_3 = 0.3577190\ldots,$$

in agreement with Table 1. We also maximized Λ numerically over all complex correlation matrices $B = D^{-1/2} X X^* D^{-1/2}$ of orders four and five, using random complex X, Ryser permanents for every cofactor, and Nelder–Mead ascent from 15 (order four) and 10 (order five) independent starting points. The largest values found were

$$1.645226619865\ (n = 4), \qquad 1.872765356\ (n = 5),$$

which agree with the optimal equicorrelation values 1.645226619865 and 1.872766516 of Table 1 to the accuracy of the ascent. No start produced a larger value. Order four is settled exactly in Section 62, and Section 66 treats order five. At order six the optimal equicorrelation matrix has $\Lambda(B_6) = 2.10437\ldots > 2$; it is therefore an explicit equicorrelation witness that the raw cofactor bound $C(B) \preceq 2(\mathrm{per}\, B) I$ fails at order six.

These computations, the order-three theorem, the local rigidity of Theorem 1, the global maximization among one-factor matrices in Theorem 3 and the exact Walsh-family maximization of Theorem 17 all identify the optimal equicorrelation matrix as the extremizer of the cofactor ratio.

7 One-factor correlation matrices

Let $C(A)_{ij}$ be the permanent of the submatrix obtained by deleting row i and column j. We study $\Lambda(A) = \lambda_{\max}(C(A))/\operatorname{per} A$. This is the raw cofactor ratio. The adjoint in [4] uses the transpose of this cofactor convention. Marcus and Merris give an inverse-matrix upper bound for the cofactor adjoint. Lavi [3, Section 4] bounds the raw cofactor ratio in small orders. Lemma 71 gives all-order bounds whose first discrepancy is the coefficient of $1/n$, with a scalar equicorrelation matrix giving the lower bound.

Feinsilver and McSorley [7, Theorem 5.3 and Lemma 5.4] give exponential moment formulas for equicorrelation subpermanents. The scalar evaluation here is consistent with those formulas. For a diagonal plus a positive rank-one factor, the positive residual case, after removing loading phases, has an inverse with nonpositive off-diagonal entries by the rank-one inverse formula. Sra's inequality [8] thus applies to its principal permanents. The present objective also uses all nonprincipal cofactors and their joint top eigenvalue. Our main step is a global concentration argument for this eigenvalue over arbitrary independent loadings. The proof includes the scalar calculation and uses no separate local Hessian theorem.

Theorem 3. *Let $A = \operatorname{diag}(1 - |z_i|^2) + zz^*$, where $z \in \mathbb{C}^n$ and $0 \le |z_i| \le 1$. Write $\Lambda(A) = \lambda_{\max}(C(A))/\operatorname{per} A$ for the raw permanental cofactor ratio. For all sufficiently large n, this class has a unique maximizing matrix up to diagonal-unitary conjugation. It is the positive equicorrelation matrix of scalar strength $b_n = 2/(n+1) + O(e^{-cn})$. The maximum is $(n+1)^2/(4n) + O(ne^{-cn})$ for an absolute $c > 0$. The class includes every diagonal-plus-one-positive-factor correlation matrix, without a small-loading assumption on any entry.*

8 Gaussian formula and the global new coordinate

Gauge the loadings to be nonnegative, put $r_i = |z_i|^2$, $a_i = 1 - r_i$, $d_i(T) = a_i + r_i T$, and $x_i(T) = r_i T/d_i(T)$ for $T > 0$. Expectations are with respect to the density proportional to $e^{-T}\prod_i d_i(T)$. Let G_i, Z_0 be independent standard circular complex Gaussians. The variables $\sqrt{a_i}G_i + \sqrt{r_i}Z_0$ have covariance A. By the complex Gaussian moment identity [9], the expectation of the product of squared moduli equals the permanent. For an off-diagonal cofactor the two omitted variables leave one conditional mean and its conjugate, while each remaining index contributes d_i. For a diagonal cofactor the factor d_i is simply omitted. Dividing by the scalar permanent and conditioning on $T = |Z_0|^2$ gives

$$M = C(A)/\operatorname{per} A = \mathbb{E}[\operatorname{diag}(a_i/d_i^2) + vv^{\mathsf{T}}], \qquad v_i = \sqrt{r_i T}/d_i.$$

All identities are for $T > 0$. The formulas below prove integrability even at singular residuals. They follow from $1/d_i = 1 - x_i + x_i/T$:

$$v_i^2 = x_i(1 - x_i) + x_i^2/T, \qquad a_i/d_i^2 = (1 - x_i)^2 + x_i(1 - x_i)/T.$$

Introduce an integer K by expanding $\prod_i d_i(T) = \sum_{k=0}^n c_k T^k$ with $c_k \ge 0$. In the joint law $K = k$ has probability proportional to $c_k k!$, and $T \mid K = k$ is Gamma of shape $k + 1$, rate one. Given T, the generating function of K is $\prod_i[(1 - x_i) + x_i u]$, so it is a sum of conditionally independent Bernoulli variables. In particular

$$\mathbb{E}\sum_i x_i/T = \mathbb{E}(K/T) = \Pr(K > 0) \le 1.$$

Therefore the expected diagonal operator norm is at most 2, the total rank-component trace is at most $n/4+1$, and

$$\mathbb{E}\sum_i (x_i - 1/2)^2 = O(1) \quad \text{if} \quad \lambda_{\max}(M) \geq n/4 + 1/2 - o(1).$$

For the last implication use $\operatorname{tr}\mathbb{E}[vv^{\mathsf{T}}] = n/4 - \mathbb{E}\sum(x_i - 1/2)^2 + \mathbb{E}\sum x_i^2/T$. Every sequence of maximizers meets the hypothesis by the scalar test in the last section.

By conditional Bernoulli variance and Cauchy–Schwarz, $\mathbb{E}(K - n/2)^2 \leq n/4 + n\mathbb{E}\sum_i (x_i - 1/2)^2 = O(n)$. The Gamma variance and mean also give $\mathbb{E}(T - n/2)^2 = O(n)$. Write $T_0 = n/2$ and use the bijective coordinates $p_i = x_i(T_0) \in [0,1]$, even at $r_i = 0, 1$. In detail

$$r_i = \frac{p_i}{T_0(1-p_i) + p_i}, \qquad x(p_i, t) = \frac{p_i t}{1 + p_i(t-1)}, \qquad t = T/T_0.$$

Choose one T with $|T - T_0| \leq C\sqrt{n}$ and $\sum_i (x_i(T) - 1/2)^2 \leq C$ by the two expectation bounds and a union bound. The derivative $\partial_t x = p(1-p)/[1 + p(t-1)]^2$ is uniformly bounded on $1/2 \leq t \leq 3/2$ for every $p \in [0,1]$. Thus $\|p - \frac{1}{2}\mathbf{1}\|_2 = O(1)$ on any near-maximizing sequence, uniformly including its outlier coordinates.

9 The two-term law with strong loading outliers

The next calculation applies to every sequence $\|p - \frac{1}{2}\mathbf{1}\|_2 = O(1)$, including sequences with some p_i equal to 0 or 1. Put $b = n^{-1}\sum p_i$, $\delta_i = p_i - b$, $u = \sqrt{n}(b - 1/2)$. Then $u, \|\delta\|_2 = O(1)$, and we claim

$$\lambda_{\max}(M) = \frac{n}{4} + \frac{1}{2} - 4u^2 - \|\delta\|_2^2 + o(1). \tag{7}$$

The error is uniform for each fixed bound on $\|p - \frac{1}{2}\mathbf{1}\|_2$. This follows from the estimates below, or by applying sequential convergence to every sequence within the bound.

The tilt density, after canceling $\prod_i d_i(T_0)$, uses $\prod_i [1 + p_i(t-1)]$. Its ratio to the scalar unnormalized density at b is e^{ℓ} with

$$\ell = \sum_i \log(1 + p_i(t-1)) - n\log(1 + b(t-1)) \leq 0.$$

On $1/2 \leq t \leq 3/2$, Taylor cancellation gives $|\ell| \leq C(t-1)^2\|\delta\|_2^2$. For the scalar density, $T + s_b$ is Gamma of shape $n+1$, conditioned to exceed $s_b = T_0(1-b)/b$. Here $s_b = n/2 + O(\sqrt{n})$, so truncation is exponentially small. The centered variable $Y = (T - (n - s_b))/\sqrt{n}$ converges with all fixed moments to $N(0,1)$, and its moments are uniformly bounded. This holds first for Gamma by the independent-exponential representation and exponential moments, then for the heterogeneous tilt because $e^{\ell} \leq 1$ globally, tends to one uniformly for bounded Y, and its normalization tends to one. All fixed polynomial $\tau = T/n$ moments outside $[1/4, 3/4]$ are exponentially small by the Gamma Chernoff bound.

We also need the contribution of $1/T$ near zero. Write $F(T) = \prod_i [1 + p_i(t-1)]$, $Z = \int e^{-T} F(T)\, dT$. Since $F'(T) = F(T)\sum_i x_i/T$, integration by parts on $[0, n/4]$ gives

$$\mathbb{E}\left[\mathbf{1}_{T<n/4} \sum_i x_i/T\right] \leq e^{-n/4} F(n/4)/Z + \Pr(T < n/4).$$

The Lebesgue density at $n/4$ is exponentially small by comparison with the scalar shifted Gamma density, using $F \le [1 + b(t-1)]^n$, the bounded normalization ratio, and Stirling's formula there. For $T > 3n/4$, $\sum_i x_i/T \le 4/3$. Thus the expected squared norm of v and the expected diagonal operator on the complement of $[n/4, 3n/4]$ are exponentially small.

On the central interval let $D(p,t) = (x(p,t) - 1/2)^2$. At $t = 1$, $D(p,1) = (p - 1/2)^2$ exactly. Taylor cancellation in strength, using bounded mixed derivatives on $[0,1] \times [1/2, 3/2]$, shows

$$\sum_i D(p_i, t) = nD(b,t) + \|\delta\|_2^2 + O(|t-1|\|\delta\|_2^2).$$

The expected error tends to zero. Also

$$\sqrt{n}(x(b,t) - 1/2) = 2u + Y/2 + o_{L^2}(1).$$

Indeed $n - s_b = n/2 + 2u\sqrt{n} + O(1)$; expand x at $(1/2, 1)$, where its derivatives are 1 and $1/4$, and use the uniformly bounded fourth moments of Y and bounded u. Thus $\mathbb{E}\sum_i D(p_i,t) = 4u^2 + 1/4 + \|\delta\|_2^2 + o(1)$. Meanwhile $\mathbb{E}\sum_i x_i^2/T \to 1/2$: on the central interval $\sum_i x_i^2/n \to 1/4$ in mean by the just-bounded square loss and Cauchy–Schwarz, while $T/n \to 1/2$ in mean and $n/T \le 4$; the tails were handled above.

For the rank and diagonal spectral correction, use the exact expression

$$v_i(T) = \frac{\sqrt{p_i(1 - p_i + p_i/T_0)}\sqrt{t}}{1 + p_i(t-1)}.$$

Choose $e = v(T_0)/\|v(T_0)\|$, $P = I - ee^{\mathsf{T}}$ and $c(t) = \sqrt{t}/[1 + b(t-1)]$. The difference $v_i(T) - c(t)v_i(T_0)$ equals

$$v_i(T_0)\sqrt{t}\left(\frac{1}{1 + p_i(t-1)} - \frac{1}{1 + b(t-1)}\right).$$

On the central interval this gives $\|Pv\| \le C|t-1|\|\delta\|_2$ without differentiating the square root at the endpoints. Hence $\mathbb{E}\|Pv\|^2 = O(n^{-1})$. The function $v_i(T_0)$ is uniformly Lipschitz on a fixed neighborhood of $1/2$; outside that neighborhood its difference from the scalar value is bounded by a constant times $|p_i - b|$. Thus $\|e - \mathbf{1}/\sqrt{n}\| = O(n^{-1/2})$. The expected diagonal operator is bounded. Its quadratic form on e tends to $1/4$ since replacing e by $\mathbf{1}/\sqrt{n}$ loses $o(1)$, the average of $(1 - x_i)^2$ tends to $1/4$, and the remaining average is bounded by $1/n$. The lower block PMP is $O(1)$ in norm. The off-diagonal block PMe has bounded norm by Cauchy–Schwarz from $\mathbb{E}\|Pv\|^2 = O(1/n)$ and $\mathbb{E}\|v\|^2 = O(n)$. The exact scalar Schur complement then shows $\lambda_{\max}(M) = e^{\mathsf{T}}Me + O(1/n) = \mathbb{E}\|v\|^2 + 1/4 + o(1)$. This gives (7).

In particular maximizing sequences satisfy $\|p - \frac{1}{2}\mathbf{1}\|_2 \to 0$. The conclusion rules out even a single loading bounded away from $1/2$ in the p coordinates.

10 Uniform local differentiation and exactness

We first prove the dimension-free estimate $\|D_p^3 \Lambda\| \le C$ on a fixed small Euclidean ball around $p = \frac{1}{2}\mathbf{1}$. The auxiliary variables $\beta_i = nr_i = 2p_i/(1 - p_i + 2p_i/n)$ have bounded Euclidean differentials through order three, since the map is entrywise and a product of unit Euclidean directions has norm at most one. It suffices to work on a small Euclidean ball in β about $\beta_* \mathbf{1}$, where $\beta_* = 2/(1 + 2/n)$, and transfer the derivative bounds back to p. The local estimates use $b = n^{-1}\sum_i \beta_i$ and $\delta_i = \beta_i - b$

on the ball $\|\beta - \beta_*\mathbf{1}\|_2 \le \varepsilon$, where $\varepsilon > 0$ is fixed and small. With $\tau = T/n$ the diagonal residuals are $a_i = 1 - \beta_i/n$ and $d_i = 1 + \beta_i(\tau - 1/n)$. Write

$$D_n(\beta_i, \tau) = \frac{(\beta_i(\tau + 1/n) - 1)^2}{4a_i d_i^2}.$$

The pointwise trace identity $\|v\|^2 = \sum_i (4a_i)^{-1} - \sum_i D_n(\beta_i, \tau)$ follows by subtraction. The scalar β-tilt translates T by $n/b - 1 = n/2 + O(\sqrt{n})$. The likelihood ratio of the heterogeneous law against this scalar reference is controlled as in the preceding section by Taylor expansion of $\sum_i \log d_i$ around the scalar mean.

For the tilt, put $X = \sqrt{n}(\tau - 1/2)$. The shifted-Gamma comparison from the preceding section, in the equivalent strength variables, applies uniformly since $b = 2 + O(n^{-1/2})$ and $\|\delta\|_2 \le \varepsilon$. To be explicit about the uniform denominator, on $\tau \in [1/4, 3/4]$ the scalar law has mass at least $1/2$ for large n and the likelihood ratio there is bounded below. Thus all fixed moments of X are uniformly bounded, and all polynomial moments of τ outside that interval are exponentially small. Before differentiating the normalized tilted measure in the strengths, subtract its log-density value at $\tau_0 = 1/2$; this does not change the normalized law. The first log-derivative in any unit direction h is

$$\sum_i h_i \left(\frac{\tau - 1/n}{1 + \beta_i(\tau - 1/n)} - \frac{\tau_0 - 1/n}{1 + \beta_i(\tau_0 - 1/n)} \right).$$

Its bound is $C|X|$ because the τ derivative of each fraction is $d_i^{-2} \le 4$ and $\|h\|_1 \le \sqrt{n}$. For higher mixed derivatives through order three, $\sum_i |h_i^{(1)} \cdots h_i^{(k)}| \le 1$ for $k \ge 2$ for unit directions. The derivative formulas for $\log(1 + \beta_i(\tau - 1/n))$ and $d_i \ge 1/2$ give uniform polynomial bounds in $|X|$. Differentiating the quotient of unnormalized integrals gives a finite sum of expectations of observable derivatives times log derivatives and of products of such expectations. It therefore preserves uniform moment bounds, even when integrands are matrix-valued, by taking operator norms and using Hölder. Pointwise bounds on the central interval with exponentially small tails yield the same n-dependent estimates after differentiation.

For the block estimates choose $u_0 = v(\tau_0)$, $e = u_0/\|u_0\|$ and $P = I - ee^{\mathsf{T}}$ as before, and write $w(\tau) = \sqrt{\beta_*\tau}/[1 + \beta_*(\tau - 1/n)]$ for the scalar component at the base strength. The residual $W = v(\tau) - [w(\tau)/w(\tau_0)]u_0$ vanishes at τ_0, and at $\beta = \beta_*\mathbf{1}$. Each component depends only on its own strength. On the central interval, coordinate differentiation and the mean value theorem in τ give $\|D^j W\| \le C|\tau - \tau_0|$ through order three, interpreted in mixed unit directions for $j > 0$; for $j = 0$ use the bound with additional factor $\|\beta - \beta_*\mathbf{1}\|_2$. The norm of an entrywise product of unit Euclidean directions is at most one. The normalization of u_0, whose norm is comparable to $\sqrt{n}$, gives bounded derivatives of e and P to order three. Thus $v_\perp = Pv = PW$ and its derivatives are $O(|X|/\sqrt{n})$, and $e^{\mathsf{T}}v$ and its derivatives are $O(\sqrt{n})$ on the central interval. On its complement strength derivatives in this ball grow at most polynomially in n, τ, since β_i is bounded below and $d_i \ge 1/2$. The estimate for the tilt therefore applies on the full interval.

Denote $D = \operatorname{diag}(a_i/d_i^2)$, $A_0 = e^{\mathsf{T}}Me$, $q = PMe$ and $R = PMP$. The preceding estimates show $\mathbb{E}_\beta\|v_\perp\|^2 = O(1/n)$ through three derivatives and $q, R = O(1)$ through three derivatives. For q use $(e^{\mathsf{T}}v)v_\perp$ and for R use $v_\perp v_\perp^{\mathsf{T}}$, along with their bounded diagonal terms. For A_0 use

$$A_0 = \sum_i (4a_i)^{-1} - \mathbb{E}_\beta \sum_i D_n(\beta_i, \tau) - \mathbb{E}_\beta\|v_\perp\|^2 + \mathbb{E}_\beta e^{\mathsf{T}}De.$$

The first term minus $n/4$ has bounded derivatives to order three, using $\|h\|_1 \le \sqrt{n}$ at order one and the product bound at higher orders. In the loss, $k_i = \beta_i(\tau + 1/n) - 1$ has ℓ^2 norm $O(1 + |X|)$ from

the local ball, and its first directional derivative has norm at most $|\tau + 1/n|$. Its higher derivatives vanish. Write the loss as $k^{\mathsf{T}} \operatorname{diag}[(4a_i d_i^2)^{-1}]k$. The central diagonal derivatives are bounded in operator norm. This shows $A_0 - n/4 = O(1)$ through order three.

The top eigenvalue is uniformly simple since $A_0 \sim n/4$ and the lower block R is bounded. It satisfies the exact scalar equation

$$\lambda = A_0 + q^{\mathsf{T}}(\lambda I - R)^{-1} q.$$

Here R can be treated on the full space because $Re = 0$, $q \perp e$, and $\lambda > 0$. The correction function and its first three partial derivatives in β at fixed $\lambda \sim n/4$ are $O(1/n)$, and each additional λ-derivative gains one factor $O(1/n)$. This follows by the resolvent product rule using the bounds on q, R. The implicit derivative coefficient at $d\lambda$ is $1 + q^{\mathsf{T}}(\lambda I - R)^{-2} q \geq 1$. Differentiating successively therefore bounds all three total derivatives uniformly.

We next evaluate the scalar maximum. All $r_i = r \in (0, 1)$ give $s = (1 - r)/r$ and $U = T + s$. By symmetry and positivity the all-ones vector is the top eigenvector. Its Rayleigh quotient is

$$\Lambda = (s + 1)\big[n\mathbb{E}(1/U) - (n - 1)s\mathbb{E}(1/U^2)\big].$$

Here U has the $\Gamma(n + 1, 1)$ law conditioned on $U > s$. For s/n in a fixed neighborhood of $1/2$, replacing the two moments by their unconditioned values $1/n$ and $1/[n(n-1)]$ gives $\Lambda = (s+1)(1 - s/n) + O(ne^{-cn})$. For the omitted inverse moments, their integrals factor as the corresponding lower-shape Gamma probabilities times the exact inverse moments, so each relative error is exponentially small. The quadratic is maximized at $s = (n - 1)/2$ with value $(n + 1)^2/(4n)$. In particular this scalar test supplies $n/4 + 1/2 + o(1)$ for localization.

Let p_n be any maximizing scalar coordinate in $[0, 1]$. It is near $1/2$ by (7) and the prior localization, and therefore interior. The quadratic estimate shows $\Lambda(p_n \mathbf{1}) = (n + 1)^2/(4n) + O(ne^{-cn})$ and $s(p_n) = (n - 1)/2 + O(ne^{-cn/2})$, which gives the asserted off-diagonal strength after decreasing c. For $F_n(p) = \Lambda(A(p))$, permutation invariance and scalar stationarity give $DF_n(p_n \mathbf{1}) = 0$.

Let h be a unit Euclidean direction, $m = n^{-1/2} \sum_i h_i$. For a fixed small $t > 0$, apply (7) at $p_n \mathbf{1}$ and $p_n \mathbf{1} \pm th$. The symmetric second difference divided by t^2 is $-8m^2 - 2(1 - m^2) + o(1)$ uniformly in h as $n \to \infty$; the mean offset cancels. The uniform C^3 bound gives an additional error at most Ct when replacing that difference by the Hessian at the center. First fix t small and then take n large. This proves $D^2 F_n \preceq -I$ at the center. The third derivative bound makes F_n strictly concave on a fixed smaller ball around the center. Global localization puts every full-class maximizing sequence in this ball, so the stationary scalar profile is the unique maximizer of the full cube in sufficiently large order. Gauge restoration accounts for precisely the stated matrix equalities.

11 Bounded fixed-rank corrections: setup and main result

For a correlation matrix A, let $C(A)_{ij}$ be the permanent with row i and column j deleted. Write $\Lambda(A) = \lambda_{\max}(C(A))/\operatorname{per} A$. Bounds of Marcus and Merris [4] and Lavi [3, Section 4] concern this raw cofactor observable. The all-order bounds of Lemma 71 agree up to and including the constant term, and Theorem 11 determines the $1/n$ coefficient. A local Hessian calculation Theorem 1 gives strict local maximality at the scalar optimum modulo diagonal-unitary conjugation. The cap-free one-factor result Theorem 3 gives exact eventual maximization for independently variable single loadings. The fixed balanced Walsh result Theorem 17 treats certain additional factors with symmetry.

Here the latent dimension r is fixed. The full matrix has diagonal one and off-diagonal entries $u_i^* u_j/n$, where $u_i \in \mathbb{C}^r$ and $\|u_i\| \le B$. Equivalently one adds a positive diagonal to a rank-at-most-r Gram matrix in all sufficiently large orders. The restriction is on the correction as well as every row at its natural scale. The cofactor spectrum depends on the whole correlation matrix.

Low-rank permanent methods and Gaussian representations include Barvinok [10, 11] and Fassino, Pistone, and Rogantin [9]. McCullagh [12] studies determinant approximations for scalar permanents. Our comparison concerns the top eigenvalue of the entire normalized cofactor matrix. The proof bounds losses due to variation of both the amplitudes and phases of the conditional cofactor vector.

Theorem 4. *Fix $r \ge 1$ and $B > \sqrt{2}$. Let $\gamma_{n,r,B}$ be the maximum of Λ in the class just defined, and let $A_{*,n}$ denote the maximizing positive equicorrelation matrix. Then*

$$0 \le \gamma_{n,r,B} - \Lambda(A_{*,n}) = O_{r,B}(n^{-3}), \qquad \gamma_{n,r,B} = \frac{n}{4} + \frac{1}{2} + \frac{1}{4n} + O_{r,B}(n^{-3}).$$

Any sequence with $\Lambda \ge n/4 + 1/2 - o(1)$ admits diagonal row gauges and a latent unit direction h such that, for $s = \|n^{-1}\sum_i u_i\|$, $\ell_i = u_i^* h$ and $w_i = P_{h^\perp} u_i$,

$$\sum_i w_i = 0, \qquad \rho_n = \frac{1}{n}\sum_i \|w_i\|^2 \longrightarrow 0, \qquad \sqrt{n}(s^2 - 2) \longrightarrow 0, \qquad \sum_i |\ell_i - s|^2 \longrightarrow 0.$$

In particular the gauged matrices converge in Frobenius norm to equicorrelation with off-diagonal $2/n$. For any near-maximizing sequence with $\Lambda > \Lambda(A_{*,n})$ the transverse energy ρ_n in this frame is $O_{r,B}(n^{-1})$.

12 Localizing a near-maximizing profile

Use $a_i = 1 - \|u_i\|^2/n$, $m_i = u_i^* y$, $d_i = a_i + |m_i|^2$, $v_i = m_i/d_i$. The latent Gaussian divided by $\sqrt{n}$ has tilted density proportional to $e^{nf(y)}$ with $f(y) = -\|y\|^2 + n^{-1}\sum_i \log d_i$. By the complex Gaussian moment identity [9] and conditioning on y, the normalized cofactor matrix, up to transpose, is $M = \mathbb{E}[D + vv^*]$, $D_{ii} = a_i/d_i^2$. Choose a unit top eigenvector e, gauge each row to make its coordinate of e nonnegative, and put $P = I - ee^*$. The exact identity

$$\Lambda = \sum_i (4a_i)^{-1} + \mathbb{E}e^* D e - \mathbb{E}\|Z\|^2 - \mathbb{E}\|Pv\|^2, \qquad Z_i = \frac{|m_i|^2 - a_i}{2\sqrt{a_i} d_i} \tag{8}$$

gives bounded expected squared losses, since $a_i \ge 1/2$, the sum is $n/4 + O(1)$ and $D \preceq 2I$. Also $|v_i| \le 1/(2\sqrt{a_i})$, and $(1/(2\sqrt{a_i}) - |v_i|)^2 \le Z_i^2$. Thus $\mathbb{E}\||v| - \mathbf{1}/2\|^2 = O(1)$.

Let $\alpha = e^* v$. Pointwise $\||v| - |\alpha| e\| \le \|Pv\|$. Choose one realization with both losses bounded by a constant, which implies $\||\alpha| e - \mathbf{1}/2\| = O(1)$. Taking norms and subtracting gives $\|e - \mathbf{1}/\sqrt{n}\| = O(n^{-1/2})$. Set $\phi = \arg\alpha$, choosing it arbitrarily at zero. The same triangle inequalities in expectation now give $\mathbb{E}\|v - e^{i\phi}\mathbf{1}/2\|^2 = O(1)$.

The tilted law has exponentially small fixed polynomial tails outside a sufficiently large fixed ball. Its original Gaussian normalization is at least $\prod_i a_i \ge e^{-2B^2}$, and the numerator is bounded by a fixed-dimensional Gaussian factor times $(1 + B^2\|y\|^2)^n$. Taking the ball large enough gives any needed exponential tail. On the ball, $m_i - 2v_i = m_i(d_i - 2)/d_i$ and $d_i - 2 = (|m_i|^2 - a_i) + 2(a_i - 1)$. Thus $\sum_i |m_i - 2v_i|^2 \le C\sum_i Z_i^2 + O(1/n)$ on the ball. The tail contribution is exponential. It follows that

$$\mathbb{E}\|m - e^{i\phi}\mathbf{1}\|^2 = O(1). \tag{9}$$

The score gives more information than this projected alignment:

$$\partial_{\bar y} f = -y + \frac{1}{n}\sum_i u_i v_i, \qquad \mathbb{E}\|\partial_{\bar y} f\|^2 = -\frac{1}{4n}\mathbb{E}\,\Delta f = O(n^{-1}).$$

The score identity follows by integrating the real gradient by parts and using $\|\nabla f\|^2 = 4\|\partial_{\bar y} f\|^2$. The same score identity underlies [13, Theorem 1]. There is no boundary term by the Gaussian tail. The Laplacian equals $-4r + (4/n)\sum_i \|u_i\|^2 a_i/d_i^2$, so the expected squared score is at most r/n. With $\bar u = n^{-1}\sum_i u_i$, the spectral alignment of v and Cauchy–Schwarz therefore yield

$$\mathbb{E}\|y - e^{i\phi}\bar u/2\|^2 = O(n^{-1}). \tag{10}$$

The operator with rows u_i^* has norm at most $B\sqrt{n}$. From (9) and (10), $\sum_i |u_i^*\bar u/2 - 1|^2 = O(1)$. Averaging gives $s^2 = 2 + O(n^{-1/2})$ and, with $h = \bar u/s$, $\|\ell - s\mathbf{1}\|_2 = O(1)$. The definition of h gives $\sum_i w_i = 0$. These conclusions were obtained without assuming an initial transverse covariance gap.

12.1 A nondegenerate limit from tightness

Take the common-phase section $y = e^{i\theta}(\sqrt{x}h + \eta)$, $\theta = \arg h^* y$, $\eta \in h^\perp$. The measure is a constant times $dx\,d\eta$ after removing θ, and the conditional means in the section are $m_i = \ell_i\sqrt{x} + w_i^*\eta$. Put $\beta = s^2$ and $x_0 = 1 - 1/\beta$. The identity $\beta/4 - x_0 = (\beta-2)^2/(4\beta) = O(1/n)$ and (10) imply tightness of $(\zeta,\xi) = (\sqrt{n}(x-x_0), \sqrt{n}\eta)$. Indeed $\sqrt{n}\,\big||h^*y| - s/2\big|$ and $\sqrt{n}\|\eta\|$ have bounded second moments, and $|h^*y| + s/2$ is tight; the product formula gives the first tightness assertion. Identify the varying $h^\perp$ with a fixed complex space of dimension $r-1$ by a unitary frame.

On every compact set of scaled coordinates, the log density after subtracting its value at $(x_0, 0)$ is

$$-\frac{1}{2}\zeta^2 + \ell_n^{\mathrm{lin}}\cdot\xi - \xi^{\mathsf T} H_n \xi + o(1), \qquad H_n = I - Q_n/2, \quad \xi^{\mathsf T} Q_n \xi = \frac{1}{n}\sum_i (\operatorname{Im} w_i^*\xi)^2. \tag{11}$$

Here ℓ_n^{lin} is a bounded real vector. For the Taylor verification, $\sum(\ell_i - s) = 0$, $\sum|\ell_i - s|^2 = O(1)$, and $a_i = 1 + O(1/n)$ yield $\partial_x f(x_0, 0) = O(1/n)$, by subtracting the scalar derivative and canceling its linear term. The η gradient is $O(n^{-1/2})$ by subtracting the constant scalar coefficient of w_i, using $\sum w_i = 0$ and $\sum|\ell_i - s| = O(\sqrt{n})$. The Hessian at the center differs by $o(1)$ from its value with all $\ell_i = s$, $a_i = 1$: coordinatewise derivatives are uniformly smooth on the bounded parameters, and their average perturbation is $O(n^{-1/2})$. In that scalar array the x Hessian is -1, mixed Hessian vanishes by centering, and transverse Hessian tends to $-2I + Q_n$ up to $o(1)$ from $\beta \to 2$. This uses $\log(1 + |1+z|^2) = \log 2 + \operatorname{Re} z + (\operatorname{Im} z)^2/2 + O(|z|^3)$. The third derivatives in a fixed neighborhood are uniformly bounded, which proves (11).

Take a convergent subsequence of the bounded matrices Q_n and vectors ℓ_n^{lin}. Let N_n be the total normalization after subtracting the local exponent and rescaling volume. From (11), a fixed box yields a positive lower bound on N_n. By tightness choose a large fixed box with probability at least $1/2$; its unnormalized mass is bounded by (11), giving an upper bound on N_n. Any subsequential limit of N_n is finite and positive. Local convergence and tightness then give the limiting probability density proportional to the limiting exponential in (11). Its integrability forces $H = I - Q/2$ positive definite, even though there was no gap assumption. Thus ζ is centered Gaussian of variance one and independent of ξ, whose mean possibly shifts but whose covariance is $H^{-1}/2$.

12.2 Transverse variance and collapse

Let R_n be the corresponding real-form matrix for $\operatorname{Re} w_i^*\xi$. Take further subsequences $R_n \to R$; $\operatorname{tr} R_n = \operatorname{tr} Q_n = n^{-1}\sum \|w_i\|^2$. On any fixed compact scaled set,

$$Z = Z^0 + \frac{1}{2\sqrt{n}}\Big(\zeta\mathbf{1} + (\operatorname{Re} w_i^*\xi)_i\Big) + o_{\ell^2}(1), \qquad Pv = Pv^0 + \frac{i}{2\sqrt{n}}(\operatorname{Im} w_i^*\xi)_i + o_{\ell^2}(1). \tag{12}$$

Superscript zero means $(x,\eta) = (x_0, 0)$. In the second identity use $\sum w_i = 0$, $\|e - \mathbf{1}/\sqrt{n}\| = O(n^{-1/2})$, and the fact that the real differential of $m/(1+|m|^2)$ at $m = 1$ vanishes on the real axis and equals $i\operatorname{Im}/2$ on the imaginary axis. In the first identity the differential of Z at $m = 1, a = 1$ is $\operatorname{Re}/2$, and $\partial_x m = s/(2\sqrt{x_0}) \to 1$. To justify replacing the derivative arrays by their common scalar-center values, their pointwise errors are bounded by $C(|\ell_i - s| + |\beta - 2| + 1/n)$; the squared sum is $O(1)$ and each derivative is multiplied by $1/\sqrt{n}$. Uniform Taylor remainders also have Euclidean norm $O(n^{-1/2})$ on each fixed scaled box. The offsets Z^0, Pv^0 are bounded in Euclidean norm by $\|\ell - s\mathbf{1}\|_2 = O(1)$ and the uniform comparison of e.

The two nonnegative quadratic expressions in (12) have bounded coefficients as polynomials in the fixed-dimensional scaled variables. On a further subsequence these coefficients converge. Their Gaussian expectations are at least the variances of their linear parts, since the squared norm of the deterministic offset plus mean is nonnegative. Equivalently this follows for each n under the fixed limiting Gaussian and then by coefficient convergence. Local density convergence on increasing boxes gives, by nonnegativity and monotone convergence on the limiting polynomials,

$$\liminf \mathbb{E}\|Z\|^2 \geq \frac{1}{4} + \frac{1}{8}\operatorname{tr}(RH^{-1}), \qquad \liminf \mathbb{E}\|Pv\|^2 \geq \frac{1}{8}\operatorname{tr}(QH^{-1}).$$

No uniform integrability of the full scaled law beyond tightness is needed for these lower bounds.

Finally $\mathbb{E}e^*De \to 1/4$. Replacing the unit vector by $\mathbf{1}/\sqrt{n}$ costs $o(1)$ because $D \preceq 2I$; (9) and the bounded derivative of $a/(a+|m|^2)^2$ in m then give the limiting average. Also

$$\sum_i (4a_i)^{-1} = n/4 + 1/2 + \operatorname{tr} Q_n/4 + o(1).$$

Substitution in (8) gives

$$\limsup(\Lambda - n/4 - 1/2) \leq -\frac{1}{8}\operatorname{tr}\left[(Q+R)(H^{-1} - I)\right] \leq -\frac{1}{16}\operatorname{tr}(Q^2).$$

The assumed near-optimality forces $Q = 0$, hence $\operatorname{tr} Q_n \to 0$ along the full sequence since all convergent subsequences obey this conclusion. The boundedness of the longitudinal error gives $\|\ell\ell^*/n - s^2\mathbf{1}\mathbf{1}^*/n\|_{\mathrm{F}} = O(n^{-1/2})$, allowing conjugation of ℓ if needed by Gram convention. The transverse Gram contribution is positive and its Frobenius norm is at most its trace $\operatorname{tr} Q_n$. Subtracting the diagonal does not increase Frobenius norm. This proves the transverse and Frobenius assertions of Theorem 4.

12.3 Longitudinal convergence

The transverse variance inequality gives $Q_n \to 0$ under $\Lambda \geq n/4 + 1/2 - o(1)$. The transverse linear coefficient in the scaled density expansion then tends to zero: after subtracting the scalar derivative coefficient, its norm is bounded by

$$Cn^{-1/2}\sum_i \|w_i\|(|\ell_i - s| + n^{-1}) = o(1).$$

This follows from Cauchy–Schwarz, $\|\ell - s\mathbf{1}\|_2 = O(1)$, and $n^{-1}\sum \|w_i\|^2 = o(1)$. Keep the squared offsets in the two variance bounds on any convergent subsequence of their polynomial coefficients. The limiting Gaussian is now centered, $Q = R = 0$, and the longitudinal variance already consumes the full $1/4$ loss budget. Hence $\|Z^0\|_2 + \|Pv^0\|_2 \to 0$. Precisely, their liminf sum of squares plus $1/4$ is a lower bound on the sum of the expected losses by the same restriction to increasing boxes used in Section 12.2, while the exact identity and near-optimality bound that total above by $1/4 + o(1)$.

The center moduli satisfy $\||v^0| - \mathbf{1}/2\|_2 = o(1)$ from the defect bound, and $Pv^0 = o_{\ell^2}(1)$. With $e \geq 0$ this implies $\|v^0 - e^{i\phi_n}\mathbf{1}/2\|_2 = o(1)$ for $\phi_n = \arg(e^* v^0)$, by taking coordinate moduli as in the preceding localization. The identity for $m - 2v$ at the bounded center gives $\|m^0 - e^{i\phi_n}\mathbf{1}\|_2 = o(1)$. Its mean $s\sqrt{x_0} = \sqrt{s^2 - 1}$ is real positive, so by averaging and taking moduli, $\sqrt{n}|s^2 - 2| \to 0$ and $\|\ell - s\mathbf{1}\|_2 \to 0$. In particular, the longitudinal vector is $o(1)$ in Euclidean norm from the scalar optimizing longitudinal vector $c_n\mathbf{1}$, where $c_n = \sqrt{nb_n} = \sqrt{2} + O(1/n)$.

Gauge the longitudinal loading real positive coordinatewise, which is legitimate for all sufficiently large n since each complex coordinate differs from s by $o(1)$. Write the resulting loadings as the real longitudinal vector $c_n\mathbf{1} + t_n h$, where $t_n = o(1)$, h is real and $\|h\|_2 = 1$, and row transverse matrix $\sqrt{n\rho_n}\, V$, where $\rho_n = n^{-1}\sum_i \|w_i\|^2 = o(1)$ and $\|V\|_{\mathrm{F}} = 1$ when $\rho_n > 0$. Use conjugate rows if needed to match the Gram convention. Before the gauge change, $V^*\mathbf{1} = 0$. Afterward $\|V^*\mathbf{1}\| \leq \|(\mathrm{diag}(e^{i\theta_i}) - I)\mathbf{1}\| = o(1)$ because the original longitudinal phases have Euclidean norm $o(1)$ and $\|V\| \leq 1$. If $\rho_n = 0$, use the one-factor maximum Theorem 3, or the scalar/local comparison below with $\alpha = 0$.

13 Uniform differentiation in the covariance chart

Fix a real unit h and a complex n-by-k matrix V with $k \leq r - 1$, $\|V\|_{\mathrm{F}} = 1$, and $\|V^*\mathbf{1}\| \leq 1$. For t near zero and $\alpha \geq 0$ near zero let $A(t, \alpha)$ have diagonal one and off-diagonal entries from

$$n^{-1}\ell\ell^* + \alpha VV^*, \qquad \ell = c_n\mathbf{1} + th.$$

We prove $\Lambda(A(t,\alpha)) - n/4 = O(1)$ with all mixed t, α derivatives through total order three uniformly bounded on a fixed small half-rectangle. The covariance parameter α is nonnegative. We first allow both signs of its square-root amplitude.

Use the amplitude $s = \sqrt{\alpha}$ at first and allow either sign. Conditional on the longitudinal latent radial variable $\tau = T/n$, T initially exponential of mean one, and an independent transverse standard complex Gaussian $\xi \in \mathbb{C}^k$, set

$$z = V\xi, \quad L_i = \|V_{i,:}\|^2, \quad m_i = \ell_i\sqrt{\tau} + sz_i, \quad a_i = 1 - \ell_i^2/n - s^2 L_i, \quad d_i = a_i + |m_i|^2.$$

The removed longitudinal phase may be absorbed into the independent circular ξ. For small $|t|, |s|$ and large n, $a_i \geq 1/2$ since $L_i \leq 1$, $\sum_i L_i = 1$. The tilt is by $\prod_i d_i$. For scalar mean strength $b = \|\ell\|^2/n = 2 + O(n^{-1/2})$, let $d_b = 1 + b(\tau - 1/n)$. Jensen gives

$$\sum_i \log d_i - n\log d_b \leq C\big(|s|\sqrt{\tau}\|\xi\| + s^2\|\xi\|^2\big). \tag{13}$$

Indeed the linear term in $\sum_i d_i - nd_b$ is $2s\sqrt{\tau}\ \mathrm{Re}\langle \ell, z\rangle$, bounded by $C|s|\sqrt{\tau}\|\xi\|$ since $\|V^*\mathbf{1}\| \leq 1$ and $\|V^*h\| \leq 1$. The quadratic term is $s^2(\|z\|^2 - 1)$, and $d_b \geq 1/2$. On $\tau \in [1/4, 3/4]$, $\|\xi\| \leq 1$, the same log ratio is bounded below: Taylor the logarithm at d_b using $d_i, d_b \geq 1/2$, the bounded linear sum, and $\|(d_i - d_b)_i\|_2 = O(1)$. This last bound follows from $\|\ell - c_n\mathbf{1}\|_2 \leq |t|$, $\|z\| \leq \|\xi\|$ and $\|L\|_2 \leq 1$. Thus the tilted normalization compared with the scalar reference is bounded above and

below. For the upper bound, absorb the first term in (13) into $\|\xi\|^2/4 + Cs^2\tau$, choose $|s|$ small so $Cs^2\|\xi\|^2 \le \|\xi\|^2/4$, and use scalar shifted-Gamma domination for τ.

In particular the tilted law has uniform moments of every fixed degree in $X = \sqrt{n}(\tau - 1/2)$, $\sqrt{\tau}$, and $\|\xi\|$. The scalar law is the translate by $n/b - 1$ of Gamma$(n+1, 1)$ conditioned to be positive after translation. Under the same upper bound an additional fixed factor $e^{Cs^2\tau}$ changes its Gamma rate by $O(n^{-1})$, which preserves these moments and bounded normalization. The lower bound uses the positive scalar mass of $[1/4, 3/4]$ and the fixed-dimensional Gaussian ball. These facts apply uniformly in h, V.

Subtract the value at $(\tau, \xi) = (1/2, 0)$ from the log tilt before differentiating its normalized expectation. The first log derivatives in t, s are bounded by a polynomial in the variables with uniform moments. To see this, compare m_i/d_i to m_0/d_0, $m_0 = c_n\sqrt{\tau}$, $d_0 = 1 - c_n^2/n + |m_0|^2$. The comparison has bounded Euclidean norm up to a polynomial in $\sqrt{\tau}, \|\xi\|$: the differences in m_i, a_i have such norms and the rational map $m/(a + |m|^2)$ is uniformly Lipschitz for $a \ge 1/2$. Thus the error in the first derivative sum involving $\partial_t m_i = h_i\sqrt{\tau}$ or $\partial_s m_i = z_i$ is bounded by Cauchy–Schwarz. The a_i derivative arrays have bounded ℓ^1 norms. The scalar component of the s derivative is bounded since $|\sum z_i| \le \|\xi\|$. The scalar component of the t derivative is $2c_n\tau\sum h_i/d_0$. Subtracting its value at $\tau = 1/2$ bounds it by $C|X|$: the scalar fraction has bounded derivative in τ and $|\sum h_i| \le \sqrt{n}$.

Higher mixed log derivatives, through at least order six, have the same polynomial bounds. The first derivatives of d_i form arrays of bounded Euclidean norm up to such polynomials. The second derivatives form arrays of bounded ℓ^1 norm: the m_i derivatives multiply two Euclidean-bounded arrays and the second a_i derivatives use h_i^2/n or L_i. All higher derivatives of d_i vanish. Each chain-rule term for $\sum_i \log d_i$ with order at least two has either two first-derivative factors, summable by Cauchy–Schwarz, or a second-derivative factor, summable in ℓ^1. Remaining factors and d_i^{-1} have polynomial sup bounds. The values subtracted at $(1/2, 0)$ have the same bounds. Differentiation of normalized expectations to order six now preserves any uniform polynomial observable bound, with its derivatives, by the quotient rule and uniform moments. Differentiation under the integral is justified on a slightly larger fixed parameter rectangle by the same scalar/Gaussian exponential domination for each fixed n.

Let $e = \mathbf{1}/\sqrt{n}$, $P = I - ee^*$ in this local chart. For $v_i = m_i/d_i$, the Euclidean norm of Pv and all parameter derivatives through order six are bounded by polynomials in $|X|, \sqrt{\tau}, \|\xi\|$. At order zero compare v_i to m_0/d_0 and apply P. At higher orders each coordinate derivative contains at least one factor from a derivative of m_i or a_i, an array of bounded Euclidean norm, and other factors have polynomial sup bounds. The same reasoning applies to $Z_i = (|m_i|^2 - a_i)/(2\sqrt{a_i}d_i)$: at order zero the common scalar baseline has Euclidean norm at most $C(1 + |X|)$ since its numerator is affine in τ with value $O(1/n)$ at $1/2$, and its remainder has bounded Euclidean norm. The diagonal matrix D and its derivatives have polynomial operator bounds. Finally the scalar $\sum(4a_i)^{-1} - n/4$ and its derivatives are uniformly bounded by bounded ℓ^1 norms of the zero-order deficits $1 - a_i$ and the parameter derivative arrays.

It follows that for $M = \mathbb{E}[D + vv^*]$, $A_0 = e^*Me - n/4$ and $R = PMP$ are bounded with derivatives through order six. For A_0 use the exact trace-defect formula and subtract $\mathbb{E}\|Pv\|^2$. The off-diagonal block $q = PMe$ and its derivatives are $O(\sqrt{n})$, since e^*v and derivatives are bounded by $\sqrt{n}$ times a polynomial and Pv is bounded as just shown. The simple top eigenvalue $\lambda \sim n/4$ satisfies

$$\lambda = n/4 + A_0 + q^*(\lambda I - R)^{-1}q.$$

The correction and its parameter derivatives at fixed $\lambda \sim n/4$ are $O(1)$, and each extra λ-derivative gains $O(1/n)$. Differentiate the equation successively, whose implicit coefficient is $1 + q^*(\lambda I -$

$R)^{-2}q \geq 1$, to obtain uniform derivatives through order six in t, s. The eigenvalue is even in s by symmetry of ξ. For an even function f, the derivative in $\alpha = s^2$ is $\frac{1}{2}\int_0^1 f''(u\sqrt{\alpha})\,du$, extending continuously to zero. Iterating this identity with bounded derivatives and powers of u proves mixed derivative bounds in t, α whenever the weighted order $i + 2j \leq 6$ for $\partial_t^i \partial_\alpha^j$. This gives the required uniform ordinary C^3 estimates on the half-rectangle.

14 Comparison with the affine Hessian

At $(0, 0)$, the full matrix gradient vanishes by the scalar stationarity and permutation and conjugation symmetry. Thus the chart Hessian is the affine Hessian evaluated on the tangent

$$tX + \alpha Y, \qquad X_{ij} = c_n(h_i + h_j)/n, \qquad Y = VV^* - \operatorname{diag}(L_i).$$

The affine Hessian Theorem 1 has the following spectral sectors. The constant real mode and real stars have negative coefficients of order n; real and imaginary zero-row-sum sectors have negative coefficients bounded away from zero; the imaginary gradients are the gauge kernel. The tangent X is entirely in the constant plus real-star spaces, with $n\|X\|_{\mathrm{F}}^2$ bounded above and below away from zero uniformly in unit h. For Y, $\|Y\|_{\mathrm{F}} \leq 1$ and $\|Y\mathbf{1}\| \leq \|V^*\mathbf{1}\| + \|L\|_2 \leq 2$. The projections of Y onto the constant and real-star spaces and the imaginary gradient space have Frobenius norm $O(n^{-1/2})$. Indeed the projection onto constant and real-star spaces is determined by the real row sums, with row-sum to Frobenius singular values $\sqrt{n-1}$ for constants and $\sqrt{(n-2)/2}$ for stars, obtained directly from $X_{ij} = a_i + a_j$. The imaginary gradient singular value is $\sqrt{n/2}$ from $X_{ij} = i(a_i - a_j)$, $\sum a_i = 0$. Thus the weighted Hessian norm of Y is bounded.

If $\max_i L_i \leq 1/[16(r-1)]$ for $r > 1$, then $\|VV^*\|_{\mathrm{F}} \geq 1/\sqrt{r-1}$ by rank and trace, while $\|\operatorname{diag} L\|_{\mathrm{F}} \leq \sqrt{\max_i L_i}$. The projection of Y onto the two zero-row-sum sectors therefore has Frobenius norm bounded below by a positive constant depending on r. The negative Hessian on $tX + \alpha Y$ then bounds $t^2 + \alpha^2$ below up to a constant. To see this, the normal part bounds α^2, and in the weighted real-star plus constant part the norm of tX is bounded below by $c|t|$ while the norm of αY is bounded above by $C|\alpha|$. A triangle inequality controls t^2 by the weighted combined norm plus α^2. The uniform C^3 remainder therefore gives a strictly smaller objective for any sufficiently small nonzero chart displacement under this leverage condition.

For an arbitrary V, the longitudinal Hessian entry remains strictly negative. The uniform derivative bounds give the upper Taylor estimate $-ct^2 + C|t|\alpha + C\alpha^2 + C(|t| + \alpha)^3$ for the excess. Thus if $|t| \geq K\alpha$ for a sufficiently large fixed K and the neighborhood is sufficiently small, the excess is nonpositive. If $|t| < K\alpha$, use the full negative semidefiniteness of the Hessian at the origin, including the gauge kernel. The quadratic term is nonpositive, while the cubic remainder is $O(\alpha^3)$. Together these give

$$\Lambda(A(t, \alpha)) \leq \Lambda(A_{*,n}) + C\alpha^3. \tag{14}$$

If the original row cap is B and $\alpha = \rho_n \geq 16(r-1)B^2/n$, then $L_i = \|w_i\|^2/(n\rho_n) \leq 1/[16(r-1)]$, so the strict comparison applies. Otherwise (14) gives excess $O_{r,B}(n^{-3})$.

A maximizing profile exists by compactness and includes the scalar test $A_{*,n}$ for sufficiently large n. Every such sequence is near-optimal for the Gaussian-score localization, so $(t, \alpha) \to (0, 0)$ and $\|V^*\mathbf{1}\| \leq 1$ eventually. Apply either case above uniformly; a contradiction sequence handles any purported failure for infinitely many n. The scalar expansion now gives the claimed sharp $1/n$ coefficient and the $O(n^{-3})$ error. Thus any improvement over the scalar maximum lies in the small-energy, relatively high-leverage regime in Theorem 4.

15 A quantitative centered transverse model

Theorem 5. *Fix $r \geq 2$, $B < \infty$, and $\epsilon > 0$. For every n, take $w_i \in \mathbb{C}^{r-1}$ with $\sum_i w_i = 0$, $\max_i \|w_i\| \leq B$, and $\Sigma_n = n^{-1}\sum_i w_i w_i^* \preceq (2-\epsilon)I$. Put $\beta = 2 + u_n/\sqrt{n}$ with bounded real u_n. For sufficiently large n, let A be the correlation matrix with Gaussian factors $(\sqrt{\beta}, w_i)$ multiplied by $n^{-1/2}$ and diagonal residual $a_i = 1 - (\beta + \|w_i\|^2)/n$. Define real symmetric Q_n, R_n on the realification of $\mathbb{C}^{r-1}$ by*

$$\eta^{\mathsf{T}} Q_n \eta = \frac{1}{n}\sum_i (\operatorname{Im} w_i^*\eta)^2, \qquad \eta^{\mathsf{T}} R_n \eta = \frac{1}{n}\sum_i (\operatorname{Re} w_i^*\eta)^2.$$

Let $H_n = I - Q_n/2$, which is uniformly positive. Then

$$\Lambda(A) = \frac{n}{4} + \frac{1}{2} - \frac{u_n^2}{16} - \frac{1}{8}\operatorname{tr}\left[(Q_n + R_n)(H_n^{-1} - I)\right] + o(1)$$

uniformly for the given caps and bounded u_n. The transverse loss is at least $\operatorname{tr}(Q_n^2)/16$, so an order-one empirical transverse covariance forces a strict order-one deficit. If one secondary real Walsh mode has strength $\kappa < 2$, the transverse loss reduces to $\kappa^2/[8(2-\kappa)]$.

15.1 Laplace law with the common phase removed

Write the latent vector divided by $\sqrt{n}$ as $y = (y_0, \eta')$, with complex Gaussian density proportional to $\exp(-n\|y\|^2)$. Remove the common phase by taking $y_0 = \sqrt{x}\, e^{i\theta}$, $\eta' = e^{i\theta}\eta$, $x \geq 0$. The measure after integration in θ is constant times $dx\, d\eta$. Each conditional mean modulus comes from $m_i = \sqrt{\beta x} + w_i^*\eta$. Put $d_i = a_i + |m_i|^2$. The tilted density is proportional to $\exp(nf_n)$ with

$$f_n(x,\eta) = -x - \|\eta\|^2 + \frac{1}{n}\sum_i \log d_i.$$

Center at $x_0 = 1 - 1/\beta$ and $\eta = 0$. For global domination set $\bar{a} = n^{-1}\sum a_i$, $q = x + \eta^*\Sigma_n\eta/\beta$, and $\phi_n(q) = -q + \log(\bar{a} + \beta q)$. Jensen and centering give

$$f_n(x,\eta) \leq \phi_n(q) - (1 - \|\Sigma_n\|/\beta)\|\eta\|^2.$$

The scalar maximum is at $q_n = 1 - \bar{a}/\beta = x_0 + O(n^{-1})$. There, the Jensen gap at the point $(q_n, 0)$ is $O(n^{-2})$ because $a_i - \bar{a} = O(n^{-1})$ and the linear terms cancel. The change of variables $x \mapsto q$ has Jacobian one and extends to $q \geq 0$ for bounding. On a fixed interval about q_n, $\phi_n(q) \leq \phi_n(q_n) - c(q - q_n)^2$, with exponential gaps/tails on its complement and a linear bound at large q. The transverse bound has a fixed quadratic gap. On a fixed box in $\zeta = \sqrt{n}(x - x_0)$, $\xi = \sqrt{n}\eta$ the exact exponent stays within $O(1)$ of $n\phi_n(q_n)$ by the Taylor calculation below, giving the required lower normalizer of order $n^{-r+1/2}\exp(n\phi_n(q_n))$. The change from q to x costs $O(\|\eta\|^2)$, so these bounds give bounded moments of every fixed order in (ζ, ξ), uniform bounds on the scaled densities on compact sets, and exponentially small tails outside sufficiently small fixed neighborhoods of $(x_0, 0)$ after any polynomial weight in $x, \|\eta\|$. For the convergence of moments it suffices to apply the uniform higher moment bounds after truncating at a fixed scaled radius.

In scaled variables the exponent relative to $f_n(x_0, 0)$ is

$$-\frac{\zeta^2}{2} - \xi^{\mathsf{T}} H_n \xi + o(1)$$

uniformly on compact scaled sets. To check, at the center the leading common conditional mean is 1 and leading $d_i = 2$. For a perturbation $m_i = 1 + \delta m_i$, the first and second terms of $\log(1 + |m_i|^2) - \log 2$ are $\operatorname{Re} \delta m_i + \frac{1}{2}(\operatorname{Im} \delta m_i)^2$. The radial Taylor expansion of $-x + \log(1 + \beta x)$ at x_0 has second derivative -1 exactly. The cross derivative with η averages to zero by $\sum w_i = 0$. At bounded u_n, replacing the common center $\sqrt{\beta - 1}$ by 1 in the transverse Hessian costs $O(n^{-1/2})$; the possible transverse linear term from a_i variations is $O(n^{-1})$ before scaling because the order-one term cancels. These estimates and the bounded third derivatives on a fixed neighborhood prove the expansion. Thus all fixed polynomial moments are given up to $o(1)$ by independent $\zeta \sim N(0, 1)$ and real $\xi \sim N(0, \frac{1}{2} H_n^{-1})$, uniformly. This is also obtained by subsequences $Q_n \to Q$, which exist by fixed dimension and uniform bounds.

15.2 Amplitude and phase losses

The conditional cofactor identity gives $M = \mathbb{E}[D + vv^*]$, with $D_{ii} = a_i / d_i^2$ and $v_i = m_i / d_i$ up to conjugating the full matrix. The common phase has dropped out. Let $e = \mathbf{1}/\sqrt{n}$, $P = I - ee^*$, and let $L_n \xi$ be the vector with entries $\operatorname{Im} w_i^* \xi$. Expand on the central neighborhood:

$$\sqrt{n}(|m_i|^2 - a_i) = u_n + \beta\zeta + 2 \operatorname{Re} w_i^* \xi + O\big(n^{-1/2}(1 + |\zeta| + \|\xi\|)^2\big).$$

The same bound with a constant independent of i, n follows from $\max \|w_i\| \le B$ and the positivity of x_0. The exact trace identity

$$\|v\|^2 = \sum_i \frac{1}{4a_i} - \sum_i \frac{(|m_i|^2 - a_i)^2}{4 a_i d_i^2}$$

therefore gives

$$\mathbb{E}\|v\|^2 = \frac{n}{4} + \frac{\beta + \operatorname{tr} Q_n}{4} - \frac{u_n^2 + 4}{16} - \frac{1}{8} \operatorname{tr}(R_n H_n^{-1}) + o(1).$$

In this computation $\sum_i \operatorname{Re} w_i^* \xi = 0$, $\operatorname{tr} Q_n = \operatorname{tr} R_n = \operatorname{tr} \Sigma_n$, and the radial Gaussian has mean zero and variance one. All denominator replacements are controlled in mean by bounded scaled moments, and the omitted tails are exponential since $a_i \ge 1/2$ eventually and the trace defect is pointwise bounded per coordinate.

At the common center the map $m \mapsto m/(1 + |m|^2)$ has differential $z \mapsto i \operatorname{Im} z/2$ at $m = 1$. Centering therefore gives

$$Pv = \frac{i}{2\sqrt{n}} L_n \xi + o_{L^2}(1), \qquad e^* v / \sqrt{n} = \tfrac{1}{2} + o_{L^2}(1).$$

For the first formula one can subtract the constant scalar v evaluated at the perturbed x with $\eta = 0, a_i = 1$, then use uniform Taylor expansion; the residual $a_i - 1$, second-order terms, and the derivative errors from the moving radial center have total Euclidean norm $O(n^{-1/2}(1 + |\zeta| + \|\xi\|)^2)$. Consequently

$$\mathbb{E}\|Pv\|^2 = \tfrac{1}{8} \operatorname{tr}(Q_n H_n^{-1}) + o(1).$$

The diagonal Rayleigh term tends to $1/4$. The lower block PMP has $O(1)$ norm. To see that the upper eigenvalue equals the e-Rayleigh quotient up to $o(1)$, the off-diagonal block must be $o(\sqrt{n})$, not only $O(\sqrt{n})$. Its division by $\sqrt{n}$ and the last two expansions reduce it to $iL_n \mathbb{E}\xi/(4\sqrt{n}) + o(1) = o(1)$, because $\|L_n\| \le C\sqrt{n}$ and the limiting transverse Gaussian is centered. The Schur complement for the $n/4 + O(1)$ Rayleigh quotient now gives $o(1)$. Subtract the transverse rank loss from the trace, add the diagonal term, and use $2 \operatorname{tr} Q_n = \operatorname{tr}(Q_n + R_n)$. This proves the formula.

Finally $H_n^{-1} - I \succeq Q_n/2$, and $\operatorname{tr}[(Q_n + R_n)Q_n] \ge \operatorname{tr}(Q_n^2)$ by positivity. For a real one-dimensional secondary direction, Q_n and R_n have disjoint ranges and the single eigenvalue of Q_n is the empirical secondary strength, giving the claimed Walsh reduction.

16 Bounded Gram corrections of any rank

For a correlation matrix A, let $C(A)_{ij}$ be the permanent obtained by deleting row i and column j. We study $\Lambda(A) = \lambda_{\max}(C(A))/\operatorname{per} A$. Marcus and Merris [4] and Lavi [3, Section 4] give earlier spectral inequalities for this raw cofactor observable. Lemma 71 bounds its unrestricted maximum up to constant order, and Theorem 1 gives the full affine Hessian at the scalar equicorrelation optimum. Exact one-factor maximization Theorem 3 includes singular diagonal residuals. The bounded fixed-rank result Theorem 4 gives an $O(n^{-3})$ upper bound on possible excess, using a local chart normalized by the transverse trace.

Here the entries of A off diagonal are $u_i^* u_j/n$ with $\|u_i\| \le B$. The diagonal entries of A are one. There is no bound on the rank of the Gram correction. Equivalently, $A = I + G - \operatorname{diag} G$, with $G \succeq 0$ and $G_{ii} \le B^2/n$. For fixed B this is also exactly the spectral condition $A \succeq (1-B^2/n)I$. One direction follows from $G \succeq 0$ and its diagonal bound. For the reverse, take $G = A - (1 - B^2/n)I$. The sequence of sharp upper bounds in this paper concerns this bounded-diagonal Gram model.

Theorem 6. *Fix $B \ge \sqrt{2}$. Let $\gamma_{n,B}$ be the maximum of Λ over the bounded Gram correction class, and $A_{*,n}$ the positive equicorrelation optimizer with off-diagonal entry b_n. Then*

$$0 \le \gamma_{n,B} - \Lambda(A_{*,n}) = O_B(n^{-3/2}), \qquad \gamma_{n,B} = \frac{n}{4} + \frac{1}{2} + \frac{1}{4n} + O_B(n^{-3/2}).$$

Any sequence with $\Lambda \ge n/4 + 1/2 - o(1)$ admits row gauges and a unit latent direction h. With $s = \|n^{-1}\sum_i u_i\|$, $\ell_i = u_i^* h$, and $w_i = P_{h^\perp} u_i$, in these gauges

$$\sqrt{n}(s^2 - 2) \to 0, \qquad \|\ell - s\mathbf{1}\|_2 \to 0, \qquad \|(w_i^* w_j/n)_{ij}\|_{\mathrm{F}} \to 0.$$

The gauged matrices converge in Frobenius norm to equicorrelation with off-diagonal entry $2/n$. Any possible excess over the exact scalar optimum in a near-maximizing sequence occurs at transverse Frobenius norm $O_B(n^{-1/2})$ in this decomposition.

The score argument uses Gaussian integration by parts [13] with a trace-bounded weight. The quadratic comparison controls the permanent-tilted law itself by a Gaussian with an adjusted covariance. It derives the gap for the comparison quadratic form from near-optimality. General Gaussian total variation estimates such as [14] compare Gaussian laws given their covariances. The local regularity uses Gaussian rotation [15] and the heat differentiation for a Gaussian expectation [16]. Its trace cancellation is specific to the conditional cofactor vector and removes a term of size up to $\sqrt{n}$ in the local chart.

17 Rank-independent localization

17.1 Weighted score estimates

Restrict the latent space to the span of the loadings so its dimension is at most n. Use a latent Gaussian y with covariance I/n before tilting. Write $m_i = u_i^* y$, $d_i = a_i + |m_i|^2$, $a_i = 1 - \|u_i\|^2/n$, $v_i = m_i/d_i$, and $Z_i = (|m_i|^2 - a_i)/(2\sqrt{a_i} d_i)$. The Gaussian moment identity [9], followed by conditioning on y, gives the normalized cofactor matrix up to transpose as $M = \mathbb{E}[D + vv^*]$ with $D_{ii} = a_i/d_i^2$, under the tilt by $\prod_i d_i$. Let U be the matrix of rows u_i^*, choose a unit top eigenvector e of M, and write $P = I - ee^*$. Make $e \ge 0$ by row gauges. The exact deficit identity

$$\Lambda = \sum_i (4a_i)^{-1} + \mathbb{E} e^* D e - \mathbb{E}\|Z\|^2 - \mathbb{E}\|Pv\|^2$$

gives both expected losses $O(1)$. The pointwise identity $Z_i^2 = (4a_i)^{-1} - |v_i|^2$ yields $\mathbb{E}\||v| - \mathbf{1}/2\|^2 = O(1)$, since $a_i = 1 + O(1/n)$ and the squared gap to the maximal modulus is bounded by Z_i^2. Let $\xi_0 = e^* v$. Pointwise $\||v| - |\xi_0| e\| \le \|Pv\|$. Choose a realization where both losses are bounded. It gives $\||\xi_0| e - \mathbf{1}/2\| = O(1)$, and $\|e - \mathbf{1}/\sqrt{n}\| = O(n^{-1/2})$ follows by taking norms. The same inequalities in expectation give $\mathbb{E}\|v - e^{i\phi}\mathbf{1}/2\|^2 = O(1)$ with $\phi = \arg e^* v$. At a zero argument choose any phase.

For $\|y\|$ large, the Gaussian tail estimate stays uniform because its original Gamma shape is at most n. The tilted normalizer relative to the Gaussian is at least $\prod_i a_i \ge e^{-2B^2}$. The tilt is at most $(1 + B^2\|y\|^2)^n$. With R large, on $\|y\| > R$ this is at most $e^{n\|y\|^2/4}$. Absorbing polynomial factors with a further $e^{n\|y\|^2/16}$, the Gaussian Chernoff bound gives exponential decay since $\mathbb{E}_0 e^{n\|y\|^2/2} \le 2^n$. Increasing R lets the residual e^{-cnR^2} beat 2^n. On the ball, use $m_i - 2v_i = m_i(d_i - 2)/d_i$ and the Z loss to obtain

$$\mathbb{E}\|m - e^{i\phi}\mathbf{1}\|^2 = O(1). \tag{15}$$

The realified potential $f = -\|y\|^2 + n^{-1}\sum \log d_i$ has uniformly bounded real Hessian in operator norm, independently of rank: each real Hessian of a log term is dominated in absolute quadratic form by $C|u_i^* h'|^2$ for a real direction h', using $a_i \ge 1/2$. For any deterministic complex positive semidefinite K, realify it and integrate by parts to get the weighted score identity (cf. [13, Theorem 1])

$$\mathbb{E}\langle \partial_{\bar{y}} f, K \partial_{\bar{y}} f\rangle = -\frac{1}{4n}\mathbb{E}\,\mathrm{tr}(\mathcal{R}(K) D^2_{\mathbb{R}} f) \le \frac{C}{n}\,\mathrm{tr}\,K.$$

The Gaussian tails justify the boundary. Here $\partial_{\bar{y}} f = -y + U^* v/n$. Choose $K = U^* U/n$, whose trace is at most B^2, and combine with the spectral alignment of v and $\|U\| \le B\sqrt{n}$. This gives

$$\mathbb{E}\|m - e^{i\phi} U\bar{u}/2\|^2 = O(1), \qquad \bar{u} = n^{-1}\sum_i u_i.$$

For the replacement of v use the operator norm $\|UU^*/n\| \le B^2$. With (15) this proves $\|U\bar{u}/2 - \mathbf{1}\| = O(1)$, hence $s^2 = 2 + O(n^{-1/2})$, $h = \bar{u}/s$, $\|\ell - s\mathbf{1}\| = O(1)$, and $\sum w_i = 0$.

Choose $K = hh^*$ for the radial direction. It gives $\mathbb{E}|h^* y - e^{i\phi} s/2|^2 = O(1/n)$: the replacement of v costs at most $\|Uh\|/n = O(n^{-1/2})$ times its $O(1)$ Euclidean error. Take the phase section $y = e^{i\theta}(\sqrt{x}\, h + \eta)$, $\theta = \arg h^* y$, with $x = |h^* y|^2$. Write $z = (w_i^* \eta)_i$, $\beta = s^2$, $x_0 = 1 - 1/\beta$, and $\zeta = \sqrt{n}(x - x_0)$. The radial estimate and $\beta/4 - x_0 = O(1/n)$ give tightness of ζ. From (15) and the radial estimate, $\mathbb{E}\|z\|^2 = O(1)$: the section vector $e^{-i\theta} m$ is within bounded mean-square distance of $e^{i(\phi-\theta)}\mathbf{1}$; $\sqrt{n}|e^{i(\phi-\theta)} - 1|$ is bounded in mean square because the complex number $h^* y$ is close to $e^{i\phi} s/2$, with s bounded below; and $\sqrt{x}\,\ell - \mathbf{1}$ is bounded in mean square by the radial estimate and the deterministic ℓ bound. For the phase bound use $|e^{i\theta} - e^{i\phi}| \le 4|h^* y - e^{i\phi} s/2|/s$ even if $h^* y = 0$. We now have joint tightness in $|\zeta|, \|z\|_2$, without tightness in the full latent vector at scale $\sqrt{n}$.

17.2 Quadratic comparison of the projected field

Under the untilted Gaussian in the phase section, the radial variable and z are independent, and z is circular complex Gaussian with covariance $\Omega = (w_i^* w_j/n)_{ij}$. It satisfies $\Omega_{ii} \le B^2/n$, $\mathrm{tr}\,\Omega \le B^2$. Write z also as a real vector when needed, with real covariance $\Gamma = \mathcal{R}(\Omega)/2$. Take the reference law ν_n that tilts only the radial variable by the homogeneous scalar denominator $d_b(x)^n$, $d_b = 1 - \beta/n + \beta x$. Then ζ converges to $N(0,1)$ with uniform moments via the shifted Gamma formula and z is unchanged. The full law μ_n has likelihood proportional to

$$L_n(x, z) = \frac{\prod_i (a_i + |\ell_i\sqrt{x} + z_i|^2)}{d_b(x)^n}.$$

On the domain $E_K = \{|\zeta|, \|z\|_2 \le K\}$, for every fixed K, $\log L_n$ is bounded in absolute value by a constant depending on K. Indeed if $b_i = a_i + |\ell_i\sqrt{x} + z_i|^2$, its difference array from d_b has bounded Euclidean norm there. Its sum is bounded since $\sum(\ell_i - s) = 0$, $\sum z_i = 0$, $\|\ell - s\mathbf{1}\|_2 = O(1)$, and the diagonal residual difference has bounded sum. Taylor expansion of the logarithms at d_b with $b_i, d_b \ge 1/2$ proves the assertion. Both measures have positive probability on some fixed E_K, by their respective tightness bounds. Consequently the likelihood normalizer $\int L_n \, d\nu_n$ is bounded above and below by positive constants. The same holds for $L_n(x_0, 0)$.

The reference Gaussian satisfies $\Pr(\max_i |z_i| > n^{-1/4}) \to 0$ by its diagonal bound and a union bound. The same holds under μ_n: first restrict to a high-probability E_K, on which the normalized likelihood is bounded, and then let $K \to \infty$. On $E_K \cap \{\|z\|_\infty \le n^{-1/4}\}$ there is the uniform expansion

$$\log[L_n(x, z)/L_n(x_0, 0)] = \tfrac{1}{2}\|\operatorname{Im} z\|_2^2 + c_n^{\mathsf{T}} z_{\mathbb{R}} + o(1), \qquad \|c_n\|_2 = O(1). \tag{16}$$

For the radial part use the log ratio at $z = 0$, whose derivative near x_0 is $O(1)$: cancel terms linear in $\ell_i - s$ and bound the diagonal residual discrepancy in ℓ^1. Over $|x - x_0| \le K/\sqrt{n}$ this costs $o(1)$. For the linear transverse term, subtract the scalar coefficient of z_i, whose sum vanishes exactly. Its residual coefficient vector at x_0 has bounded Euclidean norm by the ℓ discrepancy and the $O(n^{-1})$ diagonal residual deviations. Changing x costs $o(1)$ in Euclidean norm after this same scalar subtraction: derivatives of the residual coefficient are bounded by $C(|\ell_i - s| + 1/n)$. The leading quadratic term at $m = 1, a = 1$ is $(\operatorname{Im} z_i)^2/2$. Errors from diagonal and loading differences are at most $C\sum_i(|\ell_i - s| + n^{-1/2})|z_i|^2 = o(1)$ on this set by Cauchy–Schwarz and its sup bound. Third-order Taylor errors are bounded by $C\|z\|_\infty\|z\|_2^2 = o(1)$.

Let P_I be orthogonal projection to the imaginary coordinates of the realification, and $J_n = \Gamma^{1/2} P_I \Gamma^{1/2}$. In the reference representation $z_{\mathbb{R}} = \Gamma^{1/2} g$ with real standard Gaussian g, the comparison tilt in (16) is $\exp(\frac{1}{2} g^{\mathsf{T}} J_n g + b_n^{\mathsf{T}} g)$ with $b_n = \Gamma^{1/2} c_n$, uniformly bounded in norm. We claim a strict uniform gap $\|J_n\| \le 1 - \delta$ holds eventually along a near-maximizing sequence. Otherwise take a subsequence with leading eigenvalue $\lambda_n \ge 1 - o(1)$, and split $g = qv + g'$ into the unit leading mode v and its independent orthogonal Gaussian. Restrict $|q| \le L$, $|\zeta| \le K_0$, $\|\Gamma^{1/2} g'\| \le K_0$ and $\|\Gamma^{1/2} g'\|_\infty \le n^{-1/4}/2$ in the complex-coordinate norm. Fixed large K_0 gives uniform positive Gaussian mass for the latter two restrictions by the trace and diagonal bounds. The event in g' is even, so integrating its exponential quadratic plus linear tilt costs at least its probability by the even average and positivity of the quadratic. The radial event has positive mass. For fixed L the full event is within the domain of (16) with K depending on L, since each coordinate of $\Gamma^{1/2} v$ is $O(n^{-1/2})$ and its Euclidean norm is bounded. The q integral is bounded below by a fixed multiple of

$$\int_{-L}^{L} \exp\bigl[-(1 - \lambda_n) q^2/2\bigr] \, dq,$$

using even averaging in q as well. As $n \to \infty$ this has liminf at least $2L$. The normalized likelihood integral was bounded above, so letting $L \to \infty$ gives a contradiction.

Let $\nu_n^{(2)}$ be the reference law with the normalized quadratic tilt from (16). It keeps the radial law and makes $z_{\mathbb{R}}$ a Gaussian with covariance

$$\Gamma_2 = \Gamma^{1/2}(I - J_n)^{-1}\Gamma^{1/2}$$

and mean $\mu_z = \Gamma^{1/2}(I - J_n)^{-1} b_n$. Its normalization is uniformly bounded above and below: use the determinant formula, the gap, $\operatorname{tr} J_n \le B^2$, and $\|b_n\| \le C$. It has uniform moments of $\|z\|$ by its trace and mean bounds. Its coordinate variances are $O(n^{-1})$ by $\Gamma_2 \preceq \delta^{-1}\Gamma$, and its coordinate means are $O(n^{-1/2})$ by the diagonal bound of Γ. Thus for large K the high-probability sets of (16)

have high probability under both μ_n and $\nu_n^{(2)}$. There the unnormalized likelihood ratios agree up to $1+o(1)$ after using the factor $L_n(x_0,0)$. This proves total variation convergence of the two laws: first bound the normalization ratio by the ratios of the restricted masses, then take $n\to\infty$ at fixed K followed by $K\to\infty$.

17.3 The cofactor deficit and Frobenius convergence

On the sets of (16), uniform Taylor expansion in the section gives

$$Z = Z^0 + \frac{1}{2}(\zeta\,\mathbf{1}/\sqrt{n} + \operatorname{Re} z) + o_{\ell^2}(1), \qquad Pv = Pv^0 + \frac{i}{2}\operatorname{Im} z + o_{\ell^2}(1).$$

Indeed the second residuals in each coordinate are bounded by $C(|\zeta|/\sqrt{n}+|z_i|)^2$, which tend to zero in Euclidean norm on those sets. The first differential coefficients may be replaced by those at $m=1, a=1$: the radial errors are bounded in Euclidean norm before multiplying $\zeta/\sqrt{n}$, and the transverse coefficient errors from $|\ell_i - s|$ cost at most $C\|\ell - s\mathbf{1}\|\|z\|_\infty$. Also $\sum z_i = 0$, so replacing the projection onto $e^\perp$ by the projection onto the constant vector's complement costs $o(1)$. Both offsets Z^0, Pv^0 are bounded in Euclidean norm.

The two nonnegative squared losses under μ_n, restricted to the high-probability sets, can now be compared with their polynomial approximations under $\nu_n^{(2)}$ by total variation. On each fixed E_K the polynomials are bounded. Outside the sets their comparison-law expectations are uniformly small as $K\to\infty$ by uniform higher moments, while the true losses are nonnegative. Therefore their sum is bounded below by

$$\tfrac{1}{4} + \tfrac{1}{4}\operatorname{tr}\Gamma_2 + \|Z^0 + \tfrac{1}{2}\operatorname{Re}\mu_z\|^2 + \|Pv^0 + \tfrac{i}{2}\operatorname{Im}\mu_z\|^2 - o(1). \tag{17}$$

The negligible radial mean is absorbed into $o(1)$, and its variance tends to one. The radial and transverse comparison variables are independent. The diagonal Rayleigh contribution tends to $1/4$ by (15) and the uniform gauge comparison of e. The trace term in the deficit identity is $n/4 + 1/2 + \operatorname{tr}\Omega/4 + o(1)$, as seen by its first-order expansion in the diagonal residuals. Note $\operatorname{tr}\Gamma = \operatorname{tr}\Omega$.

The excess of comparison covariance is nonnegative and controls the Frobenius norm:

$$\operatorname{tr}(\Gamma_2 - \Gamma) \geq \operatorname{tr}(\Gamma J_n) = \operatorname{tr}(P_I\Gamma^2) = \tfrac{1}{4}\|\Omega\|_{\mathrm{F}}^2.$$

The first inequality uses $(I-J_n)^{-1} - I \succeq J_n$; the last equality is the circular realification identity $\Gamma = \mathcal{R}(\Omega)/2$, whose imaginary diagonal block in the square has half the full realified trace. Thus the near-maximal loss budget forces $\|\Omega\|_{\mathrm{F}}\to 0$ and both shifted offsets in (17) tend to zero. Moreover $\|\mu_z\| \leq \delta^{-1}\|\Gamma\|\|c_n\| = o(1)$, so the offsets themselves tend to zero. Taking moduli and then aligning v^0 to a common phase via Pv^0 gives m^0 converging in Euclidean norm to a unimodular constant vector. In detail, $\||v^0| - \mathbf{1}/2\| \to 0$ by the defect identity, and $\|Pv^0\|\to 0$ then gives $\|v^0 - e^{i\phi_n}\mathbf{1}/2\|\to 0$ by the same modulus comparison used above. The bounded center satisfies $\|m^0 - 2v^0\|\to 0$ by the identity preceding (15). Its mean is $s\sqrt{x_0} = \sqrt{s^2-1}$, real positive. Thus $\sqrt{n}(s^2-2)\to 0$ and $\|\ell - s\mathbf{1}\|\to 0$. The longitudinal Gram, the vanishing transverse Frobenius term, and diagonal subtraction give the convergence to the equicorrelation orbit. The transverse trace may persist: for $0<\tau_*<B^2-2$, take $\Omega = \tau_*(I - \mathbf{1}\mathbf{1}^*/n)/n$ and a common longitudinal loading of squared size $nb_n + \tau_*/n$, where b_n is the scalar optimum off diagonal. The resulting off-diagonals are exactly b_n, the row cap holds for large n, and $\operatorname{tr}\Omega\to\tau_*$ while $\|\Omega\|_{\mathrm{F}}\to 0$.

18 The Frobenius covariance chart

By Section 17, near-maximizers admit gauged longitudinal loadings $\ell = c_n\mathbf{1} + th$, where ℓ is real positive, $c_n = \sqrt{nb_n} = \sqrt{2} + O(1/n)$ is at the scalar optimum, h is real unit, and $t = o(1)$. The transverse Gram perturbation is αW, $\alpha = o(1)$ its Frobenius norm, $W = VV^* \succeq 0$, $\|W\|_{\mathrm{F}} = 1$. When $\alpha = 0$, use $W = 0$ in the chart. The strictly negative longitudinal Hessian below gives the scalar comparison. The original transverse directions are centered before the longitudinal gauge. After making ℓ_i positive, the gauge phases are $o(1)$ in Euclidean norm by the strong longitudinal convergence, so $\|V^*\mathbf{1}\| = o(1)$ from $\|V\| \le 1$.

We work uniformly for all real unit h, all such W with $\|V^*\mathbf{1}\| \le 1$, small real t, s, and $\alpha = s^2$. Use scalar radial variable $\tau = T/n$ with T originally standard exponential, and independent circular complex Gaussian field z with covariance W. Put

$$L_i = W_{ii}, \quad m_i = (c_n + th_i)\sqrt{\tau} + sz_i, \quad a_i = 1 - (c_n + th_i)^2/n - s^2 L_i, \quad d_i = a_i + |m_i|^2.$$

Then $a_i \ge 1/2$, $\sum_i L_i^2 \le 1$, $\sum_i L_i \le \sqrt{n}$. The normalized cofactor matrix is $M = \mathbb{E}[D + vv^*]$ under the tilt by $\prod_i d_i$, with $D_{ii} = a_i/d_i^2$, $v_i = m_i/d_i$. We prove all derivatives through order six in (t, s) of $\Lambda - n/4$ are uniformly bounded. Evenness in s then gives ordinary C^3 uniformity in (t, α) for $\alpha \ge 0$, by the even differentiation formula proved below.

18.1 Gaussian moment estimates

We use Gaussian rotation and its iteration (cf. [15]). Let $z_{\mathbb{R}} = \Gamma^{1/2} g$, g real standard Gaussian, $\Gamma = \mathcal{R}(W)/2$. For a sum $F = \sum_i f_i(z_i)$ with real or complex local functions, the Gaussian rotation estimate is

$$\|F - \mathbb{E}F\|_p \le C\sqrt{p}\,\big\|\|\nabla_g F\|\big\|_p \quad (p \ge 2).$$

Indeed use an independent copy g', Jensen, and the path $g_\theta = g\cos\theta + g'\sin\theta$, $0 \le \theta \le \pi/2$; its derivative is independent standard Gaussian for each θ. Minkowski and the conditional Gaussian moment give the estimate. The same proof for a Euclidean-vector-valued function bounds the right side by $C\sqrt{p}$ times the L^p Frobenius norm of its Jacobian, since a matrix applied to a standard Gaussian has L^p Euclidean norm at most $C\sqrt{p}$ times its Frobenius norm. Applying this to $\nabla_g F$ gives

$$\|F - \mathbb{E}F\|_p \le C\sqrt{p}\,\|\mathbb{E}\nabla_g F\| + Cp\,\big\|\|D_g^2 F\|_{\mathrm{F}}\big\|_p. \tag{18}$$

Apply (18) to the additive sums of $\log d_i$ and v_i, and their t, s derivatives through order six. Here are uniform bounds on both terms, for fixed p, polynomial in $1 + \sqrt{\tau}$. The Hessian is $\Gamma^{1/2}\,\mathrm{diag}_{\mathrm{block}}(D_i')\Gamma^{1/2}$, where each 2-by-2 block D_i' is the real-coordinate local Hessian. Each block has a uniform L^p bound of the required polynomial size because $d_i \ge 1/2$, $L_i \le 1$, $|h_i| \le 1$, and the local expressions and their finitely many derivatives are rational with polynomial numerators. For real components the squared Frobenius norm of the Hessian is at most

$$\sum_{ij} \|D_i'\|\,\|D_j'\|\,\|\Gamma_{ij}\|_{\mathrm{F}}^2,$$

by the block trace formula. Apply Minkowski and Hölder using $\sum_{ij}\|\Gamma_{ij}\|_{\mathrm{F}}^2 = \|\Gamma\|_{\mathrm{F}}^2 \le 1$. This avoids a maximum over coordinate magnitudes.

For the mean gradient write it as $\Gamma^{1/2}$ times the array of the expected local gradients in z_i. Compare that array to the same deterministic gradient at the scalar baseline $h_i = 0, L_i = 0, z_i = 0$. The local means differ by at most $C_p(1 + \sqrt{\tau})^k(|h_i| + L_i)$. For the deterministic parameters use

smoothness and $\ell_i - c_n = t h_i$; for the Gaussian evaluation Taylor expand the local gradient at $z_i = 0$, use its mean-zero linear term and the expected quadratic remainder $O(L_i)$, with polynomially bounded higher factors. This remains valid for the parameter derivatives: a positive t-order contains factors h_i, and s-derivatives use the same local Gaussian variable and L_i through $a_i = 1 - \ell_i^2/n - s^2 L_i$. The difference array has bounded Euclidean norm by $\|h\|_2 = 1, \|L\|_2 \le 1$. Multiplying a constant two-dimensional array by $\Gamma^{1/2}$ is bounded because $\|V^* \mathbf{1}\| \le 1$. This proves the asserted polynomial moment bounds. All sufficiently many local derivatives used in this comparison exist uniformly since $a_i \ge 1/2$.

For the tilt itself we also need exponential, not only fixed polynomial, moments. For $F = \sum \log d_i$, its Hessian in g has operator norm at most Cs^2 uniformly, because the second derivative of $\log(a + |m|^2)$ in m is uniformly bounded. The variance estimate for its gradient from (18)'s vector predecessor implies $(\mathbb{E}\|\nabla_g F\|^2)^{1/2} \le C(1 + \sqrt{\tau})$. In this bound the expected scalar local gradient itself is uniformly bounded by Cs; its deviation from the constant for h_i, z_i, L_i can be bounded in Euclidean norm by $C(1 + \sqrt{\tau})$ using the Lipschitz first derivative and Gaussian second-order Taylor with uniformly bounded derivatives. The Gaussian rotation estimate for the scalar gradient norm, which is Cs^2-Lipschitz, gives

$$\|F - \mathbb{E}F\|_p \le C(1 + \sqrt{\tau})\sqrt{p} + Cs^2 p, \qquad p \ge 2. \tag{19}$$

For a prescribed fixed $q \ge 1$, this implies $\mathbb{E}_g e^{q(F - \mathbb{E}_g F)} \le C_q e^{C_q \tau}$ on a small fixed amplitude interval. To see this, optimizing Markov over integers $p \ge 2$ gives a tail bounded by a constant times the sum of a Gaussian tail $\exp[-cx^2/(1+\tau)]$ and an exponential tail $\exp[-cx/s^2]$, allowing a larger constant for small x. Integrate e^{qx} against those tails and choose $s^2 < c/(2q)$. This proof of (19) is independent of rank.

18.2 Tilt normalization and radial anchoring

Use a fixed scalar reference at c_n, with $a_* = 1 - c_n^2/n$, $m_* = c_n\sqrt{\tau}$, $d_* = a_* + |m_*|^2$; denote by ν the reference radial tilted law times the original Gaussian field. Let $F_{\mathrm{rel}} = \sum \log d_i - n \log d_*$. Anchor it at its Gaussian expectation at $\tau_0 = 1/2$. This preserves the normalized tilt. For $\tau \in [1/4, 3/4]$, the difference of its Gaussian mean from the anchor and its parameter derivatives through six are $O(1 + |X|)$, $X = \sqrt{n}(\tau - 1/2)$. For other τ they are bounded by a polynomial in $1 + |X|, 1 + \sqrt{\tau}$. Here is a direct verification of the mean estimates. For a smooth local observable $f(a, m)$, define

$$\mathcal{T}_q f(a, m) = \mathbb{E}_{G \sim CN(0,1)} f(a - q, m + \sqrt{q}\, G), \qquad \mathcal{L} = -\partial_a + \tfrac{1}{4}\Delta_m.$$

For $q = s^2 L_i$ and $a = 1 - \ell_i^2/n$, the expectation of the local observable is $\mathcal{T}_q f(a, \ell_i\sqrt{\tau})$. On the chart $a - q \ge 1/2$, $\partial_q^j \mathcal{T}_q f = \mathcal{T}_q \mathcal{L}^j f$ by Gaussian heat differentiation, including at zero (cf. [16]). For $f = \log(a + |m|^2)$, the difference from $q = 0$ and all its needed amplitude derivatives are bounded by CL_i times a polynomial in $1 + \sqrt{\tau}$; it follows from $q = s^2 L_i$, the fundamental theorem and the local rational derivatives. The same bound applies after a bounded number of longitudinal derivatives and one radial derivative on the central interval. The baseline longitudinal scalar difference from c_n costs $C|h_i|$ times such a polynomial, and this statement also covers its positive-order derivatives. Thus the total raw mean differences and their derivatives are bounded by $C\sqrt{n}$ on the central interval with radial derivative, and by $C\sqrt{n}$ times a polynomial globally. This proves the anchored estimates by the mean value theorem centrally, and by $|X| \ge \sqrt{n}/4$ on the complement.

For the undifferentiated mean, a sharper global bound $C\sqrt{n}(1 + \sqrt{\tau})$ suffices for exponential integrability outside the central interval. The log function has uniformly bounded first derivatives in m, a and $\mathcal{L}\log(a + |m|^2)$ is uniformly bounded when $a \ge 1/2$. This proves the sharper bound by

the same rowwise comparison. The centered conditional exponential moment from (19), the central bound $C|X|$, and this global bound show the anchored likelihood has bounded $L^q(\nu)$ for a prescribed $q > 1$, on a fixed small amplitude interval. Indeed, for the scalar τ law $T + (n/c_n^2 - 1)$ has the Gamma$(n+1, 1)$ law conditioned to exceed $n/c_n^2 - 1$. It has bounded exponential moments in any fixed multiple of $|X|$ for sufficiently large n; outside the central interval a factor $\exp[C_q\sqrt{n}(1+\sqrt{\tau})+ C_q\tau]$ is absorbed by its exponential tails. The tilted normalization is bounded below by Jensen on $|X| \le 1$, using the Gaussian-mean anchored bound there. The normalized log likelihood derivatives have bounded moments after using (18) and the anchored mean estimates, with polynomial radial factors. Hence the quotient differentiation rule preserves uniform bounds on expectation derivatives of observables with such L^p estimates. Differentiation under the integral for each n follows directly by differentiating the finite polynomial likelihood and the rational observables with $d_i \ge 1/2$ on a slightly larger parameter rectangle.

18.3 Trace compensation in the cofactor Rayleigh quotient

Put $v_* = m_*/d_*$ and $S = \sum_i (v_i - v_*)$. Its centered conditional Gaussian fluctuations, through six parameter derivatives, have already been bounded by (18). It also has rank-independent conditionally expected size and derivatives polynomial in $1 + |X|, 1 + \sqrt{\tau}$. For the mean estimate put $a_{0,i} = 1 - \ell_i^2/n$. The scalar baseline difference $\sum_i [v(a_{0,i}, \ell_i\sqrt{\tau}) - v_*]$ has uniform bound of this size: Taylor in ℓ_i about c_n, using the first derivative

$$\frac{d}{d\ell}\frac{\ell\sqrt{\tau}}{1+\ell^2(\tau-1/n)} = \frac{\sqrt{\tau}\,[1-\ell^2(\tau-1/n)]}{[1+\ell^2(\tau-1/n)]^2}.$$

At c_n this is $O((1+\sqrt{\tau})(|\tau - 1/2| + 1/n))$. The difference at order one and the base linear term sum with $\|h\|_1 \le \sqrt{n}$; second and higher longitudinal terms sum with $\sum_i h_i^2 = 1$.

For the covariance difference, use $\mathcal{T}_q v = v + q\mathcal{L}v + O(q^2)$ and the exact cancellation

$$\mathcal{L}v(a, m) = \frac{m(|m|^2 - a)}{(a + |m|^2)^3}.$$

Indeed $\frac{1}{4}\Delta_m(m/d) = -2am/d^3$ whereas $-\partial_a(m/d) = m/d^2$. The coefficient vanishes at $a = m = 1$. Thus summing $q\mathcal{L}v$ costs at most a polynomial times

$$\sum_i L_i \left(|h_i| + |\tau - 1/2| + 1/n\right) \le 1 + |X| + n^{-1/2}.$$

Positive longitudinal derivatives of this term use $\sum_i L_i|h_i| \le 1$, and amplitude derivatives retain the compensated coefficient. The $O(q^2)$ remainder and its mixed t, s derivatives sum with $\sum_i L_i^2 \le 1$. More explicitly, write it as $q^2 \int_0^1 (1-\vartheta)\mathcal{T}_{\vartheta q}\mathcal{L}^2 v\, d\vartheta$. Differentiating $q = s^2 L_i$ and this smooth expression at fixed t, s leaves at least L_i^2 in every term, with bounded heat derivatives. This proves all the required estimates for S.

Let $e = \mathbf{1}/\sqrt{n}$, $P = I - ee^*$. For any deterministic unit b the scalar b^*Pv and its parameter derivatives through six have bounded conditional L^p norms polynomial in $1 + \sqrt{\tau}$, uniformly in b. Indeed, b^*P is contracted and annihilates the constant, so subtract v_*. Taylor at $z_i = 0$ with a_i fixed. The array of deterministic baseline differences has bounded Euclidean norm by $\|h\|, \|L\|$. The terms linear in $z_i, \overline{z}_i$ are tested against bounded coefficient Euclidean vectors, so their variance is bounded since $\|W\| \le 1$. The quadratic remainders tested against b have bounded L^p norm by $\sum_i |b_i'|L_i \le 1$ and the uniform local Gaussian moments. The same reasoning applies to all parameter derivatives: take the z Taylor expansion of the local differentiated function, whose coefficients have uniform

polynomial bounds and whose diagonal deterministic differences contain either the removed scalar baseline or a factor h_i or L_i.

The exact Rayleigh expression is

$$e^* M e = \mathbb{E}\left[n|v_*|^2 + 2\,\mathrm{Re}(\overline{v}_* S) + |S|^2/n + n^{-1}\sum_i D_{ii} \right].$$

The baseline $n|v_*|^2 - n/4$ is $O(1 + X^2)$ from its exact scalar defect, and the terms involving S are bounded with all expectation derivatives through six. For D, arbitrary unit-vector tests give uniform L^p bounds of operator bilinear forms by the local rational estimates and Cauchy–Schwarz for the deterministic test vectors. For $R = PMP$, use the uniform estimates for scalar tests of Pv twice and Hölder. For $q = PMe$, use a unit test against Pv and $e^*v = \sqrt{n}v_* + S/\sqrt{n}$. Consequently R is $O(1)$ and q is $O(\sqrt{n})$, through six parameter derivatives in operator or Euclidean norm respectively.

The Schur complement expresses the simple top eigenvalue as $e^*Me + q^*(\lambda I - R)^{-1}q$. The correction $q^*(\lambda I - R)^{-1}q$ and its fixed-λ parameter derivatives are bounded, and each λ-derivative gives a further $O(n^{-1})$. Successive implicit derivatives are bounded through order six since the coefficient of the highest derivative of λ is $1 + q^*(\lambda I - R)^{-2}q \geq 1$. The eigenvalue is even in s since the centered field is symmetric. If f is smooth and even, $\partial_\alpha f(\sqrt{\alpha}) = \frac{1}{2}\int_0^1 f''(a\sqrt{\alpha})\,da$. Iterating this formula adds bounded powers of a, so six mixed amplitude derivatives yield three mixed covariance derivatives including at zero. In particular the required C^3 covariance estimates hold with W normalized in Frobenius norm, independently of rank.

19 Affine Hessian comparison

The tangents at the scalar optimum are the real constant/star matrix $X_{ij} = c_n(h_i + h_j)/n$ and $Y = W - \mathrm{diag}(L_i)$. The affine Hessian Theorem 1 is negative semidefinite with imaginary gauge kernel. Its constant and star sectors have negative coefficients of order n, and the other transverse sectors have negative coefficients of order one. Since $\|Y\mathbf{1}\| \leq \|VV^*\mathbf{1}\| + \|L\| \leq 2$, the projections on the constant, real-star, and imaginary-gradient spaces are $O(n^{-1/2})$ in Frobenius norm. Indeed these projections are determined by the real and imaginary row sums. The row-sum map multiplies Frobenius norm by $\sqrt{n-1}$ on constant real matrices, $\sqrt{(n-2)/2}$ on real stars, and $\sqrt{n/2}$ on imaginary gradients, as is seen by summing the explicit matrices in these sectors. Therefore the norm of Y with the constant and real-star squared weights multiplied by n, and the ordinary weights on the zero-row-sum sectors, stays bounded. Also X lies in the constant and real-star spaces, with $n\|X\|_{\mathrm{F}}^2$ bounded above and below by positive constants. If $\|\operatorname{diag} W\|_{\mathrm{F}} \leq 1/4$, then the normal part of Y is bounded below, since $\|W\|_{\mathrm{F}} = 1$. The negative of the pulled-back Hessian controls $t^2 + \alpha^2$. First its normal term controls α^2, since X has zero normal component. The weighted constant and star terms then control t^2 by the triangle inequality and the bound on the weighted norm of Y. The uniform C^3 estimate makes every sufficiently small nonzero such displacement lose. Without this bound use the strong longitudinal curvature to exclude $|t| \geq C\alpha$ locally and negative semidefiniteness to give possible gain $O(\alpha^3)$ on the remaining $|t| < C\alpha$, by the following Taylor estimate. The full first derivative vanishes by permutation and conjugation symmetry and scalar stationarity. If $|t|$ is larger than a sufficiently large constant times α, the negative longitudinal quadratic dominates the cross term, the α^2 term, and the cubic remainder. Otherwise the entire quadratic is nonpositive and the remainder is $O(\alpha^3)$. By the original row cap, the diagonal of the unnormalized transverse covariance has Frobenius norm at most $B^2/\sqrt{n}$. Therefore a possible gain can only occur for $\alpha \leq 4B^2/\sqrt{n}$, and is at most $O_B(n^{-3/2})$.

The feasible class is compact by using latent dimension n. The exact scalar optimum is feasible for sufficiently large n, also at $B = \sqrt{2}$, since $nb_n = 2 - 2/(n+1) + O(ne^{-cn}) < 2$ by Proposition 2. Thus its maximizing sequences satisfy the rank-independent localization and enter this chart, with its uniform neighborhood and derivative constants. The scalar expansion Proposition 2 completes the proof.

The argument also interpolates the upper error as a function of rank. If the transverse correction has rank at most $k \geq 1$, then $\operatorname{tr} W \leq \sqrt{k}$ and

$$\| \operatorname{diag} W \|_{\mathrm{F}}^2 \leq (\max_i L_i) \operatorname{tr} W \leq B^2 \sqrt{k}/(n\alpha).$$

Hence possible excess requires $\alpha \leq 16B^2\sqrt{k}/n$ as well as $\alpha \leq 4B^2/\sqrt{n}$. Its upper bound is $O_B(\min\{\sqrt{k}/n, n^{-1/2}\}^3)$ uniformly in k. The fixed-rank $O(n^{-3})$ bound follows, while Theorem 4 also gives an explicit constant-order formula for centered fixed-gap arrays.

20 Exact maximizers under a spectral floor

For a correlation matrix A, let $C(A)_{ij}$ be the permanent with row i and column j deleted, and let $\Lambda(A) = \lambda_{\max}(C(A))/\operatorname{per} A$. This is the raw cofactor observable of Marcus and Merris [4] and Lavi [3, Section 4]. The all-correlation bounds of Lemma 71 agree through constant order. Theorem 1 determines the full affine Hessian at the positive scalar optimizer $A_{*,n} = (1-b_n)I + b_n J$, with $b_n = 2/(n+1) + O(e^{-cn})$. We use that Hessian and the rank-independent localization from Theorem 6. The latter determines the $1/n$ coefficient under the spectral floor and gives an $O(n^{-3/2})$ upper bound on possible excess over $A_{*,n}$. Here the comparison is exact.

Theorem 7. *There is an absolute $\delta > 0$ such that for all sufficiently large n, any correlation matrix A in the neighborhood*

$$\|A - A_{*,n}\|_{\mathrm{F}} + \|(A - A_{*,n})\mathbf{1}\|_2 < \delta \tag{20}$$

has $\Lambda(A) \leq \Lambda(A_{*,n})$, with equality only on the diagonal-unitary orbit of $A_{*,n}$. The same holds if the neighborhood condition applies after a diagonal-unitary conjugation. This yields the following exact global comparisons in sufficiently large orders.

1. *For each fixed $B \geq \sqrt{2}$, the unique maximizers modulo diagonal phases among correlation matrices $A \succeq (1 - B^2/n)I$ are $A_{*,n}$ and its diagonal-unitary conjugates.*

2. *Fix $B < \infty$, $M > 0$, and $0 < \epsilon \leq 2$. Consider positive semidefinite matrices*

$$A = \operatorname{diag}(1 - \beta/n - K_{ii}) + \beta J/n + K, \qquad \beta = 2 + u/\sqrt{n}, \quad |u| \leq M,$$

where $K = K^*$, $K\mathbf{1} = 0$, $\|K\| \leq 2 - \epsilon$, and $(|K|)_{ii} \leq B^2/n$, with $|K| = (K^2)^{1/2}$. The unique maximizing matrix in this centered class is $A_{*,n}$.

In both classes the maximum is the exact scalar value, with asymptotic expansion $(n+1)^2/(4n) + O(ne^{-cn})$.

The signed class includes a fixed delocalized negative transverse eigenvalue, giving correlation matrices a fixed amount below the spectral floor in part 1. For example take $K = -avv^*$ with $0 < a < 1$, v a centered real unit vector with entries $O(n^{-1/2})$, and fixed bounds admitting it. The Rayleigh quotient on v is $1 - a + O(n^{-1})$, while the matrix stays positive definite for large n.

The complex Gaussian moment method for permanents is classical [17]. We use it with different linear maps in the two paired systems. A smooth bound on the absolute product gives a locally zero-free normalized permanent. At the cofactor level the unnormalized polynomial identities handle the regions where a removed denominator is small. The resulting uniform derivatives let us use the affine Hessian in a neighborhood whose size is independent of dimension.

21 Uniform holomorphic derivatives

Let $c_n = \sqrt{nb_n}$, where b_n is the optimizing scalar off-diagonal entry, and take a real unit vector h. Let $Y = Y^*$ satisfy $\|Y\|_{\mathrm{F}} \le 1$ and $\||Y|^{1/2}\mathbf{1}\| \le 1$. For complex t, s in a small fixed polydisk, let $\ell_i = c_n + th_i$ and define a unit-diagonal matrix $A(t,s)$ with off-diagonals from $\ell\ell^{\mathsf{T}}/n + s^2 Y$. There is no conjugation on t in $\ell\ell^{\mathsf{T}}$, so the entries are holomorphic. We prove the large eigenvalue branch of the normalized transposed cofactor matrix is holomorphic in a fixed smaller polydisk, with $\lambda - n/4 = O(1)$ there uniformly in n, h, Y. For real t, s it is the simple top eigenvalue. Cauchy estimates and evenness then give uniform C^3 in $(t, \alpha = s^2)$ for real $\alpha \ge 0$.

Use paired fields $z_L = U\sqrt{\kappa}\,\eta$, $z_R = \bar{U}\sqrt{\kappa}\,\bar{\eta}$ for a spectral representation of Y, explicitly as follows. For $Y = U\,\mathrm{diag}(\kappa_j)U^*$ on the nonzero spectrum, choose one square root of each κ_j, imaginary for negative eigenvalues, and let η have independent standard circular complex coordinates. Put $z_L = U\,\mathrm{diag}(\sqrt{\kappa_j})\eta$ and $z_R = \overline{U}\,\mathrm{diag}(\sqrt{\kappa_j})\overline{\eta}$. Their bilinear cross covariance is $\mathbb{E} z_L z_R^{\mathsf{T}} = Y$. Their marginal variances are $L_i = (|Y|)_{ii}$, so $\sum L_i^2 \le 1$ and $\sum L_i \le \sqrt{n}$. Their combined real Gaussian field has covariance with operator and Frobenius norms bounded by absolute constants: represent each field via realification from the same Gaussian η, and use the singular values of the two maps, both the square roots of $|\kappa_j|/2$ repeated twice. In fact the joint covariance has Frobenius norm at most the sum of the Frobenius norms of the two covariances, by factoring through the common latent space. The variance of each of $\sum z_{L,i}, \sum z_{R,i}$ is $\mathbf{1}^*|Y|\mathbf{1} \le 1$. Write $\tau = T/n$, $X = \sqrt{n}(\tau - 1/2)$ and

$$m_i = \ell_i\sqrt{\tau}, \qquad a_{0,i} = 1 - \ell_i^2/n, \qquad d_{0,i} = a_{0,i} + m_i^2,$$
$$a_i = a_{0,i} - s^2 Y_{ii}, \qquad d_i = a_i + (m_i + s z_{L,i})(m_i + s z_{R,i}).$$

The paired circular Wick identity [17, Theorem A.4(c), Eq. (A.30)] gives the permanent as the expectation of $\prod_i d_i$. Add independent diagonal Gaussian variables with identical square roots of a_i on both sides, and use the longitudinal Gaussian of radial coordinate τ and scale ℓ_i. The conditional left-right covariance of the complete systems gives the entries of $A(t,s)$. The diagonal cofactor omits d_i, and the transposed off-diagonal cofactor indexed by i, j omits d_i, d_j and inserts $m_{L,i}m_{R,j}$, where $m_{L/R,i} = m_i + s z_{L/R,i}$. This identity remains valid for complex t, s by polynomial extension, or directly by using independent complex means and the bilinear cross covariance. For small complex t, s, every $d_{0,i}$ has modulus comparable to $1 + \tau$, with constants uniform in i, n. In particular $m_i/d_{0,i}$, $1/d_{0,i}$ are uniformly bounded.

21.1 A smooth bound for the absolute product

Put $r_i = |s|^2(|z_{L,i}|^2 + |z_{R,i}|^2)$ and choose a smooth cutoff $\chi(r)$ equal to one for $r \le \delta^2$, zero for $r \ge 4\delta^2$, and between zero and one, with δ fixed small. Let $\chi_i = \chi(r_i)$. On its support $|d_i/d_{0,i} - 1| \le 1/2$ by the bounds above and smallness of s, δ. The following is a smooth real upper bound for $\log|d_i/d_{0,i}|$:

$$F_i = \chi_i \log|d_i/d_{0,i}| + (1 - \chi_i)C r_i.$$

Here the first term is extended by zero outside the region where the logarithm is defined. Indeed for $r_i \ge \delta^2$ the ratio is bounded in modulus by $C'(1 + r_i)$, so Cr_i is an upper bound with C fixed

large. Globally $F_i \geq -\log 2$. As a function of the four real coordinates in the row, its Hessian has norm $O(|s|^2)$: on the cutoff annulus differentiation acts on bounded scaled fields sz, and outside it the function is quadratic. Its needed higher local derivatives are also uniformly bounded after taking account of these scale factors.

The conditional Gaussian mean is

$$\mathbb{E}F_i = -\operatorname{Re}\bigl[s^2(m_i/d_{0,i})^2 Y_{ii}\bigr] + O(|s|^4 L_i^2). \tag{21}$$

For this identity write

$$d_i/d_{0,i} = 1 + s(m_i/d_{0,i})(z_{L,i} + z_{R,i}) + s^2(z_{L,i}z_{R,i} - Y_{ii})/d_{0,i}.$$

The linear and direct quadratic expectations vanish. The logarithmic quadratic term has the expectation in (21) since $\mathbb{E}(z_L + z_R)^2 = 2Y_{ii}$ in that row. Odd total powers of the Gaussian fields have zero expectation. Taylor terms through weighted degree three in amplitude, treating $s^2 Y_{ii}$ as degree two, leave an absolute error bounded by $C|s|^4[(|z_{L,i}| + |z_{R,i}|)^4 + L_i^2]$ in the small region. On its complement the same bound holds for the difference between the smooth envelope and the same finite Taylor polynomial: $r_i \geq \delta^2$, the envelope grows at most quadratically in the scaled fields, and terms of degree at most three are bounded by a fixed multiple of $r_i^2 + |s|^4 L_i^2$ there. This also covers the cutoff annulus. Taking expectations proves (21).

For $F = \sum_i F_i$ we use Gaussian rotation [15]. For a function f of a standard real Gaussian, interpolation to an independent copy by rotation gives $\|f - \mathbb{E}f\|_p \leq C\sqrt{p}\,\|\|\nabla f\|\|_p$. At each interpolation angle the angular derivative is an independent standard Gaussian, so the estimate follows by Jensen and Minkowski. The vector version bounds the centered gradient by $C\sqrt{p}$ times the L^p Frobenius norm of the Hessian. Applied here these estimates give

$$\|F - \mathbb{E}F\|_p \leq C(1 + \sqrt{\tau})\sqrt{p} + C|s|^2 p. \tag{22}$$

Here are the uniform ingredients for applying that estimate to the combined paired field. Its local block Hessians are $O(|s|^2)$, so after composition with the paired covariance the Frobenius norm is $O(|s|^2)$ by the block trace formula and the paired covariance Frobenius bound. The operator norm has the same order. The mean gradient array, before multiplication by the covariance square root, differs from a constant row array by a Euclidean-bounded error polynomial in $1 + \sqrt{\tau}$. Indeed smoothness in $\ell_i, a_{0,i}, Y_{ii}$ costs $C(|h_i| + L_i)$ times such a polynomial, and Taylor in the centered row Gaussian makes the local mean-gradient error $O(L_i)$ by cancellation of the linear term. The constant array has bounded covariance-square-root image by the variance bounds on the two coordinate sums. For the envelope mean gradient the bound can be taken $C(1+\sqrt{\tau})$, using uniformly bounded first and higher derivatives of the logarithm in the scaled variables, $m_i/d_{0,i}$ uniformly bounded and its derivative in ℓ_i uniformly bounded, and $1/d_{0,i}$ uniformly bounded with the same property. The envelope outside the cutoff has no dependence on ℓ. Apply Gaussian rotation to F and the gradient norm, whose Lipschitz constant is $C|s|^2$; its mean is bounded by its L^2 norm as controlled by the mean gradient and Frobenius Hessian. This yields (22), which gives bounded conditional q-exponential moments up to $C_q e^{C_q \tau}$ for any one prescribed q, by taking the amplitude radius sufficiently small.

21.2 Analytic normalization and a zero-free neighborhood

Use the scalar reference density in τ proportional to $e^{-n\tau} d_*^n$, $d_* = 1 - c_n^2/n + c_n^2\tau$, and the original paired Gaussian. At $\tau_0 = 1/2$, set the holomorphic anchors

$$a(t) = \sum_i \log[d_{0,i}(\tau_0)/d_*(\tau_0)], \qquad b(t,s) = -s^2 \sum_i [m_i(\tau_0)/d_{0,i}(\tau_0)]^2 Y_{ii}.$$

The logs are the branches near one. Normalize the exact polynomial permanent and cofactor expectations by the common nonzero holomorphic factor $\mathcal{N} = (\operatorname{per} A_{*,n}) \exp(a+b)$. Under the scalar reference, their absolute integrands are bounded by the positive weight

$$\mathcal{W} = \exp\left(\operatorname{Re}[\sum_i \log(d_{0,i}/d_*) - a - b] + F \right)$$

times the relevant inserted numerator factors. Omitting one or two diagonal factors costs at most an additional fixed constant, from $|d_{0,i}|$ uniformly bounded below and $F_i \geq -\log 2$. The weight has bounded L^q norm for a fixed $q > 1$, with all fixed polynomial τ, X moments by Hölder after slightly increasing the chosen q. For this, the difference of the deterministic log term from its radial anchor is $O(|X|)$ on $[1/4, 3/4]$, since its radial derivative difference from the reference is $O(\sum_i |h_i|) = O(\sqrt{n})$. Elsewhere its absolute difference is $O(\sqrt{n})$: each log ratio is $O(|h_i|)$ uniformly for all τ. The conditional mean of $F - \operatorname{Re} b$ has the same bounds by (21), $\sum |Y_{ii}| \leq \sqrt{n}$, and bounded derivatives of $(m_i/d_{0,i})^2$ on the central interval. Relative centered fluctuations use (22). Finally the shifted scalar Gamma reference has bounded exponential moments of any fixed multiple of $|X|$, and tails outside this interval absorb $e^{C\sqrt{n}+C\tau}$.

In particular $\operatorname{per} A(t,s)/\mathcal{N}$ is uniformly bounded on the polydisk and equals one at the origin. By multivariable Cauchy estimates it is bounded away from zero in modulus on a smaller fixed polydisk. This uses only a uniform bound and its value at zero, and all functions are holomorphic there. No positivity for complex parameters is presumed.

21.3 Uniform cofactor blocks

Write $m_{L,i} = m_i + sz_{L,i}$, $m_{R,i} = m_i + sz_{R,i}$. Define truncated ratios $\tilde{v}_{L,i} = \chi_i m_{L,i}/d_i$, $\tilde{v}_{R,i} = \chi_i m_{R,i}/d_i$, extended smoothly as zero. Put $v_* = c_n\sqrt{\tau}/d_*$, $S_L = \sum_i (\tilde{v}_{L,i} - v_*)$, and similarly S_R. Both sums have bounded conditional L^p norms polynomial in $1 + |X|, 1 + \sqrt{\tau}$ for each fixed p, uniformly in the parameters. Their centered Gaussian fluctuations use the same double-rotation block estimate as above: local Hessian blocks are polynomially bounded, the covariance-weighted Frobenius norm is controlled by block trace, and the expected local gradients differ from a constant by $O(|h_i| + L_i)$ with polynomial factors. For the conditional means, use the same small-field expansion and outside-cutoff bound as for (21). The rowwise mean is

$$\mathbb{E}\tilde{v}_{L,i} = \frac{m_i}{d_{0,i}} + s^2 Y_{ii} \frac{m_i(m_i^2 - a_{0,i})}{d_{0,i}^3} + O\big(|s|^4 L_i^2 (1+\sqrt{\tau})^C\big). \tag{23}$$

Indeed for $v_L = p/(a+pq)$ the second-order paired generator is $(\partial_p\partial_q - \partial_a)v_L = p(pq-a)/(a+pq)^3$. The covariance is Y_{ii}, the pure Gaussian covariances are zero, and the residual change is $-s^2 Y_{ii}$. Odd degree terms have zero mean. The cutoff complement has the same quartic expectation bound, since the truncated ratio and the Taylor polynomial of degree at most three are bounded by $C(1+\sqrt{\tau})^C(r_i^2 + |s|^4 L_i^2)$ when $r_i \geq \delta^2$. This proves (23) and its right counterpart.

The coefficient in (23) is bounded by a polynomial in $1+\sqrt{\tau}$ times $|h_i| + |\tau - 1/2| + 1/n$. Hence its sum costs at most $C(1+|X|)$ with the polynomial factors, by $\sum L_i|h_i| \leq 1$, $\sum L_i \leq \sqrt{n}$. The scalar baseline differences $m_i/d_{0,i} - v_*$ obey the same order after summing: Taylor in $\ell_i - c_n = th_i$ and use that the first derivative of $\ell\sqrt{\tau}/[1+\ell^2(\tau - 1/n)]$ at c_n vanishes to order $O((1+\sqrt{\tau})(|\tau - 1/2| + 1/n))$. Its linear term sums with $\sum |h_i| \leq \sqrt{n}$ and its second remainder with $\sum h_i^2 = 1$. Thus the two sums are uniformly controlled as asserted.

For a deterministic unit vector f, the transverse scalar tests $f^* P\tilde{v}_L$ and $\tilde{v}_R^{\mathsf{T}} P f$, where $P = I - ee^*$, $e = \mathbf{1}/\sqrt{n}$, have uniformly bounded L^p norms with polynomial radial factors. Subtract v_* and

use rowwise Taylor in the local Gaussian at fixed residual a_i. The constant term has Euclidean norm bounded via $\|h\|_2, \|L\|_2$, the linear field term has bounded variance by the paired covariance operator norm, and the second remainder tested against Pf uses $\sum_i |(Pf)_i| L_i \le 1$. The truncated row ratios have smooth uniformly bounded local second derivatives with polynomial radial factors.

For the omitted-coordinate terms put $g_{L,i} = |m_{L,i}| \mathbf{1}_{r_i > \delta^2}$, and similarly g_R. These have bounded conditional L^p norms in total ℓ^1: each row L^p norm is at most $C_p(1+\sqrt{\tau})^C L_i^2$, by the Gaussian moment bound at sufficiently high order on the event $|s|^2(|z_{L,i}|^2 + |z_{R,i}|^2) > \delta^2$, then Hölder for the numerator. Sum using $\sum L_i^2 \le 1$. Their Euclidean norms are bounded by their ℓ^1 norms.

The exact off-diagonal cofactor polynomial splits by $\chi_i\chi_j$, $(1-\chi_i)\chi_j$, $\chi_i(1-\chi_j)$, and $(1-\chi_i)(1-\chi_j)$. In the first term the integrand is the full permanent product times $\tilde{v}_{L,i}\tilde{v}_{R,j}$. In the second, it factors as

$$\big[(1-\chi_i) m_{L,i} \prod_{k \ne i} d_k\big] \tilde{v}_{R,j} \qquad (i \ne j).$$

The third is the reverse version. On the fourth, omitted-factor bounds give the entrywise majorant $C\mathcal{W} g_{L,i} g_{R,j}$ relative to the scalar anchor. The bracketed vector in the second term is bounded entrywise by $C\mathcal{W} g_{L,i}$ relative to the anchor, and the same for the third. The diagonal discrepancies if one extends these formulas to all i, j are bounded as diagonal matrices: truncated ratios are bounded by a polynomial in $\sqrt{\tau}$, and the g arrays have bounded norms. The exact diagonal cofactor is itself bounded entrywise by the one-factor omission envelope.

These facts give the following anchored matrix bounds. The e, e entry of the first term minus $n/4$ times the anchored permanent is uniformly bounded: use

$$n v_*^2 + v_*(S_L + S_R) + S_L S_R / n, \qquad n v_*^2 - n/4 = O(1 + X^2),$$

with the exact scalar defect identity for the latter. The one-bad terms in this entry cost $O(1)$, since testing the bad bracket against e uses its ℓ^1 norm divided by $\sqrt{n}$, while the good factor against e is $O(\sqrt{n})$ in fixed L^p; two bad terms also cost $O(1)$. The lower block $PC^{\mathsf{T}}P$ divided by the anchor is $O(1)$ in operator norm, by testing against arbitrary deterministic unit vectors and Hölder with $\mathcal{W}$, using the transverse scalar bounds or bad Euclidean bounds. Both cross blocks divided by the anchor are $O(\sqrt{n})$ by the same estimates with one e test. All bounds persist for the normalized cofactor matrix after dividing by the zero-free scaled permanent.

21.4 The isolated eigenvalue

For complex t, s in the small fixed polydisk, the lower block R has norm $O(1)$, the upper entry is $n/4 + O(1)$, and the two cross blocks have norms $O(\sqrt{n})$. The Schur equation therefore has one root $\lambda = n/4 + O(1)$ uniformly, by Rouché on a fixed circle centered at $n/4$ sufficiently large to dominate the bounded correction. The derivative of the scalar Schur equation in λ differs from one by $O(n^{-1})$, so the root is simple and holomorphic in the parameters. For real parameters $A(t, s)$ is positive definite because its diagonal residual stays uniformly positive and the signed perturbation has small norm. The normalized cofactor matrix is then Hermitian. The lower block bound gives that this root is its largest eigenvalue. Cauchy estimates yield uniform derivatives through order six of $\lambda - n/4$. The continued matrix is even in s, so its isolated eigenvalue is even. For an even function f, $\partial_\alpha f(\sqrt{\alpha}) = \frac{1}{2}\int_0^1 f''(u\sqrt{\alpha})\, du$. Iteration of this formula yields uniform derivatives up to order three in (t, α), including at $\alpha = 0$.

22 Uniform local comparison

22.1 Phase adjustment

For a real phase vector ϕ, let $D_\phi = \operatorname{diag}(e^{i\phi_i})$ and define

$$G_A(\phi) = \operatorname{Im}(D_\phi^* A D_\phi \mathbf{1}).$$

This is always in the real subspace $\mathbf{1}^\perp$, because the sum of the coordinates of $D_\phi^* A D_\phi \mathbf{1}$ is real by Hermiticity. Restrict ϕ to that subspace. The phase equation is the torus stationarity condition for the quadratic form on unimodular vectors, as in [18, Section 4.3]. The derivative at zero is $DG_A(0) = \operatorname{Re} A - \operatorname{diag}(\operatorname{Re} A\mathbf{1})$ on real directions. At $A_{*,n}$ and on $\mathbf{1}^\perp$ this is $-nb_nI$. The derivative difference is bounded by

$$\|A - A_{*,n}\|_{\mathrm{op}} + \|(A - A_{*,n})\mathbf{1}\|_\infty.$$

The derivative at ϕ has the same form with $D_\phi^* A D_\phi$ in place of A, since the diagonal phase changes commute. Its difference from the derivative at zero is $O(\|\phi\|_2)$ uniformly on fixed small balls. Indeed $\|D_\phi - I\|_{\mathrm{op}} \le \|\phi\|_2$, $\|(D_\phi - I)\mathbf{1}\|_2 \le \|\phi\|_2$, and

$$D_\phi^* A D_\phi \mathbf{1} - A\mathbf{1} = D_\phi^* A(D_\phi - I)\mathbf{1} + (D_\phi^* - I)A\mathbf{1}.$$

Its Euclidean norm is bounded by $C\|\phi\|_2$, since $\|A\|_{\mathrm{op}}$ and $\|A\mathbf{1}\|_\infty$ are uniformly bounded under (20). The same estimate bounds the sup norm. The operator-norm difference of the conjugated matrices is $O(\|\phi\|_2)$ as well.

Since $nb_n \to 2$, the map $\phi \mapsto \phi + G_A(\phi)/(nb_n)$ is a contraction on a fixed small ball in $\mathbf{1}^\perp$, provided δ is small enough. Its displacement at zero is bounded by $\|(A - A_{*,n})\mathbf{1}\|_2/(nb_n)$, so the map sends that ball into itself. It has a fixed point of norm at most $C\|(A - A_{*,n})\mathbf{1}\|_2$. For $A' = D_\phi^* A D_\phi$ the row sums are now real and

$$\|(A' - A_{*,n})\mathbf{1}\|_2 \le C\|(A - A_{*,n})\mathbf{1}\|_2.$$

Moreover $\|A' - A_{*,n}\|_{\mathrm{F}} \le \|A - A_{*,n}\|_{\mathrm{F}} + C\|\phi\|_2/\sqrt{n}$. To see the latter, compare $D_\phi^* A_{*,n} D_\phi$ to $A_{*,n}$, using $D_\phi^*\mathbf{1}$ within $\|\phi\|_2$ of $\mathbf{1}$ and the rank-one scale $b_n = O(1/n)$. The Frobenius norm of the perturbation conjugates unchanged.

22.2 Longitudinal fit

Let $r = A'\mathbf{1} - \mathbf{1}$, real. Use the row map

$$\mathcal{R}(\ell)_i = \ell_i\big(\textstyle\sum_j \ell_j - \ell_i\big)/n.$$

At $c_n\mathbf{1}$ its value is $(n-1)b_n\mathbf{1}$ and its real derivative is $c_n[(n-2)I + \mathbf{1}\mathbf{1}^{\mathsf{T}}]/n$, uniformly invertible for large n. The quadratic remainder at $x = \ell - c_n\mathbf{1}$ has Euclidean norm at most $(n^{-1/2} + n^{-1})\|x\|_2^2$, and its derivative on a ball is bounded by twice the same constant times the ball radius, by the bilinear formula for $x_i(\sum x_j - x_i)/n$. The contraction after applying the inverse derivative gives a real solution ℓ near $c_n\mathbf{1}$ with

$$\|\ell - c_n\mathbf{1}\|_2 \le C\|r - (n-1)b_n\mathbf{1}\|_2.$$

Its coordinates are positive since $c_n \to \sqrt{2}$ and the radius is small.

Write

$$T = A' - \left[I + \ell\ell^{\mathsf{T}}/n - \operatorname{diag}(\ell_i^2/n)\right].$$

Then T is Hermitian with zero diagonal and $T\mathbf{1} = 0$ exactly. Its Frobenius norm is bounded by $\|A' - A_{*,n}\|_{\mathrm{F}} + C\|\ell - c_n\mathbf{1}\|_2/\sqrt{n}$ by the rank-one difference. Set $t = \|\ell - c_n\mathbf{1}\|_2$, $h = (\ell - c_n\mathbf{1})/t$ when nonzero, $\alpha = \|T\|_{\mathrm{F}}$, $Y = T/\alpha$ when nonzero. For zero values use arbitrary real unit h or $Y = 0$ respectively. The holomorphic chart in Section 21 applies because $Y\mathbf{1} = 0$ implies $|Y|^{1/2}\mathbf{1} = 0$. Its uniform C^3 estimate holds as functions of t, α.

22.3 Hessian comparison

The first affine derivative of Λ at $A_{*,n}$ vanishes by permutation and conjugation symmetry and scalar stationarity. The chart Hessian therefore is the affine Hessian on the tangent $tX + \alpha Y$, where $X_{ij} = c_n(h_i + h_j)/n$ for $i \neq j$, $X_{ii} = 0$. The real constant and real-star spaces contain X, with $n\|X\|_{\mathrm{F}}^2$ bounded below for every real unit h. The Hermitian normal direction Y is entirely in the real and imaginary zero-row-sum spaces. The affine Hessian Theorem 1 gives real rates of negative order n in the former constant/star spaces and negative fixed order in the latter normal spaces. Thus

$$D^2_{\mathrm{chart}}\Lambda[(t,\alpha),(t,\alpha)] \leq -c_0(t^2 + \alpha^2)$$

for some absolute $c_0 > 0$ and sufficiently large n, since these Hessian sectors are orthogonal and Y has unit norm if present. The uniform third derivative gives a remainder at most $C(|t| + \alpha)^3$. Hence for sufficiently small δ in (20) the comparison is strict unless $t = \alpha = 0$. In that case $A' = A_{*,n}$, yielding precisely the gauge equality.

23 A signed transverse loss formula

23.1 Paired moments and the absolute envelope

For the signed class of Theorem 7, we prove the quantitative formula

$$\Lambda(A) = \frac{n}{4} + \frac{1}{2} - \frac{u^2}{16} - \frac{1}{16}\operatorname{tr}\big[\operatorname{Re}(K^2)(I - \operatorname{Re}K/2)^{-1}\big] + o(1). \tag{24}$$

The error is uniform at fixed B, ϵ, M. On the nonzero spectrum write $K = U\operatorname{diag}(\kappa_j)U^*$ with U an isometry and $U^*\mathbf{1} = 0$. Put $D = \operatorname{diag}(\kappa_j, \kappa_j)$ on the realification. For a real vector ξ corresponding to a complex vector η, let $L\xi = \operatorname{Im}(U\eta)$, and put $Q = L^{\mathsf{T}}L$, so $0 \preceq Q \preceq I$. Write $S = \operatorname{diag}(\sqrt{\kappa_j}, \sqrt{\kappa_j})$ on this realification, with imaginary square roots for $\kappa_j < 0$, so $S^2 = D$. For independent standard circular complex η_j, put

$$z_L = U\sqrt{\kappa}\,\eta, \qquad z_R = \overline{U}\sqrt{\kappa}\,\overline{\eta}, \qquad Y = \sum_j |\kappa_j|\,|\eta_j|^2.$$

There is no conjugation on the selected square root in z_R. Thus $\mathbb{E}z_L z_R^{\mathsf{T}} = K$, while $\|z_L\|^2 = \|z_R\|^2 = Y$, and both vectors have zero sum. Each component of either vector has circular Gaussian marginal variance $(|K|)_{ii} \leq B^2/n$. Also $\sum_j |\kappa_j| \leq B^2$.

Let $x = T/n$ with T exponential of mean one, the radial variable of the longitudinal latent complex standard Gaussian; absorb its phase into η. Put $m = \sqrt{\beta x}$, $a_i = 1 - \beta/n - K_{ii} > 1/2$ eventually, and $d_i = a_i + (m + z_{L,i})(m + z_{R,i})$. The paired Wick identity [17, Theorem A.4(c)] gives $\operatorname{per} A = \mathbb{E}_0 \prod_i d_i$. To see why, add independent diagonal Gaussian terms $\sqrt{a_i}\,G_i$ and $\sqrt{a_i}\,\overline{G}_i$ in the two systems. Each left-right covariance is A_{ij}. Pairings entirely within the left or right system vanish by circularity, so the moments give the permanent, even though the systems are not complex conjugates. Conditioning on the latent factors yields the displayed product. A diagonal

cofactor omits d_i. An off-diagonal cofactor omits d_i, d_j and inserts $(m+z_{L,j})(m+z_{R,i})$. Where the denominators are nonzero the normalized matrix, up to transpose, is the normalized expectation with this complex weight of

$$\operatorname{diag}(a_i/d_i^2) + v_L v_R^{\mathsf{T}}, \qquad v_{L,i} = (m+z_{L,i})/d_i, \quad v_{R,i} = (m+z_{R,i})/d_i.$$

For tail estimates use the original polynomial products rather than ratios.

The scalar reference has $\bar{a} = n^{-1}\sum a_i$, $d_0(x) = \bar{a} + \beta x$, and radial weight $e^{-nx} d_0(x)^n$. The complex tilt is globally controlled by Cauchy–Schwarz,

$$|d_i| \le \sqrt{(a_i + |m+z_{L,i}|^2)(a_i + |m+z_{R,i}|^2)}.$$

Taking products and applying the arithmetic–geometric mean on each side, using zero sums and the squared norms Y, gives $\prod |d_i| \le (\bar{a} + \beta x + Y/n)^n$. If one or two factors are omitted, the same bound holds up to a fixed constant since the added envelope is bounded below, the sum over remaining terms is at most the full sum, and $(n/(n-j))^{n-j} \le e^j$. The inserted factors grow at most polynomially in x, Y. Hence the same exponential envelope bounds all original cofactor polynomials up to a fixed polynomial and at most a factor n for operator norms.

Translate the radial variable by $q = x + Y/(n\beta)$ when using the envelope. The resulting scalar q-integral, extended to $q \ge 0$, is exactly the scalar reference normalizer, with multiplier $e^{Y/\beta}$. For a fixed sufficiently small $\delta > 0$, $(1/\beta + \delta)\|K\| < 1$ by the gap. The exponential moments of Y to this rate are uniformly bounded, since $\sum |\kappa_j| \le B^2$ and Y is a weighted sum of unit exponentials. The envelope after dividing by the scalar normalizer therefore gives bounded moments of every fixed order in Y and $\zeta = \sqrt{n}(x - x_0)$, $x_0 = 1 - 1/\beta$. Indeed the scalar q has the shifted Gamma law with scaled mean at $x_0 + O(1/n)$, and the $q - x$ shift is $Y/(n\beta)$. Tails with $x \notin [1/4, 3/4]$ are exponentially small in n, using the Gamma bound and $\Pr(Y \ge cn)$ under the exponential envelope. Also on the envelope, the distribution of η is a product circular Gaussian with variances multiplied by $(1 - |\kappa_j|/\beta)^{-1}$, uniformly bounded. Thus $\max_i(|z_{L,i}| + |z_{R,i}|) > \delta_0$ has an exponentially small contribution for any fixed $\delta_0 > 0$, by the row variance and Gaussian union bound, while replacing δ_0 by $n^{-1/4}$ still gives a tail tending to zero faster than any power. These estimates hold with fixed polynomial moments and the factor n. They justify using ratio formulas on the central region and comparing moments there without any positivity of the complex likelihood.

23.2 The complex quadratic comparison

On bounded $|\zeta|, Y$ and $\max_i(|z_{L,i}| + |z_{R,i}|) \le n^{-1/4}$, the complex log relative to the scalar d_0^n, choosing the branch near each 1, is

$$\sum_i \log(d_i/d_0) = -\tfrac{1}{8}\sum_i (z_{L,i} - z_{R,i})^2 + o(1) = \tfrac{1}{2}\,\xi^{\mathsf{T}} SQS\xi + o(1).$$

The linear sums of both fields and of $a_i - \bar{a}$ are zero. At the limiting $m = 1, d_0 = 2$, the quadratic Taylor term is $z_L z_R/2 - (z_L + z_R)^2/8$. Replacing the nearby radial coefficients costs $o(1)$ in the bounded squared-field sum. The diagonal residual differences are $O(1/n)$ per coordinate, have zero sum, and give $o(1)$ by Cauchy–Schwarz. Third-order errors are $O(\max_i(|z_{L,i}| + |z_{R,i}|)Y) + o(1)$. The identity with SQS follows by writing $z_L - z_R = 2i\,LS\xi$, extending the real map L complex-linearly.

The comparison weight multiplies the original density $e^{-\|\xi\|^2}$ by $e^{\xi^{\mathsf{T}} SQS\xi/2}$, giving the complex symmetric precision $H = I - SQS/2$. Its real part is positive definite uniformly because $\|SQS\| \le \|K\|$. The complex Gaussian integral is $\det(H)^{-1/2}$ with the branch continued from $S = 0$, and its

normalized second moment is $H^{-1}/2$. These formulas follow by analytic continuation of the real Gaussian integral or integration by parts and a determinant differentiation along the line from I to H. The normalization is bounded in modulus above and below, since $\|SQS\|_1 \le \|S\|_{\mathrm{F}}^2 \le 2B^2$ and $\|SQS/2\| < 1$, by the convergent trace-log series. Its absolute weight is dominated by $e^{Y/2}$ under the original circular law, as

$$|\xi^{\mathsf{T}} SQS\xi| \le \|S\xi\|^2 = Y.$$

Thus it has the same moment and coordinate tail bounds as the original absolute envelope, and preserves independence of the radial variable. By truncating on the central sets above, the unnormalized original and comparison likelihoods relative to the scalar reference integrate to within $o(1)$, also after any fixed polynomial weight in ζ, Y. The nonzero lower bound on the comparison normalization permits the same statement for their normalized moments. The normalized comparison is complex, with uniformly bounded total variation.

23.3 Spectral comparison

Let $e = \mathbf{1}/\sqrt{n}$, $P = I - ee^*$ and $v_0 = m/d_0$. The scalar radial law has ζ tending to $N(0,1)$ with moments. At the central points the average of v_L has zero linear variation in the fields and the residual, and likewise for v_R. For the second-order terms in v_L at $m = 1, a = 1$,

$$\frac{1+z_L}{2+z_L+z_R+z_Lz_R} = \frac{1}{2} + \frac{z_L - z_R}{4} - \frac{z_Lz_R}{4} + \frac{z_R^2 - z_L^2}{8} + O((|z_L|+|z_R|)^3).$$

Interchange L, R for the other vector. Thus n times the product of their averages equals

$$nv_0^2 - \tfrac{1}{4}\sum_i z_{L,i}z_{R,i} + o(1)$$

in normalized expectation. To justify the constant-order accuracy, on the near-central sets the n-scaled errors of cubics are bounded by $o(1)$ from $\max_i(|z_L| + |z_R|)Y$; residual shifts and radial coefficient differences are $o(1)$ as in the logarithmic expansion. On the larger central neighborhood with bounded denominators, the same differences from nv_0^2 are bounded by $C(1+Y)$ using exact linear cancellation and uniformly bounded second derivatives. Then polynomial moment tails and the exponentially small outer complement justify the conclusion. The diagonal term $e^* \operatorname{diag}(a_i/d_i^2)e$ tends to $1/4$.

Use $nv_0^2 = n/(4\bar{a}) - n(\beta x - \bar{a})^2/(4\bar{a}d_0^2)$. Its first term is $n/4 + (\beta + \operatorname{tr} K)/4 + O(n^{-1})$; the expectation of its defect tends to $1/4 + u_n^2/16$ up to $o(1)$ by bounded radial moments, since $\sqrt{n}(\beta x - \bar{a}) = u_n + \beta\zeta + O(n^{-1/2})$. The transverse contribution is

$$\mathbb{E}^{(2)} \sum_i z_{L,i}z_{R,i} = \mathbb{E}^{(2)} \xi^{\mathsf{T}} D\xi = \tfrac{1}{2}\operatorname{tr}(DH^{-1}).$$

The identity before expectation uses column orthonormality of U; absolute magnitude is at most Y. These evaluations give the e-Rayleigh quotient $n/4 + 1/2 - u_n^2/16 - \operatorname{tr}[D(H^{-1} - I)]/8 + o(1)$.

The actual normalized cofactor matrix is Hermitian since A is Hermitian, and has a bounded lower block on $e^{\perp}$. On the central regions $\|Pv_L\| + \|Pv_R\| \le C(\sqrt{Y} + n^{-1/2})$, using the scalar subtraction and uniform Lipschitz estimates. The Gaussian quadratic comparison and exponential tails give the bound for their paired product in operator norm. Moreover $Pv_L = (z_L - z_R)/4 + o_{L^2}(1)$ on the central region with denominators bounded below, in the corresponding absolute-weighted L^2 norm, and similarly with opposite sign for Pv_R. This follows from the first differential and a remainder of Euclidean norm $O(\max_i(|z_L| + |z_R|)\sqrt{Y}) + o(1)$ on fixed scaled boxes. On their

complement within the bounded-denominator region use the uniform moments of Y, ζ; outside that region use the omitted-factor polynomial estimate for the cofactor blocks. Since $e^* v_L/\sqrt{n}$ and $v_R^{\mathsf{T}} e/\sqrt{n}$ tend to $1/2$, the off-diagonal block divided by $\sqrt{n}$ tends to zero by the evenness of the comparison Gaussian, which kills the linear terms. The Schur complement thus changes the Rayleigh quotient by $o(1)$.

Finally

$$H^{-1} - I = \tfrac{1}{2} S Q^{1/2} (I - Q^{1/2} D Q^{1/2}/2)^{-1} Q^{1/2} S.$$

The matrix inside the inverse is positive definite and real. Taking the trace after multiplying the displayed identity by D and using $SDS = D^2$ gives half the trace of $D^2 Q^{1/2}(I - Q^{1/2} D Q^{1/2}/2)^{-1} Q^{1/2}$. By polar factorization of L, or direct series expansion, the expression between D^2 and the trace is $L^{\mathsf{T}}(I - LDL^{\mathsf{T}}/2)^{-1} L$. In grouped real and imaginary coordinates $L = [\operatorname{Im} U, \operatorname{Re} U]$, so $LDL^{\mathsf{T}} = \operatorname{Re} K$ and $LD^2 L^{\mathsf{T}} = \operatorname{Re}(K^2)$. This proves (24).

The subtracted transverse loss is at least $\|K\|_{\mathrm{F}}^2/[16(1 + \|K\|/2)]$ since $\operatorname{Re}(K^2) \succeq 0$, its trace is $\|K\|_{\mathrm{F}}^2$, and $I - \operatorname{Re} K/2 \preceq (1 + \|K\|/2) I$. For real symmetric K the loss is $\sum_j \kappa_j^2/[8(2 - \kappa_j)]$. The complex orientation can change the loss at fixed spectrum. For centered real orthonormal v, w whose coordinates are $O(n^{-1/2})$, the two matrices $a(vv^{\mathsf{T}} - ww^{\mathsf{T}})$ and $ia(vw^{\mathsf{T}} - wv^{\mathsf{T}})$, $0 < a < 1$, have the same eigenvalues $\pm a$ and the same $|K|$. Their losses are respectively $a^2/[2(4 - a^2)]$ and $a^2/8$. The corresponding correlation matrices are positive definite for sufficiently large n.

24 Exact comparison in the two global classes

24.1 The spectral floor

Fix $B \geq \sqrt{2}$. For all sufficiently large n, the exact scalar equicorrelation optimizer is the unique maximizer modulo diagonal phases among all correlation matrices $A \succeq (1 - B^2/n) I$, without a rank bound. Equivalently $A = I + G - \operatorname{diag} G$, $G \succeq 0$, $G_{ii} \leq B^2/n$. Indeed either representation implies the other by the scalar spectral shift. The scalar optimum is feasible even at $B = \sqrt{2}$, since $n b_n < 2$ for large n.

For any sequence of maximizers the rank-independent Gaussian localization Theorem 6 gives a gauge in which

$$A = I + \ell \ell^*/n - \operatorname{diag}(|\ell_i|^2/n) + \Omega - \operatorname{diag} \Omega, \quad \|\ell - c_n \mathbf{1}\|_2 = o(1), \quad \Omega \mathbf{1} = 0, \quad \|\Omega\|_{\mathrm{F}} = o(1).$$

If the original Gram convention conjugates ℓ or Ω, use the conjugated versions; the displayed bounds persist. The bounded row cap gives $\|\operatorname{diag} \Omega\|_{\mathrm{F}} \leq B^2/\sqrt{n}$. Frobenius closeness follows at once. The row-sum difference also tends to zero in Euclidean norm: the longitudinal rank-one row vector differs from $c_n^2 \mathbf{1}$ by $O(\|\ell - c_n \mathbf{1}\|_2)$ using the sample mean, the longitudinal diagonal difference is smaller, $\Omega \mathbf{1} = 0$, and the diagonal vector of Ω has vanishing Euclidean norm. Apply (20). Compactness of the spectral class and the scalar test handle the all-sufficiently-large quantifier. Every maximizing matrix for those orders thus lies on the scalar orbit, which stays feasible under diagonal phases.

24.2 The centered signed class

Fix $B < \infty, \epsilon > 0, M > 0$. In the signed class from the paired Gaussian theorem, take $K \mathbf{1} = 0$, $\|K\| \leq 2 - \epsilon$, $(|K|)_{ii} \leq B^2/n$, $\beta = 2 + u/\sqrt{n}$, $|u| \leq M$, and keep $A = \operatorname{diag}(1 - \beta/n - K_{ii}) + \beta J/n + K$ positive semidefinite. Assume $0 < \epsilon \leq 2$ so the zero correction is allowed. For all sufficiently large n, its exact maximum is the scalar optimum and equality requires a matrix on the scalar diagonal-unitary orbit. The scalar optimum itself lies in this class. The paired Gaussian loss (24) implies

$u = o(1)$, $\|K\|_{\mathrm{F}} = o(1)$ along any maximizing sequence. The Frobenius difference from $A_{*,n}$ is $o(1)$. Its row-sum difference is

$$(\beta - nb_n)(1 - 1/n)\mathbf{1} - \operatorname{diag}(K)\mathbf{1},$$

which has Euclidean norm $o(1)$ by $u = o(1)$ and the row cap. Thus (20) applies even though matrices in this signed class may violate the earlier spectral floor by a fixed amount. The class is compact and includes the scalar test. The only orbit representative in this class is $A_{*,n}$ itself for large n. Indeed any such representative has the form $(1-b_n)I+b_nvv^*$, $|v_i| = 1$, while every row sum in the signed class is real and $3 + o(1)$ uniformly. Thus $b_nv_i(v^*\mathbf{1}) = (A\mathbf{1})_i - (1 - b_n) > 0$ for each i, forcing all v_i to be equal. The uniqueness concerns the matrix and allows multiple representations.

25 Exact maxima with bounded-size block noise

For a complex correlation matrix A, let $C(A)_{ij}$ be the permanent with row i and column j deleted. Set $\Lambda(A) = \lambda_{\max}(C(A))/\operatorname{per} A$, the raw cofactor ratio used in [3]. Let $J = \mathbf{1}\mathbf{1}^*$ and let $A_{*,n} = (1 - b_n)I + b_nJ$ denote the exact scalar equicorrelation optimizer of Theorem 1. It has $nb_n < 2$ in all sufficiently large orders.

Theorem 8. *Fix $k \in \mathbb{N}$ and $B_0 \geq 2$. Let $B = \bigoplus_b B_b$ be block diagonal with positive semidefinite complex correlation blocks of order at most k. There is no bound on the number of blocks, and the partition can depend on n. For $0 \leq \beta \leq B_0$, put*

$$A = (1 - \beta/n)B + \beta J/n.$$

For all sufficiently large n the unique maximizing matrix of Λ over this class is $A_{*,n}$, and the maximum is the exact scalar value. It equals $(n+1)^2/(4n) + O(ne^{-cn})$. Allowing diagonal-unitary conjugates of the class gives equality precisely on the scalar orbit.

Permanents of a block-diagonal matrix plus a correction of bounded rank occur in the computation of permanents and deleted minors [19, Proposition 2.1]. Scalar inequalities and asymptotics for block matrices are also studied in [20, 21]. Here the objective is the top eigenvalue of the complete normalized cofactor matrix. The block partition can contain many independent, strongly correlated components. A perfectly correlated pair gives a singular example outside the near-identity spectral floor and bounded signed diagonal cap of Theorem 7. Its exact constant-order loss below is 13/98. The uniform local theorem of Theorem 7 is the final comparison used here, after a separate localization argument for block noise.

26 A creation-operator inequality for one missing factor

Use the circular complex Gaussian Fock inner product $\langle f, g\rangle = \mathbb{E}\overline{f}g$ on holomorphic polynomials. Let G_i be a centered Gaussian linear form of variance $a > 0$, $L_i = \mu + G_i$ a multiplication operator, and $L_i^* = \bar{\mu} + D_i$ its adjoint, with D_i the corresponding directional derivative. In the Gaussian Fock inner product, $D_i = \sum_j \overline{g_{ij}}\partial_{z_j}$ if $G_i(z) = \sum_j g_{ij}z_j$. The commutator $[L_i^*, L_i] = a$ gives

$$\|L_iF\|^2 = \|L_i^*F\|^2 + a\|F\|^2, \qquad \frac{|\langle L_iF, F\rangle|}{\|L_iF\|^2} \leq \frac{1}{2\sqrt{a}}. \tag{25}$$

The ratio follows by Cauchy–Schwarz and the scalar inequality $\sqrt{y}/(y + a) \leq 1/(2\sqrt{a})$ for $y \geq 0$. More precisely for

$$p = \|F\|^2, \qquad y = \|L_i^*F\|^2/p, \qquad z = \langle F, L_i^*F\rangle/p$$

the squared deficit equals

$$\frac{1}{4a} - \frac{|z|^2}{(y+a)^2} = \frac{(y-a)^2}{4a(y+a)^2} + \frac{y-|z|^2}{(y+a)^2}. \tag{26}$$

Here $y - |z|^2 = \|(I - P_F)D_i F\|^2/p$, with P_F the orthogonal projection onto F. Equality in (25) forces $D_i F$ to be collinear with F, hence zero by degree, and then $|\mu|^2 = a$.

Take a block B_b with $\operatorname{Cov} G = aB_b$ and $F_i = \prod_{j\neq i}(\mu + G_j)$, $\mu > 0$. Then $D_i F_i = 0$ if and only if the row of B_b off diagonal vanishes. Indeed divide the derivative by F_i in the rational function field. The summands are $a(B_b)_{ji}/(\mu + G_j)$, with the conjugation convention immaterial. Equal affine factors correspond to identical Gaussian forms because the constant μ is nonzero, so their coefficients are equal and add to a nonzero multiple unless they vanish. Distinct affine factors have independent simple rational poles: multiply by their common product and reduce modulo any one affine factor, using that all other affine factors are nonzero in that integral domain. This proves the characterization, also for singular blocks. For the total zero of (26), the equality condition already makes $\mu > 0$.

27 Conditional cofactor matrix and uniform block coercivity

Condition on the common latent Gaussian, remove its circular phase, write $\tau = T/n$ with T originally exponential, $a = 1 - \beta/n$, $\mu = \sqrt{\beta\tau}$, $d = a + \mu^2$. In block b let $P_b = \|\prod_{i\in b}(\mu + G_i)\|^2$ under the block noise, and let v_i be the inner product of F_i with the full product divided by P_b, oriented to agree with the transpose of the raw cofactor matrix. The conditional permanent is $\prod_b P_b$, so the radial tilt has density proportional to $e^{-T}\prod_b P_b$. By the circular Gaussian moment formula [9], raw cofactors are Gram entries of deleted products. Within a block their normalized Gram is C_b; cross-block entries are $v_i\overline{v}_j$ by factorization. The normalized full cofactor matrix, up to transpose, is

$$M = \mathbb{E}[R + vv^*], \qquad R = \bigoplus_b (C_b - v_b v_b^*). \tag{27}$$

The matrices $C_b - v_b v_b^*$ are positive semidefinite by projection away from the block's full product. Their operator norms are bounded by k/a for all τ, since each cofactor diagonal $\|F_i\|^2/P_b \le 1/a$ by the creation identity. Each $|v_i|$ is bounded by (25). Let $\delta_i = 1/(4a) - |v_i|^2$ and $e_b = \|B_b - I\|_{\mathrm{F}}$.

On every compact range $1/2 \le a \le 2$, $0 \le \mu \le L$, fixed k, there is $c_{k,L} > 0$ such that in each block

$$\sum_{i\in b} \delta_i \ge c_{k,L}\big(e_b^2 + |b|(\mu^2 - a)^2\big). \tag{28}$$

Furthermore, on a compact range of positive μ the sum of the second terms in (26) is at least ce_b^2. Here are details of uniformity. Covariance can be realized using the positive square root of B_b and standard independent coordinates, so all polynomial norms and inner products are continuous on the compact set of block correlations. Denominators are bounded below by positive constants by iterating the creation identity in (25). For the total deficit the only zeros are $B_b = I, \mu^2 = a$ by the equality characterization. For the sum of the second terms at $\mu > 0$, any zero means every $D_i F_i = 0$ by the degree argument, so $B_b = I$. Around $B_b = I$ the map from off-diagonal row coefficients to $(I - P_{F_i})D_i F_i$ is uniformly bounded below in norm. At independence this is the map sending those coefficients to the projected $F_{ij} = \prod_{l\neq i,j}(\mu + G_l)$. These products are linearly independent by evaluation at $G_j = -\mu$ separately, and their span has degree less than F_i, so projection out of F_i is injective there. The lowest singular value at identity is bounded below on compact a, μ, including

$\mu = 0$, by continuity; it persists for B_b near identity using the positive square root there. The coefficients in the derivative are exactly the row entries $a(B_b)_{ji}$ off diagonal, so the second terms of (26) bound $\sum_{i\neq j} |(B_b)_{ij}|^2$ locally. Also $y_i - \mu^2 = O(e_b)$, since $D_i F_i$ is linear in those coefficients with bounded product polynomials. Hence a sufficiently small multiple of $|b|(\mu^2 - a)^2$ is bounded by a linear combination of the two deficit terms, by squaring $|\mu^2 - a| \leq |y_i - a| + Ce_b$. This gives (28) near the zeros; compactness away from the zeros gives (28) globally, and gives the separate variance bound for positive μ.

28 Global localization and the radial law

Consider any sequence with $\Lambda \geq n/4 + 1/2 - o(1)$. By (27) and $R \preceq (k/a)I$, $\mathbb{E}\sum \delta_i = O(1)$ for bounded β. The tilted law has exponentially small polynomial tails outside $0 \leq \tau \leq L_0$ for L_0 fixed large depending on k, B_0. Indeed $P_b \geq a^{|b|}$ by repeated creation removal, and $P_b \leq C^{|b|}(1+\tau)^{|b|}$ by bounded-size Gaussian moments. The product bound with $e^{-n\tau}$ and the positive lower normalizer $\prod_b a^{|b|} \geq c > 0$ gives the tail. Thus (28) in expectation and exponential tails imply

$$\sum_b e_b^2 = O(1), \qquad n\mathbb{E}(\mu^2 - a)^2 = O(1). \tag{29}$$

Take the scalar radial reference proportional to $e^{-n\tau} d^n$, and let $\mathcal{E}(\tau) = \sum_b \log(P_b/d^{|b|})$ be the log of the relative tilt. The bound $|\mathcal{E}| \leq C\sum_b e_b = O(\sqrt{n})$ holds globally in τ. Indeed P_b is a polynomial in μ^2 of degree $|b|$, with nonnegative coefficients since differing polynomial degrees are Fock-orthogonal. Its leading coefficient is one and the constant at least $a^{|b|}$. Matrix-entry derivatives of the polynomial on the covariance segment $I + t(B_b - I)$ have bounded coefficients by Wick expansion and bounded size. Thus the log derivatives along this segment are bounded uniformly for all $\mu \geq 0$, and the mean value estimate follows. The same reasoning controls second derivatives of $\log P_b$ and their radial derivatives on fixed compact τ-intervals.

Equation (29) and tightness force β bounded away from zero. Along any convergent subsequence $\beta \to \beta_* > 0$, the scalar reference concentrates at $(1 - 1/\beta_*)_+$ by the strictly maximized exponent $-\tau + \log(1+\beta_*\tau)$. The relative factor $e^{O(\sqrt{n})}$ leaves this concentration unchanged, since probabilities off a fixed neighborhood are exponentially small and the relative normalization is at least $e^{-C\sqrt{n}}$. The second bound in (29) therefore gives $\beta_*(1 - 1/\beta_*)_+ = 1$, so $\beta \to 2$.

We next obtain the scalar fluctuation law despite the possible $O(\sqrt{n})$ relative shift of the normalizer. Put $x_0 = 1 - 1/\beta$, $X = \sqrt{n}(\tau - x_0)$. The variable $T + na/\beta$ under the scalar law is Gamma$(n+1, 1)$ conditioned to exceed na/β, giving $X \to N(0,1)$ with all fixed moments. First variation of the block log along $I + t(B_b - I)$ at $t = 0$ is

$$f(\tau) S_b, \qquad f(\tau) = a\mu^2/d^2, \qquad S_b = \sum_{i\neq j \in b} (B_b)_{ij}.$$

This follows directly from Wick: the unique linear off-diagonal covariance pair contributes $a(B_b)_{ij}$ and a product μ^2 from the two remaining unmatched variables, with the other factors d. The remainder and its first radial derivative are $O(e_b^2)$ on a fixed interval about $1/2$, by the uniform log-polynomial derivatives. In particular

$$f'(x_0) = O(|\beta - 2| + 1/n)$$

because $f' = a\beta(a - \mu^2)/d^3$. With $\sum_b |S_b| \leq C\sqrt{n}$, Taylor on the fixed interval gives

$$|\mathcal{E}(\tau) - \mathcal{E}(x_0)| \leq C\left[(n^{-1/2} + |\beta - 2|)|X| + X^2/\sqrt{n}\right]. \tag{30}$$

Thus the relative tilt tends to one after anchoring on every fixed X-box. On a fixed central radial interval the scalar density in X is bounded by Ce^{-cX^2} by the uniform negative second derivative at its saddle and the scalar normalizer lower bound. Outside it the $e^{O(\sqrt{n})}$ relative shift preserves exponential polynomial tails. This proves $X \to N(0,1)$ with all fixed moments also under the full tilt. Then (29) gives $u_n = \sqrt{n}(\beta - 2) = O(1)$ from $\sqrt{n}(\mu^2 - a) = u_n + \beta X + O(n^{-1/2})$.

29 From variance to exact maximality

Write $Z_i = (y_i - a)/(2\sqrt{a}(y_i + a))$ using (26), and Z_i^0 evaluated at x_0. On each fixed scaled box,

$$Z = Z^0 + \frac{X}{2\sqrt{n}}\mathbf{1} + o_{\ell^2}(1), \qquad \|Z^0\| = O(1). \tag{31}$$

Indeed at $B_b = I$, $y_i = \mu^2$, and the radial derivative of Z_i tends to $1/2$. At general B_b it differs by $O(e_b)$ on the fixed central interval, since all conditional norm and inner-product ratios are bounded smooth polynomials in the block entries and μ, with denominators bounded below. Their aggregate Euclidean perturbation is $O((\sum e_b^2)^{1/2}) = O(1)$. Each fixed per-coordinate Taylor remainder is $O(X^2/n)$, and the common scalar derivative differs from $1/2$ by $o(1)$. At the center the scalar defect is $O(n^{-1/2})$ per row by bounded u_n; the e_b perturbations give the bounded offset. By the centered Gaussian limit, nonnegativity and restriction to increasing fixed boxes, (31) gives

$$\mathbb{E}\sum_i Z_i^2 \geq \tfrac{1}{4} + \|Z^0\|^2 - o(1).$$

For full rigor pass to subsequences of the bounded coefficients of the quadratic polynomial in X and let the box grow. The second terms in (26) give an additional $c\sum e_b^2 - o(1)$ because the central interval has probability tending to one and μ stays positive there.

We also have $\mathbb{E}e^*Re = 1/4 + o(1)$ for any unit top eigenvector e. In a block with no noise correlation v is the constant vector $v_0 = \mu/d$. On the central interval $v - v_0\mathbf{1}$ has norm $O(1)$ because each block discrepancy is $O(e_b)$ in fixed dimension, using the positive polynomial denominators and (29). The exponential tails and the pointwise bounds from (25) extend this to bounded second moments. Hence $M - \mathbb{E}|v_0|^2 J$ has norm $O(\sqrt{n})$ by (27), and the large top eigenvalue has eigenvector within $O(n^{-1/2})$ of the constant vector up to phase by projection of the eigenvalue equation. The bounded R permits that replacement. For the constant vector the quadratic form of R is the average of $\mathbf{1}_b^* R_b \mathbf{1}_b$. On the central interval it differs from a/d^2 by $O(n^{-1}\sum e_b) = o(1)$ because $R_b = (a/d^2)I$ at independence and its fixed-dimensional entries are Lipschitz in B_b there. The expectation tends to $1/4$.

The exact spectral deficit identity from (27), dropping the nonnegative loss $\mathbb{E}\|(I - ee^*)v\|^2$, now yields

$$\Lambda \leq n/4 + 1/2 - \|Z^0\|^2 - c\sum_b e_b^2 + o(1).$$

Consequently near-maximizers have $\sum_b e_b^2 \to 0$ and $\|Z^0\| \to 0$. With the first improvement the center vector differs by $o(1)$ in Euclidean norm from the scalar center at independence, whose norm is $|u_n|/4 + o(1)$. Hence $u_n \to 0$.

For a maximizing sequence, $\|A - A_{*,n}\|_{\mathrm{F}} \to 0$ and $\|(A - A_{*,n})\mathbf{1}\|_2 \to 0$ now follow. The block residual has Frobenius norm $(\sum_b e_b^2)^{1/2} = o(1)$; its row-sum norm is at most $\sqrt{k}$ times that norm by blockwise Cauchy–Schwarz. The scalar longitudinal row difference is $o(1)$ in Euclidean norm because $\sqrt{n}(\beta - nb_n) \to 0$. The uniform row-sum neighborhood theorem Theorem 7 then gives

exact scalar maximality modulo gauge. The scalar test is feasible, including when $B_0 = 2$, because $nb_n < 2$ eventually. Compactness of the finite union over partitions gives existence. To identify the literal maximizing matrix, for $n > 2k$ all off-block entries equal the same real $\beta/n > 0$ for any orbit matrix in the class. There are at least three blocks, and the entrywise representation $b_n w_i \bar{w}_j$, $|w_i| = 1$, forces $\beta/n = b_n$ by its modulus and equal phases across every distinct block. The off-block graph is connected, so all phases are equal. This proves the statement.

30 The loss of a fixed core

Take one fixed core B_b with all other rows independent before the common factor. In that case the constant-order loss at $\beta = 2 + u/\sqrt{n}$ is exactly $|b|/4 - \sum_{i \in b} |v_i(a = 1, \mu = 1)|^2$, in addition to $u^2/16$. The outside constant block has eigenvalue $n/4 + 1/2 - u^2/16 - |b|/4 + o(1)$, and the core Schur correction tends to $\sum |v_i(1,1)|^2$ because its cross block is $\sqrt{n - |b|}\, v_b(1,1)/2 + o(\sqrt{n})$. For a fixed nonidentity core this loss is strictly positive by (25), including singular cores. For a 2-by-2 core of correlation r, the fixed-core ratio has $|v_1| = |v_2| = |2 + r|/(4 + |r|^2 + 2 \operatorname{Re} r)$.

For example take $B = J_2 \oplus I_{n-2}$, where J_2 is the perfectly correlated pair. The cofactor mean magnitude is $3/7$ on each of the two core rows, so the additional loss is

$$\frac{1}{2} - 2\left(\frac{3}{7}\right)^2 = \frac{13}{98}.$$

The vector $e_1 - e_2$ is in the kernel of A, and the example has two rows with correlation one. The theorem admits arbitrarily many such singular blocks of bounded size before adding the common factor.

31 Exact maxima under a vanishing spectral floor

For an n-by-n Hermitian correlation matrix A, write $C(A)_{ij} = \operatorname{per} A(i \mid j)$, where the permanent omits row i and column j. Set $\Lambda(A) = \lambda_{\max}(C(A))/\operatorname{per} A$. The raw permanental adjoint was studied by Marcus and Merris [4]. Lavi used a one-row bordering calculation at fixed orders [3], and its recursive use in Lemma 71 gives an all-order bound with the correct linear and constant terms. Creation operators also enter the entry-weighted cofactor inequality of Aouichaoui and Zhang [22]. The affine Hessian and uniform local results (Theorems 1 and 7) give the equicorrelation optimizer

$$A_{*,n} = (1 - b_n)I + b_n J, \qquad b_n = \frac{2}{n+1} + O(e^{-cn}), \qquad \Lambda(A_{*,n}) = \frac{(n+1)^2}{4n} + O(ne^{-cn}).$$

The uniform local theorem asserts exact comparison and equality only on the scalar phase orbit when both $\|A - A_{*,n}\|_{\mathrm{F}}$ and $\|(A - A_{*,n})\mathbf{1}\|_2$ are small in a diagonal phase gauge. Its global spectral-floor application takes $A \succeq (1 - B^2/n)I$ for fixed B.

Theorem 9. *There is an absolute $\varepsilon > 0$ such that for all sufficiently large n the exact maximum of Λ over correlation matrices $A \succeq n^{-\varepsilon} I$ is $\Lambda(A_{*,n})$. The equality matrices are exactly the diagonal-unitary conjugates of $A_{*,n}$. In particular the same holds under any fixed spectral floor $a \in (0,1)$, for all sufficiently large n depending on a.*

We prove a quantitative localization. For each fixed C there is $K = K_C$ such that, for $0 < a \le 1$ with na sufficiently large, every correlation matrix $A \succeq aI$ with $\Lambda(A) \ge n/4 + 1/2 - C/n$ obeys, in a phase gauge,

$$\|A - A_{*,n}\|_{\mathrm{F}} + \|(A - A_{*,n})\mathbf{1}\|_2 \le \frac{C_C}{\sqrt{na^{2K+1}}}. \tag{32}$$

The exponent obtained below is qualitative. The localization uses the positive spectral floor in a lower bound on the deleted cofactor matrices.

32 Bordering estimates

Use $M = C(A)^{\mathsf{T}}/\operatorname{per} A$. These normalized matrices are positive semidefinite by Lemma 69. With $m = n - 1$, write

$$L_i = \frac{C(A_{\widehat{i}})^{\mathsf{T}}}{\operatorname{per} A_{\widehat{i}}}, \qquad \lambda_i = \lambda_{\max}(L_i), \qquad x_i = M_{ii}, \qquad z_i = A_{-i,i}, \qquad r_i = M_{-i,i}.$$

Lemma 70 gives $r_i = x_i L_i z_i$ and $z_i^* L_i z_i = (1 - x_i)/x_i$. Let c_j be the recursive upper bound of Lemma 71, with $c_j = (1 + \sqrt{1 + (j-1)c_{j-1}})/2$ for $j \geq 2$ and $c_j = j/4 + 1/2 + O(j^{-1})$ as $j \to \infty$. Set $\lambda = \lambda_{\max}(M)$, $\Delta_i = c_m - \lambda_i$. The basic row norm bound is

$$\|r_i\|_2^2 \leq \lambda_i x_i(1 - x_i) \leq c_m x_i(1 - x_i). \tag{33}$$

Since the transpose convention makes M transform by the same diagonal-unitary conjugation as A, choose a phase gauge with top unit vector e of M real and nonnegative.

Set $q_i = e_i^2$. Cauchy–Schwarz in row i gives $(\lambda - x_i)^2 q_i \leq c_m x_i(1 - x_i)(1 - q_i)$. The maximum of $x(1-x)/(\lambda - x)^2$ on $0 \leq x \leq 1$ is $1/[4\lambda(\lambda - 1)]$ when $\lambda > 1$. Consequently

$$q_i \leq \frac{c_m}{c_m + 4\lambda(\lambda - 1)} = n^{-1} + O_C(n^{-3}).$$

The expansion uses both terms of c_m and λ. Summing the excess above $1/n$ gives a total $O_C(n^{-2})$, which also bounds the sum of the negative deviations. In particular $q_i \geq 1/n - O_C(n^{-2})$ and $\|e - \mathbf{1}/\sqrt{n}\|_2^2 = O_C(n^{-3})$.

The absolute row sums s_i of M are bounded by $f(x_i) = x_i + \sqrt{(n-1)c_m x_i(1 - x_i)}$. This function has maximum c_n at $1/2 + O(n^{-1})$ and second derivative at most $-cn$ for an absolute $c > 0$. Since $\lambda \leq \sum_i s_i q_i$ by the inequality $2e_i e_j \leq q_i + q_j$, we have $\sum_i (c_n - s_i) = O_C(1)$. It follows that $\sum_i (x_i - 1/2)^2 = O_C(n^{-1})$.

Write $E_i = e_{-i}/\sqrt{1 - q_i}$ and $R_i = (\lambda - x_i)e_i/\sqrt{1 - q_i}$. The row projection onto E_i is R_i. Its summed square satisfies

$$\sum_i R_i^2 \;\geq\; (\lambda - 1/2)^2 \frac{n}{n-1} - O_C(1) = \frac{n^2 + n}{16} - O_C(1).$$

Indeed $\sum_i |x_i - 1/2| = O_C(1)$, $q_i/(1 - q_i) = O_C(n^{-1})$, and $\sum_i q_i/(1 - q_i) \geq n/(n-1)$ by Jensen. Meanwhile $\sum_i c_m x_i(1 - x_i) \leq n c_m/4 = (n^2 + n)/16 + O(1)$. By (33),

$$\sum_i \left[c_m x_i(1 - x_i) - \|r_i\|^2\right] + \sum_i \|r_i - R_i E_i\|^2 = O_C(1). \tag{34}$$

Each summand is nonnegative, and the first sum bounds $\sum_i \Delta_i x_i(1 - x_i)$. Thus $\sum_i \Delta_i = O_C(1)$.

33 A lower cofactor bound from independent noise

Lemma 10. *Let $T \succeq aI$ be a correlation matrix, $a > 0$, and put $L = C(T)^{\mathsf{T}}/\operatorname{per} T$, $y_j = L_{jj}$. Then*

$$L \succeq a \operatorname{diag}(y_j^2).$$

Proof. Use the circular complex Gaussian moment representation of the permanent [17, Theorem A.4]. Write the Gaussian forms as $\mu_j + \sqrt{a}Z_j$, where μ has covariance $T - aI$ and the Z_j are independent standard circular variables, independent of μ. Put $d_j = a + |\mu_j|^2$ and $P = \prod_j d_j$. Conditionally on μ, the permanent integrand is P. The transposed cofactor Gram matrix divided by P equals

$$vv^* + a\,\mathrm{diag}(d_j^{-2}), \qquad v_j = \mu_j/d_j.$$

Its diagonal is $1/d_j$; opposite conjugation conventions give the transpose and preserve the inequality. Thus L is the tilted expectation of this matrix with weight $P/\mathbb{E}P$, and y_j is the tilted expectation of $1/d_j$. Drop vv^* and apply Jensen. □

Fix a sufficiently small absolute $\delta_0 > 0$. Let $G = \{i : \Delta_i \le \delta_0\}$. The complement has at most K_C indices by (34). Apply the eigenvector cap above to the deleted matrix L_i of order m, for $i \in G$. For its top unit vector the cap becomes $1/m + O(\Delta_i/m^2 + m^{-3})$ on the maximal squared component. Normalization then bounds every squared component below by $1/(2m)$ for small fixed δ_0. Applying the same concave row-sum bound at order m gives

$$\sum_{j\neq i} \big((L_i)_{jj} - 1/2\big)^2 \le C_0\Delta_i + O(m^{-1}).$$

In detail, the sum of row-sum deficiencies is at most $2m\Delta_i$, and the curvature of the row-bound function is at least of order m. With a smaller fixed δ_0 the diagonal is therefore at least $1/4$ for every good deletion. Lemma 10 gives $L_i \succeq (a/16)I$ on these minors.

For every i, $\Delta_i = O_C(1)$. Bounding each off-diagonal squared row norm of L_i by $c_{m-1}/4$ and each squared diagonal by 1 gives $\|L_i\|_{\mathrm{F}}^2 \le m^2/16 + O(m)$. Since $\lambda_i^2 = m^2/16 + O_C(m)$, the other eigenvalues are $O_C(\sqrt{n})$.

34 Converting bordering loss into matrix alignment

Let I_0 be any set of indices with $L_i \succeq tI$ for $i \in I_0$, $0 < t \le 1$. We show

$$\sum_{i\in I_0} \big(\|z_i - b_n\mathbf{1}\|_2^2 + |\mathbf{1}^*(z_i - b_n\mathbf{1})|^2\big) \le \frac{C_C}{nt}. \tag{35}$$

For each $i \in I_0$, choose the phase of the top unit vector u_i of L_i so that $z_i = \alpha_i u_i + v_i$, $\alpha_i \ge 0$, $v_i \perp u_i$. The deficit

$$D_i = \lambda_i x_i(1 - x_i) - \|r_i\|^2 = x_i^2\langle v_i, L_i(\lambda_i I - L_i)v_i\rangle$$

has total $O_C(1)$ by (34). The transverse eigenvalues lie between t and $O_C(\sqrt{n})$, while x_i stays bounded away from zero. Thus

$$\sum_{i\in I_0} \|v_i\|^2 \le C_C/(nt).$$

Put $r_i^{\perp} = x_i L_i v_i$. Since $\mu^2 \le O_C(n^{-1/2})\mu(\lambda_i - \mu)$ on the transverse spectrum, $\sum_{i\in I_0} \|r_i^{\perp}\|^2 = O_C(n^{-1/2})$.

Put $B_i = x_i\lambda_i\alpha_i$ and $d_i = r_i - R_iE_i$, where $d_i \perp E_i$. As $r_i = B_iu_i + r_i^{\perp}$, we have

$$B_i^2 = R_i^2 + \|d_i\|^2 - \|r_i^{\perp}\|^2.$$

The positive R_i are uniformly comparable to $\sqrt{n}$. It follows that B_i are likewise comparable, and $|B_i - R_i| \le C(\|d_i\|^2 + \|r_i^\perp\|^2)/\sqrt{n}$. From $B_i(u_i - E_i) = d_i - r_i^\perp + (R_i - B_i)E_i$,

$$\sum_{i\in I_0} \|u_i - E_i\|^2 = O_C(n^{-1}), \qquad \sum_{i\in I_0} \|\mathbf{1} - \sqrt{m}\, u_i\|^2 = O_C(1). \tag{36}$$

The second estimate uses $\|E_i - \mathbf{1}/\sqrt{m}\| = O_C(n^{-3/2})$ from the parent vector.

We also have

$$\sqrt{m}\,\alpha_i = \frac{\sqrt{m}\, e_i(\lambda - x_i)}{x_i\lambda_i\sqrt{1-e_i^2}} + O_C\!\left(\frac{\|d_i\|^2 + \|r_i^\perp\|^2}{n}\right).$$

The leading factor differs from 2 by squared error summed over I_0 of order $1/n$. This follows from $\sum_i (x_i - 1/2)^2 = O_C(n^{-1})$, $\sum_i (\sqrt{m}e_i - 1)^2 = O_C(n^{-1})$, $(\lambda - x_i)/\lambda_i = 1 + O_C(n^{-1})$, and $(1-e_i^2)^{-1/2} = 1 + O_C(n^{-1})$. The last displayed error has squared sum $O_C(n^{-2})$. Thus $\sum_{i\in I_0} |\sqrt{m}\,\alpha_i - m b_n|^2 = O_C(n^{-1})$ and $\alpha_i = O_C(n^{-1/2})$ uniformly. Decompose

$$z_i - b_n\mathbf{1} = v_i + \alpha_i(u_i - \mathbf{1}/\sqrt{m}) + (\alpha_i - \sqrt{m}\, b_n)\mathbf{1}/\sqrt{m}.$$

This bounds the sum of squared column errors by $C_C/(nt)$. For the column sums, orthogonality gives $|\mathbf{1}^* v_i| \le \|\mathbf{1} - \sqrt{m}u_i\|\|v_i\|$. The maximum of the first squared factor is $O_C(1)$ by (36). For the middle term use $|\mathbf{1}^* u_i - \sqrt{m}|^2 \le \|\mathbf{1} - \sqrt{m}u_i\|^2$ and $\alpha_i^2 = O_C(n^{-1})$. The final term is the scalar error just estimated. This proves (35).

35 Exceptional principal block and the second pass

Apply (35) to G with $t = a/16$. It follows that the good principal matrix A_G has Frobenius error $O_C((na)^{-1/2})$ from its corresponding equicorrelation block, and hence $A_G \succeq I/2$ once na is sufficiently large. For any exceptional j, $\|A_{G,j}\|^2 = O_C((na)^{-1})$, by the same estimate and the squared size $O(n^{-1})$ of a scalar column on G.

Write $P_S = \operatorname{per} A_S$ for a principal index set S. Realize A as the Gram matrix of unit vectors v_i; P_S is the squared complex Gaussian Fock norm of the product of the corresponding forms. Let W be the span of the good vectors and decompose each exceptional vector as $v_j = w_j + w_j^\perp$, $w_j \in W$. The projection formula and $A_G \succeq I/2$ give $\|w_j\|^2 \le 2\|A_{G,j}\|^2 = O_C((na)^{-1})$.

Start from the product of good forms. After multiplying by k exceptional forms, the W-degree is at most $|G| + k$ and the $W^\perp$-degree at most k. Multiplication by a linear form of vector norm b costs at most $b\sqrt{q+1}$ on degree q of its coordinate subspace. This follows by unitary rotation to one coordinate and the normalized Gaussian monomial basis, and holds with the maximal bound on the indicated graded subspaces. Thus each exceptional multiplication has norm at most $C_C a^{-1/2}$ on the stages used here: the W part costs $O_C(a^{-1/2})$ and the orthogonal part costs at most $\sqrt{k+1}$. Since there are at most K exceptional factors, for $J' \subset G^c$,

$$P_G \le P_{G\cup J'} \le C_C a^{-K} P_G. \tag{37}$$

The lower bound uses the creation-annihilation identity $\|LF\|^2 = \|L^*F\|^2 + \|F\|^2$ for multiplication by a unit-norm Gaussian form.

For a pair $i, j \notin G$, the diagonal $(L_i)_{jj}$ is the ratio of the two permanents deleting i, j and deleting i, respectively. Both index sets contain G. Equation (37) bounds their ratio below by $c_C a^K$. If $i \in G$, the bound already holds by the good-minor argument. If $j \in G$, the deleted diagonal ratio is $(x_j/x_i)(L_j)_{ii}$, with x_j/x_i bounded below. This proves $(L_i)_{jj} \ge c_C a^K$ for every

pair. Lemma 10 now yields $L_i \succeq c_C a^{2K+1} I$ for every deletion. Apply (35) to all indices with this lower bound. Hermitian symmetry makes the squared column-sum norm equal to the squared row-sum norm, proving (32).

Finally, let C be absolute so the equicorrelation quotient satisfies the required lower bound, and choose $0 < \varepsilon < 1/(2K_C + 1)$. For $a = n^{-\varepsilon}$ the scalar candidate is feasible, and $na^{2K_C+1} \to \infty$. The feasible set is compact and Λ continuous, since the permanent is positive. Every maximizer satisfies the near-optimality threshold, so (32) places it in the uniform neighborhood of Theorem 7. Its exact comparison and equality statement give Theorem 9. A fixed positive floor eventually defines a subclass containing the comparator.

36 The unrestricted maximum in large dimensions

Let A be a complex Hermitian positive semidefinite matrix with unit diagonal, of order n. Write $C(A)_{ij} = \operatorname{per} A(i \mid j)$, where $A(i \mid j)$ deletes row i and column j. We use the transposed cofactor matrix

$$H(A) = C(A)^{\mathsf{T}} / \operatorname{per} A, \qquad \Lambda(A) = \lambda_{\max}(H(A)), \qquad \gamma_n = \max_A \Lambda(A).$$

The permanent is positive on correlation matrices, including singular ones. Indeed, the squared Fischer norm of a product of nonzero Gram linear forms is the permanent, and the deleted products also give positivity of H. Compactness gives the maximum. For diagonal unitary D, $H(D^*AD) = D^*H(A)D$.

The isolated-row argument of Lavi [3, Section 4], applied recursively in Lemma 71, gives the upper bound $\gamma_n \le c_n \le n/4 + 1/2 + 3/(4n)$. The scalar equicorrelation family gives the lower bound $n/4+1/2+1/(4n)$. The classical estimate of Marcus and Merris [4] uses the input matrix as a metric. The question here concerns the largest Euclidean eigenvalue of the unweighted cofactor matrix.

Let $A_{*,n} = (1 - b_n)I + b_n J$ be the positive scalar optimizer, where $J = \mathbf{1}\mathbf{1}^*$ and $b_n = 2/(n + 1) + O(e^{-an})$ for some $a > 0$ Proposition 2.

Theorem 11. *The unrestricted maximum satisfies*

$$\gamma_n = \frac{n}{4} + \frac{1}{2} + \frac{1}{4n} + o(n^{-1}).$$

For a sequence with $\gamma_n - \Lambda(A) = o(n^{-1})$ there are diagonal unitary gauges in which

$$\|A - A_{*,n}\|_{\mathrm{F}} + \|(A - A_{*,n})\mathbf{1}\|_2 = o(1).$$

For every fixed B, the same two distances are $O_B(1)$ whenever $n(c_n - \Lambda(A)) \le B$, in gauges making a top cofactor eigenvector positive.

Theorem 12. *For all sufficiently large n, $\gamma_n = \Lambda(A_{*,n})$, and every maximizer is a diagonal unitary conjugate of $A_{*,n}$. In particular the error in the expansion in Theorem 11 is $O(ne^{-a'n})$ for some $a' > 0$. There are absolute $\eta > 0$ and $K < \infty$ such that, for all sufficiently large n, every A with $\gamma_n - \Lambda(A) \le \eta/n$ has a gauge with*

$$\|A - A_{*,n}\|_{\mathrm{F}} + \|(A - A_{*,n})\mathbf{1}\|_2 \le K\sqrt{\gamma_n - \Lambda(A)}.$$

Theorem 11 uses only permanental bordering and scalar inequalities. Theorem 12 additionally uses the uniform local comparison of Theorem 7, including its quadratic estimate. Order three is settled exactly in Theorem 73, order four exactly in Section 62, and order five is treated in Section 66.

37 Control by the recursive deficit

For a correlation matrix T of order m, let T_i be the principal deletion of i, T_{ij} the deletion of i, j, and $z_{j|i} = T_{[m]\setminus\{i,j\},j}$. Write

$$\lambda = \Lambda(T), \quad d = c_m - \lambda, \quad \kappa = md, \quad H = H(T), \quad x_i = H_{ii},$$
$$H_i = H(T_i), \quad \lambda_i = \Lambda(T_i), \quad d_i = c_{m-1} - \lambda_i.$$

Use $z_i = T_{[m]\setminus\{i\},i}$ for the full bordering column at i, and $r_i = H_{-i,i}$ for the cofactor column off the diagonal. Lemma 70 gives

$$r_i = x_i H_i z_i, \qquad z_i^* H_i z_i = (1 - x_i)/x_i, \qquad 0 < x_i \le 1. \tag{38}$$

In particular, with $c = c_{m-1}$,

$$\|r_i\|_2^2 \le \lambda_i x_i(1 - x_i) \le c x_i(1 - x_i), \qquad D_i = \lambda_i x_i(1 - x_i) - \|r_i\|_2^2 = x_i^2 z_i^* H_i(\lambda_i I - H_i) z_i. \tag{39}$$

The recursive envelope and the row bound are

$$c_1 = 1, \quad c_m = \tfrac{1}{2}(1 + \sqrt{1 + (m-1)c}), \quad f_m(x) = x + \sqrt{(m-1)c\,x(1-x)}.$$

The function f_m has maximum c_m at

$$X_m = \frac{c_m}{2c_m - 1}, \qquad h_m = \frac{c_m - X_m}{m-1} = \sqrt{\frac{cX_m(1-X_m)}{m-1}}, \qquad H_m^\circ = (X_m - h_m)I + h_m J.$$

The explicit upper bound $u_m = m/4 + 1/2 + 3/(4m)$ of Lemma 71 obeys $\frac{1}{2}(1 + \sqrt{1 + (m-1)u_{m-1}}) = u_m + O(m^{-3})$. On the interval between c_{m-1} and u_{m-1}, the derivative of the recursion is $1/2 + O(m^{-1})$, by the lower bound of Lemma 71. Iterating this contraction gives $u_m - c_m = O(m^{-3})$. Consequently

$$X_m = \tfrac{1}{2} + \tfrac{1}{m} + O(m^{-2}), \qquad h_m = \tfrac{1}{4} + \tfrac{1}{4m} + O(m^{-2}). \tag{40}$$

Denote $U(T) = \sum_i (x_i - X_m)$ and by T' a uniformly chosen principal deletion.

Lemma 13. *There is an absolute K such that for all $m \ge 2$,*

$$\sum_i (x_i - X_m)^2 \le Kd, \quad |U(T)|^2 \le K\kappa, \quad \sum_i (d_i + D_i) \le K\kappa, \quad \mathbb{E}[\kappa(T') \mid T] \le K\kappa.$$

For $d \le d_0$, with $d_0 > 0$ absolute and sufficiently small, let e be any top unit eigenvector. Then

$$\|H - [(X_m - h_m)I + m h_m e e^*]\|_{\mathrm{F}}^2 \le K\kappa.$$

In a positive gauge for e one also has $\|e - \mathbf{1}/\sqrt{m}\|_2^2 \le Kd^2/m$ and $\|H - H_m^\circ\|_{\mathrm{F}}^2 \le K\kappa$.

Proof. Write $q_i = |e_i|^2$. When $d \le c_m/2$ and m is large, the row eigenvalue equation and (39) give

$$(\lambda - x_i)^2 q_i \le c x_i(1 - x_i)(1 - q_i), \qquad q_i \le Q = \frac{c}{c + 4\lambda(\lambda - 1)}.$$

Here the maximum of $x(1-x)/(\lambda - x)^2$ is $1/[4\lambda(\lambda - 1)]$. At $d = 0$ the cap is exactly $1/m$, since $4c_m(c_m - 1) = (m-1)c$. The mean value theorem gives $Q - 1/m \le Kd/m^2$ on this range. Hence the total positive deviation of q_i from $1/m$ is at most Kd/m, and at most Kd indices have $q_i < 1/(2m)$.

Let $a_i = x_i + \|r_i\|_1$. The Rayleigh quotient gives $\lambda \le \sum_i q_i a_i \le c_m$. Since $f_m'' \le -2\sqrt{(m-1)c}$, we have $c_m - a_i \ge c_0 m (x_i - X_m)^2$. Sum this on the indices with $q_i \ge 1/(2m)$ and bound the other $O(d)$ terms by one. This gives the diagonal and U estimates.

For each i, the squared projection of r_i onto e_{-i} is $P_i = (\lambda - x_i)^2 q_i/(1-q_i)$. Because $q_i \le Q$, the same scalar optimization gives

$$0 \le c x_i(1-x_i) - P_i \le c x_i (1-x_i) \frac{1 - q_i/Q}{1-q_i}.$$

The sum is $O(md)$, since $m - 1/Q = O(d)$ and $Q = O(1/m)$. It bounds both the total $D_i + d_i x_i(1-x_i)$ and the total squared residual from these projections. By the diagonal estimate only $O(d)$ indices have $x_i \notin [1/4, 3/4]$. Their d_i's are at most $c = O(m)$. This proves the gap estimates. When $d > c_m/2$, md has order m^2, and the first claims follow from $x_i \in [0,1]$, $d_i \le c$ and $D_i \le c/4$. A larger absolute constant includes the finitely many small orders, where the same cap proof applies near $d = 0$ for $m \ge 2$ since $c_m > 1$.

For $d \le d_0$ all $q_i \ge 1/(2m)$. In the positive gauge, the bounds $\sum_i |q_i - 1/m| = O(d/m)$ and $\max_i |q_i - 1/m| = O(d/m)$ give $\|e - \mathbf{1}/\sqrt{m}\|_2^2 = O(d^2/m)$. On column i the projection is $e_{-i}\bar{e}_i(\lambda - x_i)/(1-q_i)$. The reference multiplier $m h_m = (c_m - X_m)/(1 - 1/m)$ differs from this multiplier by $O(d + |x_i - X_m| + m|q_i - 1/m|)$. The sum of squared column errors is bounded by the sum of the squared multiplier errors times q_i, hence is $O(md)$. On the diagonal the difference is $x_i - X_m - h_m(mq_i - 1)$, with squared sum $O(md)$. This proves the rank-one comparison. Replacing ee^* by J/m in the positive gauge costs $O(md^2)$ in squared Frobenius norm. □

38 Two overlapping deletions

All estimates in this section are uniform for $\kappa(T) \le B$, with B fixed. Work in the positive cofactor gauge of $T = (t_{ij})$. Set $y_{ij} = (H_i)_{jj}$ and $p_{ij} = \operatorname{per} T_{ij}/\operatorname{per} T$. For distinct i, j put $B_{ij} = H(T_{ij})$, $v_{j|i} = B_{ij}^{1/2} z_{j|i}$ and $v_{i|j} = B_{ij}^{1/2} z_{i|j}$, using the same ordering of the remaining coordinates. In addition to (38), direct permanent expansion gives

$$p_{ij} = x_i y_{ij} = x_j y_{ji}, \qquad H_{ij} = p_{ij}\big(t_{ij} + v_{i|j}^* v_{j|i}\big), \qquad \|v_{j|i}\|_2^2 = \frac{1 - y_{ij}}{y_{ij}}, \tag{41}$$

In the cofactor $C(T)_{ji}$ the surviving row i can use column j directly, giving the first summand inside parentheses. If it uses a column in the common minor, another row uses column j. Summing these terms gives $z_{i|j}^* B_{ij} z_{j|i}$, with the transpose convention of (38).

Lemma 14. *For fixed B, every T with $\kappa(T) \le B$ satisfies*

$$\|T - I\|_{\mathrm{F}} = O_B(1), \qquad \|(T-I)\mathbf{1} - 2\mathbf{1}\|_2 = O_B(1), \qquad \mathbb{E}[U(T') \mid T] = U(T) + \tfrac{1}{2} + o_B(1).$$

Proof. Choose a small absolute $\delta > 0$ and restrict initially to $G = \{i : d_i \le \delta\}$. Its complement has size $O_B(1)$ by Lemma 13. On these children all $y_{ij} \in [1/4, 3/4]$ for large m, and

$$\sum_{i \in G} \sum_{j \ne i} |y_{ij} - X_{m-1}|^2 = O_B(1). \tag{42}$$

This follows from the diagonal estimate by choosing δ small. Also $x_i = 1/2 + O_B(m^{-1/2})$ for every i.

For each $i \in G$, $\lambda_i = m/4 + O_B(1)$ and $\lambda_2(H_i) = O_B(\sqrt{m})$. Indeed (39) applied to H_i bounds its squared Frobenius norm by $m^2/16 + O(m)$, and subtracting λ_i^2 gives the claim. Let u_i be a top unit eigenvector of H_i, with phase chosen so that its inner product with $\mathbf{1}/\sqrt{m-1}$ is nonnegative. For non-top eigenvalues μ, $\mu^2 \le [\lambda_2(H_i)/(\lambda_i - \lambda_2(H_i))]\mu(\lambda_i - \mu)$, so (39) controls the squared orthogonal projection of r_i by $O(D_i)$. Since $\sum_i \|r_i - h_m\mathbf{1}\|^2 = O_B(1)$, projecting $h_m\mathbf{1}$ off each top direction gives

$$\sum_{i \in G} \|u_i - \mathbf{1}/\sqrt{m-1}\|_2^2 = O_B(m^{-1}).$$

Here a unit vector with nonnegative overlap has squared distance at most twice its squared orthogonal component. The rank-one comparison in Lemma 13, with this angle estimate, gives in the inherited gauge

$$\sum_{i \in G} \|H_i - H^{\circ}_{m-1}\|_{\mathrm{F}}^2 = O_B(m). \tag{43}$$

For $i, j \in G$, $i \ne j$, let $d_{ij} = c_{m-2} - \Lambda(T_{ij})$. The global gap bound on T_i gives $d_{ij} \le K(m-1)d_i$. With δ sufficiently small this places the top eigenvalue of B_{ij} above $m/5$, with its gap to the second at least $c_1 m$ for some $c_1 > 0$. The latter follows as before from $\|B_{ij}\|_{\mathrm{F}}^2 \le m^2/16 + O(m)$, since $2/25 > 1/16$. Summing child reserves $D_j(T_i)$ over $i \in G$ gives $O_B(m)$. Formula (39) and $y_{ij} \ge 1/4$ therefore bound the sum of squared non-top projections of $v_{j|i}$ by $O_B(1)$.

For the top projections compare the cofactor columns in the same coordinates. By (38), $B_{ij}z_{j|i}$ is the off-diagonal column of H_i at j, divided by y_{ij}. Equations (42) and (43) imply that the squared errors in comparison with $(h_{m-1}/X_{m-1})\mathbf{1}$ sum to $O_B(m)$. The same holds in the reverse direction. On the top eigenspace of B_{ij} the squared norm of the difference of the two v's is the squared norm of the corresponding projection of $B_{ij}(z_{j|i} - z_{i|j})$ divided by the top eigenvalue. Thus

$$\sum_{\substack{i,j \in G \\ i \ne j}} \|v_{j|i} - v_{i|j}\|_2^2 = O_B(1). \tag{44}$$

All pairs in sums are ordered. This estimate uses the weighted vectors $B_{ij}^{1/2}z$, so it holds at singular double minors as well.

By (41), the v's have uniformly bounded norm on these pairs. The difference $v_{i|j}^* v_{j|i} - 1$ is bounded in modulus by a constant times $\|v_{j|i} - v_{i|j}\|_2 + |y_{ij} - 1/2|$. Its squared sum is bounded by (42) and (44). Similarly p_{ij} is bounded below and its squared deviations from $1/4$ have bounded sum, using $p_{ij} = x_i y_{ij}$. Lemma 13 bounds the squared deviations of H_{ij} from $1/4$ on all off-diagonal pairs, since $h_m - 1/4 = O(m^{-1})$. The pair identity gives $\|T_G - I\|_{\mathrm{F}} = O_B(1)$. For an omitted index i, positivity of T implies $T_{G,i}T_{G,i}^* \preceq T_G$, so $\|T_{G,i}\|_2^2 \le \|T_G\|_{\mathrm{op}}$. The bounded number of omitted indices then proves $\|T - I\|_{\mathrm{F}} = O_B(1)$ on the full matrix.

The exact Laplace row identity is

$$x_i + \sum_{j \ne i} t_{ij}\,\overline{H_{ij}} = 1. \tag{45}$$

Replace $\overline{H_{ij}}$ by h_m in this identity. The error vector has ℓ^1 norm at most $\|T - I\|_{\mathrm{F}}\|H - H_m^{\circ}\|_{\mathrm{F}}$. Also $\sum_i |x_i - 1/2| = O_B(1)$. With $h_m = 1/4 + O(m^{-1})$, this yields both the asserted row-sum norm estimate and

$$\sum_{i \ne j} t_{ij} = 2m + O_B(1). \tag{46}$$

To extract the drift, take real parts in (41) and use the norms of both weighted vectors:

$$p_{ij}\,\mathrm{Re}\,t_{ij} = \mathrm{Re}\,H_{ij} - \tfrac{1}{2}(x_i + x_j) + p_{ij} + \tfrac{1}{2}p_{ij}\|v_{j|i} - v_{i|j}\|_2^2. \tag{47}$$

Write $X = X_m$, $Y = X_{m-1}$, $h = h_m$, $\xi_i = x_i - X$, and $\zeta_{ij} = y_{ij} - Y$. Sum (47) over distinct $i, j \in G$. There are $O_B(m)$ omitted pairs. By (46) and the full Frobenius bound, the sum of $\mathrm{Re}\,t_{ij}$ on the retained pairs is $2m + O_B(\sqrt{m})$. Replacing p_{ij} by $XY = 1/4 + O(1/m)$ costs $O_B(1)$ by Cauchy–Schwarz and the squared deviation bound above. Thus the left side sums to $m/2 + O_B(\sqrt{m})$.

On the right the weighted pair energy contributes $O_B(1)$. The sum of $\mathrm{Re}(H_{ij} - h)$ over all off-diagonal pairs is $O_B(1)$: the test on $\mathbf{1}/\sqrt{m}$ differs from the top Rayleigh quotient by at most $\lambda\|e - \mathbf{1}/\sqrt{m}\|_2^2$, and the trace differs from mX by $U(T) = O_B(1)$. Removing the omitted pairs costs $O_B(\sqrt{m})$. In $p_{ij} = (X+\xi_i)(Y+\zeta_{ij})$ the sum of cross terms is $O_B(1)$ by (42) and $m\sum_i |\xi_i|^2 = O_B(1)$. Therefore, writing $g = |G|$, the right side equals

$$g(g-1)(h - X + XY) - (g-1)(1-Y)\sum_{i\in G}\xi_i + X\sum_{\substack{i,j\in G\\ i\neq j}}\zeta_{ij} + O_B(\sqrt{m}).$$

By (40), $h - X + XY = 1/(4m) + O(m^{-2})$. The sum of ξ_i over the omitted indices is $O_B(m^{-1/2})$. The sum of ζ_{ij} over good children but omitted coordinates is $O_B(\sqrt{m})$ by (42). Finally the omitted children contribute $O_B(\sqrt{m})$ to $\sum_i U(T_i)$, since each $d_i = O_B(1)$ and Lemma 13 gives $|U(T_i)| = O_B(\sqrt{m})$. Divide by m to obtain

$$\tfrac{1}{2} = \tfrac{1}{4} - \tfrac{1}{2}U(T) + \tfrac{1}{2}\mathbb{E}[U(T') \mid T] + o_B(1).$$

This gives the drift. □

39 Iteration and the sharp third term

Use a positive cofactor gauge whenever $d \le d_0$. Define the nonnegative residual

$$\mathcal{R}(T) = m\left(\sum_i q_i L_i + \sum_i q_i a_i - \lambda\right) + \frac{m}{4}\sum_i (x_i - \bar{x})^2, \qquad L_i = \sqrt{(m-1)\lambda_i x_i(1-x_i)} - \|r_i\|_1, \tag{48}$$

where $\bar{x} = m^{-1}\sum_i x_i$. The formula uses the squared moduli of a top unit eigenvector and also defines a nonnegative residual outside this range. By the row inequality $\mathcal{R}(T) \le K\kappa(T)$ globally: the terms in parentheses are bounded by d, and Lemma 13 controls the variance. In the bounded-κ range,

$$\kappa(T) \ge \frac{1}{2}U(T)^2 + \frac{1}{2}\mathbb{E}[\kappa(T') \mid T] + \mathcal{R}(T) - o_B(1). \tag{49}$$

Indeed when $\kappa \le B$, $mq_i = 1 + O_B(m^{-1})$ uniformly. The scalar concavity gives $c_m - f_m(x_i) \ge \sqrt{(m-1)c}\,(x_i - X_m)^2$. After weighting by mq_i, the sum is at least $(1/2 - o_B(1))m\sum_i (x_i - X_m)^2$. Split the latter into $U(T)^2$ and $m\sum_i (x_i - \bar{x})^2$. The other scalar loss is

$$\sqrt{(m-1)x_i(1-x_i)}(\sqrt{c} - \sqrt{\lambda_i}) = \big(\tfrac{1}{2} + o_B(1)\big)d_i$$

uniformly, using $x_i = 1/2 + O_B(m^{-1/2})$, $d_i = O_B(1)$ and $c \sim m/4$. Its sum gives the child term since $\mathbb{E}[\kappa(T') \mid T] = \frac{m-1}{m}\sum_i d_i$. The residual terms are the remaining losses in the weighted Rayleigh bound.

Lemma 15. *Uniformly for $\kappa(T) \le B$ as $m \to \infty$,*

$$\kappa(T) \ge \tfrac{1}{2} + (U(T) + \tfrac{1}{2})^2 + \mathcal{R}(T) - o_B(1).$$

Proof. Start at a sequence of matrices with $\kappa_0 \le B$ and $U_0 \to u$ along a subsequence. Delete one uniformly chosen index at each step, with U_k, κ_k the statistics at depth k. For every fixed k, Lemma 13 gives bounded $\mathbb{E}\kappa_k$ and $\mathbb{E}U_k^2$. The drift in Lemma 14 on parents with $\kappa_k \le L$ has a uniform error $o_L(1)$. The contribution from $\kappa_k > L$ to the U_k and U_{k+1} means is $O_{B,k}(L^{-1/2})$ by Markov and Cauchy–Schwarz. Taking $m \to \infty$, then $L \to \infty$, shows $\mathbb{E}U_k \to u + k/2$.

For fixed depth s, keep (49) only on paths with $\kappa_j \le L$ at every step through the parent. The excluded probability through depth k is $O_{B,k}(L^{-1})$ and the excluded U_k mean is $O_{B,k}(L^{-1/2})$, by the same global estimates. Iterate (49), retaining $\mathcal{R}(T)$ at the root and dropping subsequent nonnegative residuals. At each step the nonnegative next κ term can be restricted to the surviving paths. Jensen or Cauchy–Schwarz bounds the retained second moments below by the square of the retained mean. Taking the two limits as above gives

$$\liminf_{m\to\infty}(\kappa_0 - \mathcal{R}(T)) \ge \sum_{k=0}^{s-1} 2^{-k-1}(u + k/2)^2.$$

As $s \to \infty$ the right side tends to $(u + 1/2)^2 + 1/2$, since the probabilities 2^{-k-1} have mean 1 and variance 2 in k. Every bounded sequence of initial U_0's has a convergent subsequence. This proves the stated uniform error by contradiction. □

For maximizers the scalar lower bound of Lemma 71 and the expansion of c_n give $\kappa \le 1/2 + O(n^{-2})$. Lemma 15 gives $\kappa \ge 1/2 - o(1)$, which proves the sharp coefficient in Theorem 11.

40 Alignment at the sharp value

Consider a sequence with $\kappa(T) = 1/2 + o(1)$, equivalently $\gamma_m - \Lambda(T) = o(m^{-1})$. Lemma 15 gives $U(T) \to -1/2$ and $\mathcal{R}(T) \to 0$. The budget and drift give $\limsup \mathbb{E}\kappa(T') \le 3/4$ and $\mathbb{E}U(T') \to 0$. In fact

$$\mathbb{E}\kappa(T') \to 3/4, \qquad \mathbb{E}U(T')^2 + \mathbb{E}\mathcal{R}(T') \to 0, \qquad \lim_{L\to\infty} \limsup_{m\to\infty} \mathbb{E}[\kappa(T'); \kappa(T') > L] = 0. \tag{50}$$

To see this, apply the profile to children with $\kappa(T') \le L$, in the form $\kappa(T') \ge 3/4 + U(T') + U(T')^2 + \mathcal{R}(T') - o_L(1)$. Markov and the global U^2 bound make the discarded probability and mean $O(L^{-1})$ and $O(L^{-1/2})$. Keep the full $\kappa(T')$ contribution on the discarded event. The upper bound $3/4 + o(1)$ then forces the last bound in (50) and vanishing retained nonnegative terms as $L \to \infty$. The global bounds $U^2 + \mathcal{R} \le K\kappa$ control the discarded parts. This also gives the lower bound on $\mathbb{E}\kappa(T')$.

In particular $\max_i d_i = o(1)$. Otherwise a child with d_i bounded below would contribute a positive amount $(m-1)d_i/m$ to the tail expectation in (50). Thus eventually every child belongs to the set G used in Lemma 14. Lemma 13 applied to each child now also gives $\max_{i,j} d_{ij} = o(m)$.

We next check the cofactor losses used in the pair argument. In a positive gauge with $\kappa \le B$, one has

$$\|\operatorname{off}(H - H_m^\circ)\|_{\mathrm{F}}^2 \le K\mathcal{R}(T) + O_B(m^{-1}), \qquad \sum_i D_i \le K\mathcal{R}(T). \tag{51}$$

For the second inequality L_i bounds $\sqrt{m-1}(\sqrt{\lambda_i x_i(1-x_i)} - \|r_i\|_2)$, which dominates D_i up to an absolute constant. Since $mq_i \ge 1/2$, the residual controls their sum. For the first inequality, set

$s_i = \|r_i\|_1/(m-1)$ and $s_i^* = \sqrt{\lambda_i x_i(1-x_i)/(m-1)}$. The sum of squared deviations of $|H_{ji}|$ from s_i on column i equals $\|r_i\|_2^2 - (m-1)s_i^2 \le KL_i$. Also $(m-1)(s_i - s_i^*)^2 = L_i^2/(m-1) \le KL_i$, and

$$|s_i^* - h_m| \le K\big(d_i/m + |x_i - X_m|(|x_i - X_m| + m^{-1})\big).$$

The last estimate follows by comparing the square roots near $1/4$, using $x_i(1-x_i) - X_m(1-X_m) = (x_i - X_m)(1 - x_i - X_m)$. Lemma 13 shows that $\sum_i (m-1)|s_i^* - h_m|^2 = O_B(m^{-1})$. Moreover

$$\sum_i q_i a_i - \lambda \;\ge\; \sum_{i,j} e_i e_j \big(|H_{ij}| - \operatorname{Re} H_{ij}\big).$$

This follows by $2e_ie_j \le q_i + q_j$. Since $e_ie_j \ge 1/(2m)$ and $|H_{ij}| \le 1$, the residual bounds the sum of squared differences between H_{ij} and $|H_{ij}|$. These estimates prove (51).

At the root, (51) gives $\|\operatorname{off}(H - H_m^\circ)\|_{\mathrm{F}}^2 = o(1)$ and $\sum_i D_i = o(1)$. The variance term in (48) gives $m\sum_i (x_i - \bar{x})^2 = o(1)$. In the children, (50) and the same variance term imply

$$\sum_i \sum_{j \ne i} |y_{ij} - X_{m-1}|^2 = o(1). \tag{52}$$

Indeed for each child the sum equals its variance sum plus $U(T')^2/(m-1)$. The average of $(m-1)$ times the variance sum is bounded by $4\mathbb{E}\mathcal{R}(T')$.

In the children's own positive gauges the summed squared cofactor errors from H_{m-1}° are $o(m)$. On $\kappa(T') \le L$ this follows from (51), (52) and (50). On the remaining children the total is bounded by a constant times $m\mathbb{E}[\kappa(T'); \kappa(T') > L]$ by Lemma 13, since $\max_i d_i = o(1)$. The summed child reserves $\sum_i \sum_{j \ne i} D_j(T_i)$ are likewise $o(m)$, using (51) on bounded $\kappa(T')$ and the global gap bound on the tails.

To transfer the cofactor comparison into the parent gauge, the top direction estimate in the proof of Lemma 14 now gives $\sum_i \|u_i - \mathbf{1}/\sqrt{m-1}\|_2^2 = o(m^{-1})$, by the root off-diagonal and reserve bounds. Undoing the child phase change costs at most $K(m-1)^2\|u_i - \mathbf{1}/\sqrt{m-1}\|_2^2$ in squared Frobenius norm for H_{m-1}°. This follows from $|\operatorname{phase}(u_{i,j}) - 1| \le 2\sqrt{m-1}|u_{i,j} - 1/\sqrt{m-1}|$ and the rank-one form of the reference. Thus the errors in the inherited gauge still sum to $o(m)$.

Repeat the spectral comparison proving (44), with (52), the $o(m)$ inherited squared errors and the $o(m)$ child reserves. All pairs participate, with a spectral gap of order m by $\max d_{ij} = o(m)$ and the Frobenius bound on B_{ij}. Now the sum in (44) is $o(1)$. By (41), the squared deviations of $v_{i|j}^* v_{j|i}$ from $(1 - X_{m-1})/X_{m-1}$ sum to $o(1)$. The p_{ij}'s are bounded below, and their squared errors from $\bar{x}X_{m-1}$ also sum to $o(1)$, by (52) and the root variance bound. Since $\bar{x} = 1/2 + 1/(2m) + o(m^{-1})$, the off-diagonal scalar predicted by (41) is

$$\frac{h_m}{\bar{x}X_{m-1}} - \frac{1 - X_{m-1}}{X_{m-1}} = 2/m + o(m^{-1}).$$

Consequently $\|T - A_{*,m}\|_{\mathrm{F}} = o(1)$. For the row sums, (45) and (51) give an error vector of ℓ^1 norm $o(1)$ when H_{ij} is replaced by h_m. Also $\|x - \mathbf{1}/2\|_2 = O(m^{-1/2})$. It follows that $\|(T - I)\mathbf{1} - 2\mathbf{1}\|_2 = o(1)$, and $(m-1)b_m = 2 + O(m^{-1})$ gives the required row-sum comparison. Lemma 14 proves the $O_B(1)$ assertion in the broader band. This completes Theorem 11.

41 Exact comparison for large orders

The uniform local theorem Theorem 7 provides an absolute radius in the sum of Frobenius and unnormalized row-sum distances, within which $\Lambda(A) \le \Lambda(A_{*,n})$ and equality requires the scalar

orbit. Any sequence of unrestricted maximizers enters this radius by Theorem 11. Hence all maximizers for sufficiently large n lie on the orbit. The exponential error then follows from the scalar evaluation Proposition 2.

There is also a fixed η/n band in the local neighborhood, for some absolute $\eta > 0$. Otherwise there would be a sequence with $n \to \infty$ and $n(\gamma_n - \Lambda(A)) \to 0$ outside the neighborhood in every gauge, contradicting Theorem 11. In the local theorem's real-row-sum gauge, its longitudinal loading difference has norm t and its centered transverse perturbation has Frobenius norm α. The proof gives a lower bound $c(t^2+\alpha^2)$ on the objective loss and an upper bound $K(t+\alpha)$ on the sum of matrix and row-sum distances in that gauge. Combining them proves the quantitative assertion of Theorem 12.

42 Walsh families: the model and the cofactor identity

For a correlation matrix B define $C(B)_{ij} = \operatorname{per} B(i,j)$, with row i and column j deleted, and put $\Lambda(B) = \lambda_{\max}(C(B))/\operatorname{per} B$. This is the raw, unweighted cofactor convention of Theorem 1. The unrestricted bounds of Lemma 71 agree at orders n and one and leave a gap in the coefficient of $1/n$. The local result of Theorem 1 controls all directions near the scalar optimizer. Here we obtain exact maximization in a fixed finite-factor class by treating both its near-optimal limits and its local maximization in strength coordinates.

The raw cofactor bounds of Marcus and Merris [4] and Lavi [3] provide precedents for the spectral observable. Complex Gaussian permanent moments are classical [9], and Gaussian optical photon counts also involve positive semidefinite permanents [23]. McCullagh [12, Section 3] analyzes scalar permanent approximations for matrices with two blocks and shrinking between-block entries. We use the Gaussian law tilted by the permanent to represent the normalized raw cofactor matrix. Its phase covariance gives the limiting cofactor spectrum when the two blocks or more blocks approach decoupling.

Let $G = (\mathbb{Z}/2\mathbb{Z})^k$ be fixed, of size m, and assume n is a multiple of m. Index the rows by equal fibers over $x \in G$. Each character χ gives the sign vector s_χ on these rows and $s_0 = \mathbf{1}$ for the trivial character. We study

$$B_n = a_n I + \frac{1}{n}\sum_{\chi} \beta_{\chi,n} s_\chi s_\chi^T, \qquad a_n = 1 - \frac{1}{n}\sum_{\chi}\beta_{\chi,n}, \qquad \beta_{\chi,n} \geq 0.$$

The strengths are uniformly bounded, so $a_n > 0$ for all sufficiently large n. Character gauges are conjugations by diagonal matrices with a character's signs on the fibers. They permute the strength labels and preserve Λ.

Take independent standard circular complex Gaussians G_χ with $\mathbb{E}|G_\chi|^2 = 1$. Set

$$z_x = n^{-1/2}\sum_{\chi}\sqrt{\beta_{\chi,n}}\chi(x)G_\chi, \qquad d_x = a_n + |z_x|^2.$$

The row variables $X_i = \sqrt{a_n}\epsilon_i + z_{x_i}$, with independent standard circular ϵ_i, have covariance B_n. Wick's formula and conditional integration give $\operatorname{per} B_n = \mathbb{E}\prod_x d_x^{n/m}$. Write $\mathbb{E}_{\mathrm{t}}$ for expectation after tilting the latent Gaussian law by this product. For the raw cofactor the unpaired variables at $i \neq j$ are $\overline{X}_i X_j$ and its diagonal expectation contains $1/d_{x_i}$. It follows that

$$\frac{C(B_n)}{\operatorname{per} B_n} = \mathbb{E}_{\mathrm{t}}\big[\operatorname{diag}(a_n/d_{x_i}^2) + vv^*\big], \qquad v_i = \overline{z}_{x_i}/d_{x_i}. \tag{53}$$

Indeed the rank-one term supplies $\bar{z}_{x_i} z_{x_j}/(d_{x_i} d_{x_j})$ off the diagonal, and the diagonal completes it to $1/d_{x_i}$.

Group translations and within-fiber permutations commute with the cofactor. The fiber-constant space is diagonalized by the characters. Its orthogonal complement sees a diagonal operator of norm $O(1)$ by (53). Write $\langle f \rangle_G = m^{-1} \sum_x f_x$ for a fiber average. The uniform-mode eigenvalue has the exact defect identity

$$\mathbb{E}_{\mathrm{t}} \left\langle a_n / d_x^2 \right\rangle_G + \frac{n}{4a_n} - n\mathbb{E}_{\mathrm{t}} \left\langle \frac{(|z_x|^2 - a_n)^2}{4a_n d_x^2} \right\rangle_G - n\mathbb{E}_{\mathrm{t}} \left\langle |v_x - \langle v \rangle_G|^2 \right\rangle_G . \tag{54}$$

This follows by subtracting the fiber variance from $\left\langle |v|^2 \right\rangle_G$ and using

$$\frac{u}{(a+u)^2} = \frac{1}{4a} - \frac{(u-a)^2}{4a(a+u)^2}.$$

43 A dominant character near strength two

Theorem 16. *Suppose $\beta_{0,n} = 2 + u/\sqrt{n} + o(n^{-1/2})$ and $\beta_{\chi,n} \to \kappa_\chi < 2$ for all $\chi \neq 0$. Then*

$$\Lambda(B_n) = \frac{n}{4} + \frac{1}{2} - \frac{u^2}{16} - \sum_{\chi \neq 0} \frac{\kappa_\chi^2}{8(2 - \kappa_\chi)} + o(1).$$

Proof. Rotate by the phase of G_0, let $T = |G_0|^2$, and use the independent rotated secondary Gaussians $H_\chi = X_\chi + iY_\chi$. Write $\beta = \beta_{0,n}, \kappa_{\chi,n} = \beta_{\chi,n}$ and

$$g_0 = \sqrt{\beta T/n}, \quad h_x = \sum_{\chi \neq 0} \sqrt{\kappa_{\chi,n}} \chi(x) H_\chi, \quad z_x = g_0 + n^{-1/2} h_x, \qquad X = (T - n(1 - 1/\beta))/\sqrt{n}.$$

The common rotation leaves all observables of (53) unchanged. The tilted variables X, X_χ, Y_χ converge with polynomial moments to independent centered Gaussians of variances $1, 1/2, 1/(2 - \kappa_\chi)$, respectively. For domination, AM–GM and orthogonality bound the tilted product by $(a_n + [\beta T + \sum \kappa_{\chi,n} |H_\chi|^2]/n)^n$. With $\widetilde{T} = T + \sum (\kappa_{\chi,n}/\beta) |H_\chi|^2$, the Gaussian densities leave a radial factor $e^{-\widetilde{T}} (a_n + \beta \widetilde{T}/n)^n$ and secondary factors $e^{-(1 - \kappa_{\chi,n}/\beta)|H_\chi|^2}$. These rates have a uniform positive lower bound. The radial factor is a translated Gamma density of shape $n+1$, with lower cutoff at $n a_n/\beta$ before translation and with center differing by $O(1)$ from $n(1 - 1/\beta)$ in $\widetilde{T}$. Its rescaled centered moments are bounded by the Gamma moment generating function. The shift from $\widetilde{T}$ to T is controlled by the secondary Gaussian factors. This envelope is an upper bound in integral of the same order as the true normalizer, since the local expansion below gives a comparable lower mass on a fixed compact fluctuation box.

With $q = a_n + \beta T/n = \beta + \beta X/\sqrt{n} + O(n^{-1})$, the radial log density has limit $-X^2/2$, up to its constant. In the fiber log expansion use $d_x = q + 2g_0 \operatorname{Re} h_x/\sqrt{n} + |h_x|^2/n$. Orthogonality cancels the linear term in $\operatorname{Re} h_x$. The two order-one terms give

$$\tfrac{1}{2} \sum_\chi \kappa_\chi (X_\chi^2 + Y_\chi^2) - \tfrac{1}{2} \sum_\chi \kappa_\chi X_\chi^2.$$

With the original Gaussian costs this is precisely the asserted covariance. The expansions are uniform on compact fluctuation boxes, and domination as above gives tightness and uniform integrability of polynomial moments.

On the subspace orthogonal to the uniform vector, the rank expectation in (53) has trace $n\mathbb{E}_{\mathrm{t}}\left\langle |v_x - \langle v\rangle_G|^2\right\rangle_G = O(1)$. This follows from the uniformly Lipschitz map $z \mapsto \bar{z}/(a_n + |z|^2)$ and the bounded moments of h_x. The uniform-mode eigenvalue is asymptotic to $n/4$ and is therefore the top eigenvalue for large n.

In (54) the average diagonal tends to $1/4$ and $n/(4a_n) = n/4 + (2 + \sum \kappa_\chi)/4 + o(1)$. The first deficit tends to $(u^2+4+2\sum \kappa_\chi)/16$, since $\sqrt{n}(|z_x|^2 - a_n)$ tends to $u + 2X + 2\operatorname{Re} h_x$ in moments and $d_x \to 2$. For the second deficit, the rescaled centered v_x tends to $-i \operatorname{Im} h_x/2$. This is the derivative at $z = 1, a = 1$, which vanishes in the real direction. Orthogonality and the limiting imaginary variances give the second deficit $\sum \kappa_\chi/[4(2-\kappa_\chi)]$. The positive denominators and moment bounds justify taking expectations in these expressions. Substituting the two deficits gives the formula. □

44 Global rigidity and exact maximization

Theorem 17. *Fix the Walsh group and any cap $B > 2$. For all sufficiently large n divisible by m, the maximum of $\Lambda(B_n)$ over $0 \le \beta_{\chi,n} \le B$ equals the scalar equicorrelation maximum. Every maximizer, after a character gauge, has only the trivial strength nonzero and is the scalar maximizer. In particular, the maximum is $(n+1)^2/(4n) + O(ne^{-cn})$ for some $c > 0$.*

44.1 Localization of two-term near-maximizers

Suppose first that the strengths are uniformly bounded and $\Lambda(B_n) \ge n/4 + 1/2 - o(1)$. The top eigenvector must lie in the fiber-constant space, since its complement has bounded norm. Gauge a maximizing character to the trivial one. Then (54) equals $\Lambda(B_n)$. Its two nonnegative deficits are bounded, since $a_n = 1 - O(n^{-1})$ and the diagonal term is bounded. Hence in tilted probability, for each of the fixed number of fibers,

$$|z_x|^2 = 1 + O_{\mathbb{P}}(n^{-1/2}), \qquad z_x - \langle z\rangle_G = O_{\mathbb{P}}(n^{-1/2}). \tag{55}$$

Indeed the first defect controls $(|z_x|^2 - a_n)/d_x$ at this scale, which also controls $|z_x|^2 - a_n$ by solving for it. The second defect controls $v_x - \langle v\rangle_G$; then use $z_x = d_x \bar{v}_x$ and $d_x = 2 + O_{\mathbb{P}}(n^{-1/2})$.

In rotated latent coordinates as in Theorem 16, (55) implies $\beta T/n = 1 + O_{\mathbb{P}}(n^{-1/2})$ and $\sqrt{\kappa_{\chi,n}} H_\chi = O_{\mathbb{P}}(1)$. The scaled latent vector $(G_\chi/\sqrt{n})$ is tight for bounded strengths. With $R = n^{-1}\sum |G_\chi|^2$, the Gaussian radial cost and AM–GM give the exponential tail bound with exponent $n[-R + \log(1 + BR)]$, up to a fixed-degree radial volume factor. The permanent normalizer is bounded below by $a_n^n \ge c > 0$. In particular β cannot tend to zero along near-maximizers.

Pass to a subsequence on which the strengths converge, with trivial limit $\beta_* > 0$. The scaled latent vector concentrates on the maximizers of

$$-\sum_\chi |w_\chi|^2 + \frac{1}{m}\sum_x \log\left(1 + \left|\sum_\chi \sqrt{\beta_{\chi,*}}\chi(x) w_\chi\right|^2\right).$$

This follows by uniform convergence of the exponents on compacts, the displayed tail bound, and integration of a small ball near a maximum for a lower bound. A limiting latent vector must give all z_x equal with modulus one by (55). Its positive-strength nontrivial coordinates vanish by character orthogonality, and the zero-strength coordinates vanish at a maximum since they only enter the negative quadratic term. At the trivial-axis maximizer $|w_0|^2 = 1/\beta_*$, so the radial derivative is $-1 + \beta_*/2$. Thus $\beta_* = 2$. No secondary limit can exceed two, since moving to that pure character at the same radius would then increase the exponent.

The secondary H_χ are tight when their strengths have positive limits, by (55). To see tightness when $\kappa_{\chi,n} \to 0$, condition on T and the other rotated variables. With arbitrarily high probability the trivial latent value $g_0 = \sqrt{\beta T/n}$ is in a fixed compact neighborhood of one, and all other $\sqrt{\kappa_{j,n}} H_j$ are bounded by a fixed M. Denote the field with χ omitted by $z_x^0 = g_0 + O(M/\sqrt{n})$, and write $d_x^0 = a_n + |z_x^0|^2$. The change in the log tilt is bounded above using $\log d_x - \log d_x^0 \le (d_x - d_x^0)/d_x^0$. Its sum with multiplicities is at most $C_M \sqrt{\kappa_{\chi,n}} |H_\chi| + C\kappa_{\chi,n} |H_\chi|^2$, because the linear term in the constant value of $\bar{z}_x^0/d_x^0$ cancels by the zero character sum. On $|H_\chi| \le 1$, Taylor expansion and the same cancellation give an absolute bound C_M. Thus the conditional normalizer has a positive lower bound after dividing by the value at $H_\chi = 0$, and the original Gaussian yields a common conditional Gaussian tail. This proves full secondary tightness.

Set $X_n = (T - n/\beta)/\sqrt{n}$, tight by (55). On any compact of its coordinates and the H_χ, the log of the unnormalized density divided by its central value has expansion

$$L_n X_n - \tfrac{1}{2} X_n^2 - \sum_\chi X_\chi^2 - \sum_\chi (1 - \kappa_\chi/2) Y_\chi^2 + o(1), \qquad L_n = \sqrt{n} \left(\frac{\beta}{a_n + 1} - 1 \right),$$

where κ_χ denotes the limiting secondary strengths on the subsequence. This is the same character cancellation as in Theorem 16, with radial center at n/β. If L_n were unbounded, moving any compact box a fixed distance in X_n in its sign direction would make the density on the shifted box arbitrarily larger than on the original, along a subsequence. This would force the original compact mass to vanish, contradicting tightness. Hence $\sqrt{n}(\beta - 2)$ is bounded. Pass further so that L_n converges. If some $\kappa_\chi = 2$, the local limiting unnormalized density is constant in Y_χ, so its full integral is infinite. Integrating over arbitrarily large compact boxes then makes the full normalizer divided by the central density diverge, forcing every fixed compact to have vanishing probability. This again contradicts tightness. Thus all $\kappa_\chi < 2$.

Theorem 16 now applies at every subsequential limit u of $\sqrt{n}(\beta - 2)$. Its nonnegative defect $u^2/16 + \sum_\chi \kappa_\chi^2/[8(2 - \kappa_\chi)]$ must be zero at a near-maximizer. We conclude $\sqrt{n}(\beta - 2) \to 0$ and $\max_{\chi \ne 0} \kappa_{\chi,n} \to 0$ after the chosen gauges.

44.2 A uniform strength neighborhood

Let $\beta_n^* = n b_n$, with b_n the exact scalar maximizer from Theorem 1, so $\beta_n^* = 2 + O(n^{-1})$. Parametrize the neighborhood by $\beta = \beta_n^* + u/\sqrt{n}$ and $\kappa_{\chi,n} = \psi_\chi^2$, with u, ψ in fixed sufficiently small symmetric intervals. Let $F_n(u, \psi)$ be (54) minus $n/4$. On this rectangle the uniform mode is the top eigenvalue for large n, uniformly, by the moment and complement bounds of Theorem 16. We prove

$$F_n \;\longrightarrow\; \tfrac{1}{2} - \frac{u^2}{16} - \sum_\chi \frac{\psi_\chi^4}{8(2 - \psi_\chi^2)} \quad \text{in } C^4(u, \psi).$$

Take the fixed radial center $T_0 = n(1 - 1/\beta_n^*)$, and $X = (T - T_0)/\sqrt{n}$. Put $q = a_n + \beta T/n$, $q_0 = a_n + \beta T_0/n$, and write $d_x = q + r_x$, where

$$r_x = 2 g_0 \operatorname{Re} h_x/\sqrt{n} + |h_x|^2/n, \qquad h_x = \sum_\chi \psi_\chi \chi(x) H_\chi.$$

Divide the unnormalized density in (X, H) by its value at the center. Its log tilt terms decompose as

$$n \log(q/q_0) + n \langle r_x \rangle_G / q + n \langle \log(q + r_x) - \log q - r_x/q \rangle_G .$$

The first term has polynomially bounded derivatives through order four, uniformly on the rectangle. In fact $q/q_0 = 1 + (\beta/q_0)X/\sqrt{n}$ is bounded below and each positive-order parameter derivative of β/q_0 is $O(n^{-1/2})$. The second term uses $\langle r_x \rangle_G = n^{-1} \sum \psi_\chi^2 |H_\chi|^2$. In the third term use the integral Taylor remainder with numerator r_x^2 and denominators $(q + t r_x)^2$, $0 \le t \le 1$, all bounded below by a_n^2. The factors $\sqrt{n}\, r_x$ and their parameter derivatives are polynomially bounded in (X, H), since $T/n \le C(1 + |X|)$ and derivatives of $g_0 = \sqrt{\beta T/n}$ differentiate only $\sqrt{\beta}$. This proves polynomial bounds for all differentiated density factors after taking out the density.

The diagonal contribution and the first defect in (54) have the same bounds after differentiation, since $d_x \ge 1/2$ and $\sqrt{n}(|z_x|^2 - a_n)$ is polynomially bounded with its derivatives. For the last defect, subtract the common value $v(g_0)$ before centering over the fibers. The expression $\sqrt{n}[v(z_x) - v(g_0)]$ is an integral of the real differential of v applied to h_x along the line from g_0 to z_x. The rational map $v(z) = \bar{z}/(a_n + |z|^2)$ and its fixed-order derivatives have polynomial bounds for $a_n \ge 1/2$, which also bounds all differentiated versions of this defect.

The preceding Gamma envelope is uniform on the rectangle. Its integral relative to the density value at the center is bounded in $dX\, dH$ by Stirling's formula, since the translated Gamma center is within $O(\sqrt{n})$ of T_0 and the T-to-X Jacobian is $\sqrt{n}$. Its bounded moments of all orders give uniform integrability of the polynomial derivative bounds. The true relative integral has a uniform positive lower bound on a compact fluctuation box by the local expansion. Differentiation of the normalized integrals is therefore justified, and convergence of the differentiated expansions on compact fluctuation boxes gives uniform convergence through order four. In these fixed-center coordinates, the radial limiting log tilt is $uX/4 - X^2/2$. After recentering X by $u/4$, the dominant-character calculation gives exactly the displayed limit for F_n.

The functions are separately even in every ψ_χ, by changing the sign of the corresponding H_χ. Thus C^4 convergence gives C^2 convergence on the closed strength rectangle $\kappa_\chi = \psi_\chi^2 \ge 0$. For a one-variable even function f, the first derivative of $f(\sqrt{\kappa})$ is $\frac{1}{2}\int_0^1 f''(t\sqrt{\kappa})dt$, and the second is $\frac{1}{4}\int_0^1 \int_0^1 t^2 f^{(4)}(st\sqrt{\kappa})ds\, dt$, both continuous at zero. Apply the same formula in each variable for mixed derivatives. The limiting Hessian in the strength coordinates is uniformly negative definite on a small fixed rectangle, so the exact objective is strictly concave there for large n.

Its gradient at $u = 0, \kappa = 0$ is zero exactly. The scalar coordinate is stationary by the interior scalar optimum. At the same equicorrelation point, simplicity and permutation invariance make the derivative on real symmetric zero-diagonal perturbations proportional to the sum of off-diagonal entries. Since the scalar perturbation has nonzero such sum and derivative zero, the derivative in every secondary direction is zero. Thus the unique strength maximizer in this neighborhood is the scalar point.

The capped family is compact and contains the scalar point for large n. Every maximizing sequence is two-term near-optimal; the localization above places it, after a gauge, inside the strength rectangle for sufficiently large n. The strict concavity then proves Theorem 17, including equality. The scalar value and its error follow from Theorem 1.

45 The critical quartic regime at strength one

The scale in this section concerns growth of the ratio, before the optimization near strength two. The scalar radial transition is related to incomplete Gamma asymptotics [24]. Quartic critical laws for vector spin systems [25] are another precedent for the limiting form of the exponent. Here the law controls the largest eigenvalue of a permanental cofactor.

Theorem 18. *If all $\beta_{\chi,n} \to b_\chi < 1$, then*

$$\Lambda(B_n) \longrightarrow \max_\chi (1 - b_\chi)^{-1}.$$

Suppose a nonempty character set J has $\beta_{\chi,n} = 1 + u_\chi/\sqrt{n} + o(n^{-1/2})$, with all other limits $b_\chi < 1$. Let $Y = (Y_\chi)_{\chi \in J}$ be the complex vector of density proportional to

$$\exp\left[\sum_{\chi \in J} u_\chi |Y_\chi|^2 - \frac{1}{2m} \sum_{x \in G} \left| \sum_{\chi \in J} \chi(x) Y_\chi \right|^4 \right].$$

Then

$$n^{-1/2} \Lambda(B_n) \longrightarrow \lambda_{\max}(\mathbb{E}[YY^*]) = \max_{\chi \in J} \mathbb{E}|Y_\chi|^2.$$

For one critical character of detuning u, the value is $u + \varphi(u)/\Phi(u)$, with φ, Φ the standard real normal density and distribution function.

Proof. The exact fiber identity

$$\frac{n}{m} \sum_x (d_x - 1) = \sum_\chi \beta_{\chi,n} (|G_\chi|^2 - 1)$$

follows from orthogonality. In the subcritical regime, the log tilt converges to $\sum b_\chi (|G_\chi|^2 - 1)$ at fixed G_χ. The bound $\log d_x \le d_x - 1$ gives a common Gaussian majorant with all moments, since the strengths are uniformly below one. The latent limits then have independent circular variances $(1 - b_\chi)^{-1}$. In (53) the diagonal tends to the identity. In the orthonormal fiber basis the rank expectation has matrix $(n/m)\mathbb{E}_{\mathrm{t}}[vv^*]$ using the m fiber values. Here $v_x = \bar{z}_x + o(n^{-1/2})$ in tilted L^2, so this rank contribution has limiting eigenvalues $b_\chi/(1 - b_\chi)$ by character orthogonality. The subcritical formula follows.

For critical characters set $G_\chi = n^{1/4} Y_\chi$ and leave the other Gaussians unscaled. At fixed rescaled variables, with $\psi_x = \sum_{\chi \in J} \chi(x) Y_\chi$, one has $n^{1/4} z_x \to \psi_x$. The exact linear sum cancels the order-$\sqrt{n}$ Gaussian weight of the critical variables and gives the detuning terms. The second log term tends to the quartic term in the statement. The unscaled modes have independent limiting Gaussian weights $e^{-(1-b_\chi)|G_\chi|^2}$.

For domination put $T = n^{-1/2} \sum_J \beta_{\chi,n} |Y_\chi|^2 + n^{-1} \sum_{\chi \notin J} \beta_{\chi,n} |G_\chi|^2$. AM–GM bounds the tilt by $(1 + T)^n$. Use $\log(1 + T) \le T - c \min(T^2, T)$ for $T \ge 0$. After the original Gaussian costs, the joint exponent has the upper bound

$$C\|Y\|^2 - c' \sum_{\chi \notin J} |G_\chi|^2 - c'' \min(\|Y\|^4, \sqrt{n}\, \|Y\|^2).$$

This gives a common integrable Gaussian upper envelope for all sufficiently large n, also for any fixed polynomial moment. The local exponents converge uniformly on compact sets and the limit has positive normalizing mass there. The quartic term is coercive, since $m^{-1} \sum_x |\psi_x|^4 \ge \|Y\|^4$ by Jensen and orthogonality. Thus this is the joint limiting law.

After division by $\sqrt{n}$ the bounded diagonal term of (53) vanishes. Since $n^{1/4} v_x \to \bar{\psi}_x$ in tilted L^2, the normalized rank part on the fiber basis tends to $m^{-1}\mathbb{E}[\bar{\psi}\, \bar{\psi}^*]$. Its nonzero eigenvalues are those of the conjugate of $\mathbb{E}[YY^*]$ by the orthonormal Walsh columns. The covariance is diagonal in the character basis, by invariance under group translation of the fiber index in the quartic sum. For one critical character its squared modulus has density proportional to $e^{ut - t^2/2}$ on $t \ge 0$, giving the stated truncated normal mean. □

46 Degenerate supercritical strengths

Theorem 19. *Suppose all m strengths satisfy $\beta_{\chi,n} = \gamma + \tau_\chi/n + o(n^{-1})$ with $\gamma > 1$. Put $W_{x,\chi} = \chi(x)/\sqrt{m}$ and $Q = W\,\mathrm{diag}(\tau_\chi)W^*$. Let $\Xi = (e^{i\theta_x})_{x\in G}$ have density proportional to*

$$\exp\left[\frac{\gamma-1}{m\gamma^2}\Xi^* Q\Xi\right]$$

with respect to uniform product angles on the torus. Then

$$\frac{\Lambda(B_n)}{n} \longrightarrow \frac{\gamma-1}{m\gamma^2}\lambda_{\max}(\mathbb{E}[\Xi\Xi^*]).$$

If $m > 1$ and all detunings are finite, this value is strictly below $(\gamma-1)/\gamma^2$. At equal detunings it is $(\gamma-1)/(m\gamma^2)$. If $m = 2$, it is

$$\frac{\gamma-1}{2\gamma^2}\left(1 + \left|\frac{I_1(\mu)}{I_0(\mu)}\right|\right), \qquad \mu = \frac{\gamma-1}{2\gamma^2}(\tau_0 - \tau_1),$$

where I_0, I_1 are modified Bessel functions.

Proof. In the latent fiber coordinates $z = (z_x)$ the covariance is $(m/n)W\,\mathrm{diag}(\beta_{\chi,n})W^*$. Thus the tilted density is proportional to

$$\exp[-(n/m)z^*W\,\mathrm{diag}(\beta_{\chi,n}^{-1})W^*z]\prod_x (a_n + |z_x|^2)^{n/m}.$$

At reference strength γ, the squared radii have independent translated Gamma densities with rate $n/(m\gamma)$, restricted to positive $t_x = |z_x|^2$. Their centers tend to $\gamma - 1 > 0$. The angles are uniform and independent. Expanding the inverse strengths multiplies this density by $\exp[z^*Qz/(m\gamma^2) + o(1)\|z\|^2]$. The radial reference variables have uniformly bounded exponential moments at each fixed order, from the Gamma formula and the separated lower cutoff. This controls the tilt, and radial concentration gives the displayed limiting phase density, with convergence of the moments used in the cofactor identity. Phase models of this form are studied in [26].

On the fiber basis the rank term in (53) is $(n/m)\mathbb{E}_{\mathrm{t}}[vv^*]$. Its n^{-1} limit has the spectrum asserted because v_x tends to $\sqrt{\gamma-1}\,\overline{\Xi_x}/\gamma$ in tilted expectations. The diagonal part is bounded. The phase covariance has trace m and is positive definite under the everywhere positive torus density. Thus for $m > 1$ its top eigenvalue is strictly below m. Equal detunings give a scalar Q and independent uniform angles. For two fibers the difference angle has density proportional to $e^{\mu\cos\theta}$, giving the stated I_1/I_0 Fourier moment. □

Theorems 17–19 concern different strength regimes in the same fixed Walsh family. In particular Theorem 17 determines the third term only for a fixed bounded-rank correction with balanced Walsh fibers.

47 Lower comparisons for the normalized cofactor matrix

For A of order n, define $C(A)_{ij} = \operatorname{per} A(i \mid j)$, where the minor omits row i and column j. We use the adjoint convention $C(A)^{\mathsf{T}}$ of Marcus and Merris [4]. A Hermitian correlation matrix has unit diagonal and is positive semidefinite. The normalized adjoint is positive semidefinite by the Gaussian deletion-product representation. Marcus and Merris give the upper comparison $C(A)^{\mathsf{T}}/\operatorname{per} A \preceq nA^{-1}$ for invertible A. Other comparisons use subpermanent sums [27] or entry-weighted cofactors [22]. We compare the normalized raw adjoint from below with its original matrix.

Theorem 20. *Let $n \geq 2$. For every real correlation matrix A of rank at most two,*

$$\frac{C(A)^{\mathsf{T}}}{\operatorname{per} A} \succeq \frac{1}{n2^{n-2}}\, A. \tag{56}$$

The coefficient is optimal in every order, including on the subset of rank-two matrices whose entries are strictly positive. The sharp quotient is approached by matrices converging to the all-ones matrix.

Proof. Let X, Y be independent standard circular complex Gaussian variables. Realize A as the Gram matrix of forms $L_j = a_j X + b_j Y$ with a_j, b_j real and $a_j^2 + b_j^2 = 1$. Use the homogeneous polynomial norm $\|P\|^2 = \mathbb{E}|P(X,Y)|^2$, under which $\|X^k Y^l\|^2 = k!l!$ and the monomials are orthogonal. The circular Gaussian moment identity [17, Theorem A.4] gives the permanent and the cofactor quadratics

$$\operatorname{per} A = \left\| \prod_j L_j \right\|^2, \qquad x^{\mathsf{T}} C(A)^{\mathsf{T}} x = \|D_x\|^2, \qquad D_x = \sum_j x_j \prod_{l \neq j} L_l.$$

Here x is real. Each matrix in (56) is real symmetric, so real quadratic tests suffice.

Put $w = (1, i)$. Each $L_j(w)$ has modulus one. Therefore

$$|D_x(w)|^2 = \left| \prod_j L_j(w) \right|^2 \left| \sum_j x_j \overline{L_j(w)} \right|^2 = x^{\mathsf{T}} A x.$$

For degree $d = n - 1$ the homogeneous evaluation kernel at w is $K_w(X,Y) = (X - iY)^d / d!$. This follows by expanding in the normalized monomial basis. The evaluation kernel at $\bar{w} = (1, -i)$ has the same squared norm $2^d/d!$. The two kernels are orthogonal, since $K_w(\bar{w}) = 0$ for $d \geq 1$. The values of a real polynomial at these points are conjugate, so Bessel's inequality applied to the two normalized kernels gives

$$|D_x(w)|^2 \leq \frac{2^{d-1}}{d!}\, \|D_x\|^2. \tag{57}$$

Every entry of A has absolute value at most one, and consequently $\operatorname{per} A \leq n!$. Combining this bound with (57) proves (56).

For sharpness let $t_1, \ldots, t_n$ be distinct positive real parameters and take

$$L_j = \frac{X + t_j Y}{\sqrt{1 + t_j^2}}.$$

The deletion polynomials form a basis of the real binary forms of degree d: evaluation at $X = -t_j, Y = 1$ isolates the j-th deletion. Choose a real x so that $D_x = K_w + K_{\bar{w}}$. Then (57) is an equality and $D_x(w) \neq 0$. The relative quadratic ratio is

$$\frac{x^{\mathsf{T}} C(A)^{\mathsf{T}} x}{\operatorname{per} A\, x^{\mathsf{T}} A x} = \frac{(n-1)!}{2^{n-2} \operatorname{per} A}.$$

Let the distinct parameters tend to zero at fixed n. These matrices have rank two and positive entries, and their permanents tend to $n!$. This proves sharpness. □

For $n \geq 3$ no fixed matrix attains the sharp relative ratio on a vector with $x^{\mathsf{T}} A x > 0$. In rank two some transposition contributes a term of absolute value strictly less than one to the permanent sum, and so $\operatorname{per} A < n!$. In rank one, diagonal sign conjugation reduces the matrix to the all-ones matrix and the normalized adjoint is A/n. In order two the sharp ratio $1/2$ is attained by the rank-one matrices.

48 A positive comparison in every real rank

Theorem 21. *There is an absolute constant K such that for every real correlation matrix A of order $n \geq 2$,*

$$\frac{C(A)^{\mathsf{T}}}{\operatorname{per} A} \succeq (Kn)^{-Kn} A.$$

Proof. Choose real unit Gram vectors $v_i \in \mathbb{R}^n$. Among the signed sums $s = \sum_i \epsilon_i v_i$, $\epsilon_i \in \{-1, 1\}$, choose one of maximal norm. Flipping the i-th sign and expanding the squared norm gives $\epsilon_i \langle v_i, s\rangle \geq 1$. Hence $e = s/\|s\|$ satisfies $|\langle v_i, e\rangle| \geq 1/n$ for every i. Conjugate A by the diagonal matrix of the signs so the inner products are positive. This transforms the normalized transposed cofactors by the same congruence, as is seen directly on the deleted minors. In an orthonormal coordinate system based at e write

$$v_i = a_i(1, t_i), \qquad a_i \geq 1/n, \qquad t_i \in \mathbb{R}^{n-1}, \quad \|t_i\| \leq n, \qquad a_i^2 = \frac{1}{1 + \|t_i\|^2}.$$

Use the Gaussian polynomial model on these coordinates with $L_i(X, Y) = a_i(X + t_i \cdot Y)$. For real x use the same deleted polynomial D_x as before and set $\alpha_i = x_i/a_i$. The rational function

$$R_x(z) = \frac{D_x(1, z)}{(\prod_i a_i) \prod_i (1 + t_i \cdot z)} = \sum_i \frac{\alpha_i}{1 + t_i \cdot z}$$

is analytic at zero. If β is a multi-index of order j, the coefficient of z^β is $(-1)^j \binom{j}{\beta} \sum_i \alpha_i t_i^\beta$. On the complex polydisc of coordinate radius $r = (2n^{3/2})^{-1}$ all $|t_i \cdot z| \leq 1/2$ and the denominator of R_x has modulus at least $(2n)^{-n}$. For a homogeneous polynomial of degree $n - 1$, the Gaussian evaluation-kernel norm at $(1, z)$ is $\|(1, z)\|^{n-1}/\sqrt{(n-1)!}$, by the same monomial calculation in n variables. This is at most 2 on the polydisc. Cauchy's estimate therefore gives

$$\max_{|\beta| \leq 2n-1} \left| \sum_i \alpha_i t_i^\beta \right| \leq 2(2n)^n r^{-(2n-1)} \|D_x\|. \tag{58}$$

To recover the original linear combination of Gram vectors, put $s_i = 1 + \|t_i\|^2$ and use the polynomial

$$h(s) = \frac{1 - \prod_{i=1}^n (1 - s/s_i)}{s}.$$

It has degree at most $n - 1$ and $h(s_i) = s_i^{-1}$, including at repeated nodes. Its coefficient of s^j has absolute value at most $\binom{n}{j+1}$ because each $s_i \geq 1$. For a real unit output vector u, the polynomial

$$P_u(t) = (u \cdot (1, t))\, h\left(1 + \sum_l t_l^2\right)$$

has degree at most $2n - 1$ and coefficient ℓ^1 norm at most $\sqrt{n} \sum_{j=0}^{n-1} \binom{n}{j+1} n^j \leq \sqrt{n}(n+1)^n$. At each node it matches $u \cdot (1, t_i)/s_i$, so

$$u \cdot \sum_i x_i v_i = \sum_i \alpha_i P_u(t_i).$$

Bound this scalar product by the coefficient ℓ^1 norm and (58), and take the supremum over u. The result is $\|\sum_i x_i v_i\| \leq (K_0 n)^{K_0 n} \|D_x\|$ for an absolute K_0. Squaring, using the Gaussian cofactor Gram identity and $\operatorname{per} A \leq n!$, and enlarging the absolute constant proves the theorem. □

The proof allows arbitrary rank and repeated Gram vectors. The optimal all-rank constant is at most $1/(n2^{n-2})$ by Theorem 20. The next example shows that the rank-two constant can be too large in higher rank.

49 An upper witness in higher real rank

For $n \geq 4$ let $m = n-2$ and $k = n-1$. Choose the k unit vertices τ_j of a centered regular simplex in $\mathbb{R}^m$, so that $\sum_j \tau_j = 0$ and $\sum_j \tau_j \tau_j^{\mathsf{T}} = (k/m)I$. Take the Gaussian forms $L_0 = X$ and

$$L_j = \frac{X + \delta\, \tau_j \cdot Y}{\sqrt{1+\delta^2}}, \qquad 1 \leq j \leq k.$$

They give real positive-entry correlation matrices of rank $n-1$ for sufficiently small $\delta > 0$. In the affine chart of the preceding proof take $\alpha_0 = -k$ and $\alpha_j = 1$ for $j \geq 1$, with $x_i = a_i \alpha_i$. The constant and linear terms of the rational sum vanish, and its quadratic term is $\delta^2 (k/m) \sum_l z_l^2$. Thus in Fock norm,

$$D_x = \delta^2 (k/m) X^{n-3} \sum_{l=1}^{m} Y_l^2 + O(\delta^3).$$

The original Gram combination has only its first coordinate, equal to $-k + k/(1+\delta^2)$. The monomial norms, together with $\operatorname{per} A \to n!$, give

$$\frac{x^{\mathsf{T}} C(A)^{\mathsf{T}} x}{\operatorname{per} A\, x^{\mathsf{T}} A x} \longrightarrow \frac{2(n-3)!}{mn!} = \frac{2}{n(n-1)(n-2)^2}.$$

For example this is $1/24$ in order four, below the real rank-two constant $1/16$.

50 Complex rank two and the scope of the comparison

The complex analogue has no positive constant in any fixed order $n \geq 3$. This follows from a roots-of-unity example (Lemma 74). For $0 < t < 1$, let u consist of all n-th roots of unity and $A_t = (1-t)J + tuu^*$, where J is the all-ones matrix. Lemma 74 gives

$$u^* A_t u = n^2 t, \quad u^* C(A_t)^{\mathsf{T}} u = n\, n!\, t^{n-1}, \quad \operatorname{per} A_t = n! \big[(1-t)^n + t^n\big].$$

Thus the relative ratio in direction u equals

$$\frac{t^{n-2}}{n[(1-t)^n + t^n]} \longrightarrow 0.$$

The positive definite correlation matrices $(1-\varepsilon)A_t + \varepsilon I$, $\varepsilon \downarrow 0$, give the same vanishing limit in the unrestricted positive-definite complex class. In order two $C(A)^{\mathsf{T}} = A$ and $\operatorname{per} A = 1 + |A_{12}|^2$, so the complex sharp constant is $1/2$.

The sharp rank-two witnesses in Theorem 20 rule out a lower bound of the form $cn^{-s}A$ for absolute $c > 0$ and fixed s in the normalized raw cofactor comparison, even in real rank two. It also rules out that polynomial lower bound in the unrestricted real positive-definite class, by a sufficiently small diagonal interpolation of each sharp example. For each order the examples can be chosen as close to J as desired. Their largest normalized cofactor eigenvalue can therefore be kept bounded, since that eigenvalue is 1 at J.

51 The exponential scale of real cofactor domination

Let A be a real positive semidefinite matrix with unit diagonal, and put $C(A)_{ij} = \operatorname{per} A(i,j)$, where the minor deletes row i and column j. The permanental adjoint is $C(A)^T$. Marcus and Merris [4] give an upper comparison involving A^{-1}. For the reverse comparison with A, Theorem 21 obtains $(Kn)^{-Kn}$ in arbitrary real rank, Theorem 20 gives the exact real rank-two constant $1/(n2^{n-2})$, and Section 50 gives zero as the best complex constant for each $n \geq 3$. We determine the real exponential scale in arbitrary rank.

Theorem 22. *There is an absolute $K > 1$ such that all real correlation matrices of order $n \geq 2$ satisfy*

$$\frac{C(A)^T}{\operatorname{per} A} \succeq K^{-n} A.$$

The optimal all-rank real constant c_n satisfies

$$K^{-n} \leq c_n \leq 1/(n2^{n-2}), \qquad \log(1/c_n) = \Theta(n).$$

52 A chart normalized by the permanent

Take real unit Gram vectors v_i in $\mathbb{R}^r$, $r \leq n$, and let $L_i(Z) = v_i \cdot Z$ be their complex-linear forms. For a real test vector x, use the circular Gaussian norm and write

$$P(Z) = \prod_i L_i(Z), \quad D(Z) = \sum_i x_i \prod_{j \neq i} L_j(Z).$$

Gaussian pairings give the identities

$$\|P\|^2 = \operatorname{per} A, \qquad \|D\|^2 = x^T C(A)^T x, \qquad x^T A x = |\textstyle\sum_i x_i v_i|^2.$$

The orthogonal Gaussian monomials also give

$$|H(z)| \leq \|H\| \frac{\|z\|^d}{\sqrt{d!}}$$

for every homogeneous polynomial of degree d.

Choose a real unit direction e maximizing $|P|$ on the real unit sphere. This maximum is positive, as a finite union of proper real hyperplanes cannot cover the sphere. We claim

$$|P(e)| \geq 4^{-n} \|P\| / \sqrt{n!}. \tag{59}$$

For any complex unit $z = a + ib$, the frequency-n coefficient of $P(a\cos t + b \sin t)$ is $2^{-n} P(z)$, by homogeneity. The real curve lies in the unit ball since $\|a\|^2 + \|b\|^2 = 1$, so $|P(z)| \leq 2^n |P(e)|$. For a standard circular Gaussian vector, the direction is uniform on the complex unit sphere and independent of the radius. The radius has moment $\mathbb{E}\|Z\|^{2n} = (r)_n = n!\binom{r+n-1}{n}$. The binomial coefficient is at most 4^n since $r \leq n$. This proves (59).

Write $v_i = a_i(1, t_i)$ in orthonormal coordinates based at e. Then $a_i \neq 0$, t_i is real, and $a_i^2 = (1 + \|t_i\|^2)^{-1}$. With $\alpha_i = x_i / a_i$ put

$$R(z) = \frac{D(1,z)}{P(1,z)} = \sum_i \frac{\alpha_i}{1 + t_i \cdot z}.$$

Pad the transverse space by zeros to an odd dimension $d = 2m+1 \geq 3$ with $m \leq n$. Gaussian norms are unchanged since the polynomials ignore the added variables. For a real unit vector $\omega \in \mathbb{R}^d$ and $|y-1| \leq 1/4$, choose $s = \sqrt{y}$ of positive real part. For every real q,

$$|1 + isq| \geq \frac{\Re s}{|s|} \geq \frac{1}{2}, \qquad \|(1, is\omega)\| \leq \frac{3}{2}.$$

The first bound is obtained by minimizing the squared modulus over real q. Thus $|P(1, is\omega)| \geq 2^{-n}|P(e)|$. The evaluation inequality for D and (59) give

$$\sup_{\omega,\, |y-1|\leq 1/4} |R(i\sqrt{y}\,\omega)| \leq K_0^n \sqrt{n}\, \frac{\|D\|}{\|P\|} \tag{60}$$

for an absolute K_0, regardless of the sizes of t_i.

53 Recovering both Gram coordinates

For a real unit transverse direction u, let

$$J_0(y) = \mathbb{E}_\omega R(i\sqrt{y}\,\omega), \qquad J_1(y) = \frac{i}{\sqrt{y}} \mathbb{E}_\omega \big[(u \cdot \omega) R(i\sqrt{y}\,\omega)\big],$$

where ω is uniform on the real unit sphere. Both are holomorphic in a neighborhood of the disk from (60): the branch of $\sqrt{y}$ is holomorphic and the denominators have the same uniform lower bound on a slightly larger disk. Let $q_i = \|t_i\|$ and define

$$F_0(v) = \sum_i \frac{\alpha_i}{1 + q_i^2 v}, \qquad F_1(v) = \sum_i \frac{\alpha_i (u \cdot t_i)}{1 + q_i^2 v}.$$

Let $B = \mathrm{B}(1/2, m)$ be the beta integral. For real $y > 0$, the one-coordinate distribution of ω gives

$$J_j(y) = B^{-1} \int_0^1 z^{j-1/2} (1-z)^{m-1} F_j(yz)\, dz, \qquad j = 0, 1. \tag{61}$$

Indeed the density of the coordinate ξ in the direction t_i is $B^{-1}(1-\xi^2)^{m-1}$ on $[-1,1]$. Pairing its two signs gives $(1 + y q_i^2 \xi^2)^{-1}$ for J_0. For J_1 the conditional mean of $u \cdot \omega$ is $(u \cdot t_i)\xi / q_i$. The odd half of the rational factor is $-i\sqrt{y}\, q_i \xi / (1 + y q_i^2 \xi^2)$. Multiplication therefore adds the factor $(u \cdot t_i)\xi^2$. Terms with $q_i = 0$ give zero in J_1. The change of variable $z = \xi^2$ gives (61).

Multiplying (61) by $y^{m+j-1/2}$ and changing variables gives

$$y^{m+j-1/2} J_j(y) = B^{-1} \int_0^y (y-v)^{m-1} v^{j-1/2} F_j(v)\, dv.$$

Differentiating m times at $y = 1$ gives $(m-1)! F_j(1)/B$. This follows by differentiation of the polynomial kernel until the last step, which is the fundamental theorem of calculus; the endpoint at zero is integrable. The holomorphic expression on the left is bounded on the disk by a constant times $(5/4)^{m+1}$ times the right side of (60). Cauchy's bound for the mth derivative on the disk therefore yields

$$|F_j(1)| \leq C m B\, 4^m (5/4)^{m+1} K_0^n \sqrt{n}\, \frac{\|D\|}{\|P\|} \leq K_1^n \frac{\|D\|}{\|P\|}.$$

Here $B \leq 2$, $m \leq n$, and K_1 is absolute.

Finally $F_0(1) = \sum_i x_i a_i$ and $F_1(1) = \sum_i x_i a_i (u \cdot t_i)$, by the identity for a_i^2. These are the longitudinal and transverse components of the original Gram combination. Taking the supremum over unit u and squaring proves the claimed Loewner inequality for some absolute K; real tests suffice since the matrices are real symmetric. The upper bound on c_n follows from the rank-two example cited above.

54 Near-rank-one limits along orthogonal rays

For a real correlation matrix A of order n, put $C(A)_{ij} = \operatorname{per} A(i,j)$, where the minor deletes row i and column j. Define

$$r(A) = \inf_{x^T A x > 0} \frac{x^T C(A)^T x}{(\operatorname{per} A) x^T A x}.$$

The infimum of $r(A)$ over all real correlation matrices is the optimal all-rank constant c_n. The rank-two infimum is $1/(n2^{n-2})$ Theorem 20, and the same reference gives an upper witness based on a regular simplex Section 49. Here the Gram vectors may have higher rank while all converge to one unit vector.

Let $d = n - 1$ and $m_1, \dots, m_s$ be positive integers of sum d. For each j choose distinct positive numbers $t_{j,1}, \dots, t_{j,m_j}$, and let the slopes $t_i \in \mathbb{R}^s$ be zero and the vectors $t_{j,l} e_j$. For $\epsilon > 0$, take the Gram matrix A_ϵ of the vectors $v_i = a_i(1, \epsilon t_i)$ with $a_i = (1 + \epsilon^2 \|t_i\|^2)^{-1/2}$. The rank is $s+1$ and the entries are positive.

Theorem 23. *For this family,*

$$\lim_{\epsilon \downarrow 0} r(A_\epsilon) = \frac{1}{n \max\{E, O_1, \dots, O_s\}},$$

where

$$E = 1 + \sum_{j=1}^{s} \sum_{\substack{2 \le k \le m_j \\ k \text{ even}}} \binom{d}{k}, \qquad O_j = \sum_{\substack{1 \le k \le m_j \\ k \text{ odd}}} \binom{d}{k}.$$

Proof. Let $X, Y_1, \dots, Y_s$ be independent standard circular complex Gaussian variables, with $Y = (Y_1, \dots, Y_s)$. Use their Gaussian polynomial norm. With $L_i = a_i(X + \epsilon t_i \cdot Y)$, $P = \prod_i L_i$, and $D_x = \sum_i x_i \prod_{h \ne i} L_h$, the Gaussian identities Section 47 give

$$\|P\|^2 = \operatorname{per} A_\epsilon, \quad \|D_x\|^2 = x^T C(A_\epsilon)^T x, \quad \left\| \sum_i x_i v_i \right\|^2 = x^T A_\epsilon x.$$

Writing $x_i = a_i \alpha_i$, one has

$$D_x(1, Y) = P(1, Y) \sum_i \frac{\alpha_i}{1 + \epsilon t_i \cdot Y}, \qquad \sum_i x_i v_i = \sum_i \frac{\alpha_i (1, \epsilon t_i)}{1 + \epsilon^2 \|t_i\|^2}.$$

We choose a test basis depending on ϵ. One test uses the central slope alone with $\alpha_0 = 1$; its two limits are $D_x \to X^d$ and $\sum_i x_i v_i \to e_0$, where e_0 is the longitudinal unit vector. For every j and $1 \le k \le m_j$, solve for weights w_l on axis j with

$$\sum_{l=1}^{m_j} w_l t_{j,l}^h = \mathbf{1}_{\{h=k\}}, \qquad 1 \le h \le m_j.$$

The matrix is a Vandermonde with invertible column scalings since all nodes are distinct and nonzero. Put $\alpha_{j,l} = \epsilon^{-k} w_l$, set the central coefficient to $-\epsilon^{-k} \sum_l w_l$, and put all other coefficients to zero. The noncentral coordinates on each axis form an invertible scaled Vandermonde inverse, so these tests and the central one form a basis.

For the (j, k) test, expansion at $X = 1$ shows that the first nonzero moment in the rational sum is degree k. On any fixed compact polydisc of Y, the terms past degree k are $O(\epsilon)$ after scaling, since the denominators are uniformly bounded away from zero for small ϵ. As $P(1, Y) \to 1$, coefficient extraction and homogenization give $D_x \to (-1)^k X^{d-k} Y_j^k$ in Gaussian norm. In the Gram combination the longitudinal powers of $\epsilon t_{j,l}$ are even and the transverse powers are odd. Expanding the displayed reciprocal shows that the limit is $(-1)^{k/2} e_0$ for even k and $(-1)^{(k-1)/2} e_j$ for odd k. All lower-degree terms cancel, and the next powers beyond k vanish in the limit.

The limiting deletion polynomials are orthogonal, with squared norms $d!$ and $(d-k)!k!$. Thus their Gram matrix in the moving basis converges to a positive diagonal matrix. The Gram-vector map converges to the coordinate map just computed. Its squared operator norm relative to the deletion norm therefore converges, by continuous matrix inversion. The limiting coordinate functionals act on disjoint subsets of an orthogonal basis, with squared norms $E/d!$ and $O_j/d!$. Hence the squared operator norm tends to their maximum. This norm is $\sup_{x \neq 0} x^T A_\epsilon x / \|D_x\|^2$ for sufficiently small ϵ. Since $\operatorname{per} A_\epsilon \to n!$, taking its reciprocal and dividing by the permanent gives the result. □

55 Optimizing the ray counts

Theorem 24. *For $d \geq 4$, the minimum of the limits in Theorem 23 over positive counts of sum d is*

$$\frac{1}{n(2^{d-1} + I_d - 1)}, \qquad I_d = \begin{cases} \binom{d}{a} & d \text{ even}, \\ \binom{d-1}{a} & d \text{ odd}, \end{cases}$$

where a is a nearest even integer to $d/2$ with $0 < a < d$. Two rays of counts $a, d - a$ attain this limit.

Proof. For every pattern $O_j \leq 2^{d-1}$. To maximize E, put $\ell_j = \lfloor m_j/2 \rfloor$ and $G(b) = \sum_{k=1}^{b} \binom{d}{2k}$. The sum of the ℓ_j is at most $L = \lfloor d/2 \rfloor$. Because G increases, it suffices to optimize $\sum G(\ell_j)$ at total L. Every list of positive budgets of sum L can be realized by even ray counts $2\ell_j$, putting the possible remaining odd node on one ray.

If at least three positive budgets remain, take the smallest two $x \leq y$. Then $x + 2y \leq L$. By symmetry and unimodality of the binomial coefficients,

$$G(x + y) - G(x) - G(y) = \sum_{k=1}^{y} \left[\binom{d}{2(x+k)} - \binom{d}{2k} \right] \geq 0,$$

since $2k \leq 2(x + k) \leq d - 2k$. Repeated merging leaves at most two positive budgets without decreasing the value. The one-ray value is 2^{d-1}.

For two rays write the counts as $a, d - a$, with a even; if d is odd put its extra parity node on the second ray. Reflect the second even binomial sum through $k \mapsto d - k$. For even d, the two sums then cover all even indices and both count a, giving $2^{d-1} + \binom{d}{a} - 1$. For odd d, the reflected sum runs over odd indices above a, and the value becomes

$$2^{d-1} + \sum_{k=0}^{a} (-1)^k \binom{d}{k} - 1 = 2^{d-1} + \binom{d-1}{a} - 1.$$

The equality is Pascal's identity with cancellation. Both expressions are largest at the indicated balanced even cutoff. For odd d, the nearest allowable even integer to $d/2$ is also a nearest even integer to $(d-1)/2$, with ties permitted for the latter. The added binomial coefficient exceeds one for $d \geq 4$, proving the comparison with one-ray and transverse values. □

The two-ray family has rank three, so the theorem supplies the advertised strict improvement over the sharp rank-two comparison for every $n \geq 5$. The published simplex bound $2/[n(n-1)(n-2)^2]$ is stronger in orders 5 through 8, and the two-ray bound is stronger from order 9. For $d = 4, \ldots, 9$ this follows by evaluating the two denominators. For $d \geq 10$, $2^{d-1} > d(d-1)^2/2$ because the inequality holds at $d = 10$ and the polynomial grows by a factor less than two at each step. In order 4 the public simplex example already has rank three. The balanced binomial coefficient has order $2^{d-1}/\sqrt{d}$, which gives the stated relative improvement. This exact optimization pertains to the orthogonal-ray degenerations.

56 Sharp small-order bounds under phase and symmetry constraints

For a complex correlation matrix A, let $C(A)_{ij}$ be the permanent of the minor formed by deleting row i and column j. Write

$$\Lambda(A) = \frac{\lambda_{\max}(C(A)^{\mathsf{T}})}{\operatorname{per} A}.$$

The cofactor matrix is Hermitian positive semidefinite and the denominator is positive, by Lemma 69. Both the quotient and its defining spectrum are invariant under diagonal unitary conjugation and coordinate permutations. The unrestricted maximum in order three is Theorem 73. The unrestricted large-order maximum is Theorem 12; here we settle structured cases at orders three and four.

57 A phase constraint in order three

Theorem 25. *For a complex order-three correlation matrix with* $\operatorname{Re}(A_{12}A_{23}A_{31}) \leq 0$*,*

$$\Lambda(A) \leq \frac{5+\sqrt{10}}{6}.$$

Equality holds exactly for equal-modulus matrices whose three cyclic correlations, after diagonal unitary conjugation, are $re^{\pm i\pi/6}$*, with* $r = (\sqrt{10}-1)/(3\sqrt{3})$*.*

Proof. Use the order-three equalization of Lemma 72. It holds at fixed squared-modulus sum and fixed real triangle product. For clarity, if the squared moduli are x, y, z, with $s = x+y+z$, $p = xy+xz+yz$, $q = xyz$ and real triangle product u, the first two cofactor characteristic coefficients depend only on s, u. The determinant depends on the remaining modulus variables through $4(p-q)$. Since $p - q = (1-x)(1-y)(1-z) + s - 1$, equalizing x, y, z maximizes that determinant, strictly unless they already agree. The real triangle product can be kept because the product of the squared moduli increases. The principal minor constraints are preserved, including $1 - s + 2u \geq 0$. Increasing the cofactor determinant with the first two coefficients fixed strictly increases its largest root. The permanent $1 + s + 2u$ is unchanged.

For the equal-modulus matrix choose the cyclic entries $be^{i\theta}$ with $|\theta| \le \pi/3$, and put $a = b\cos\theta$, $d = b^2$. Direct Fourier diagonalization gives cofactor eigenvalues

$$1 - b^2 + 2bc_j + 4b^2c_j^2, \qquad c_j = \cos(\theta + 2\pi j/3).$$

The maximum is at $c_0 = \cos\theta$: the quadratic is convex in c_j, and the maximum and minimum cosines sum to at least $-1/2$, so the value at the larger one is at least that at the smaller one for $b \le 1$. Thus the quotient is

$$\frac{1 - d + 2a + 4a^2}{1 + 8a^3 + 3(1-2a)d}.$$

Its derivative in d has the sign of $16a^3 - 4$. The nonpositive triangle product and the determinant condition give $d \le 1/3$ and hence $a \le 1/\sqrt{3} < 4^{-1/3}$. The case $a = 0$ has $b = 0$. Otherwise the product $4a^3 - 3ad \le 0$ gives $d \ge 4a^2/3$. The quotient therefore increases on replacing d by $4a^2/3$. The resulting value at $w = 2a$ is

$$\frac{1 + w + (2/3)w^2}{1 + w^2},$$

whose maximum is the top eigenvalue of $\left(\begin{smallmatrix} 1 & 1/2 \\ 1/2 & 2/3 \end{smallmatrix}\right)$. Its positive optimizing ratio is $w = (\sqrt{10}-1)/3$, which gives the stated r, θ. The corresponding correlation matrix is feasible because $3d < 1$ and the real triangle product is zero. The preceding strict comparisons give the equality conditions. □

58 Phase positivity in unrestricted order four

Theorem 26. *If an order-four correlation matrix satisfies $\Lambda(A) \ge 329/200$, all its principal triangles have positive real correlation product. In a diagonal phase gauge every off-diagonal entry then has positive real part.*

Proof. Write $\lambda = \Lambda(A)$, and for a unit top eigenvector let q_i be the squared modulus of its i-th coordinate. For the i-th principal deletion put $x_i = \operatorname{per} A_i / \operatorname{per} A$ and $\gamma_i = \Lambda(A_i)$. Lemma 70 shows that x_i is the normalized cofactor diagonal and the squared norm of the normalized off-diagonal row is at most $\gamma_i x_i(1 - x_i)$. Cauchy–Schwarz applied to the eigenvector row gives

$$(\lambda - x_i)^2 q_i \le \gamma_i x_i (1 - x_i)(1 - q_i).$$

For $\lambda > 1$, direct differentiation shows that $x(1-x)/(\lambda - x)^2 \le 1/(4\lambda(\lambda-1))$ on $[0,1]$. Consequently $q_i \le \gamma_i/(4\lambda(\lambda-1) + \gamma_i)$.

The unrestricted order-three maximum is below $178/125$. Using its scalar formula from Theorem 73, the inequality is equivalent to positivity of $p(a) = 53 - 250a + 159a^2 + 356a^3$ for $0 \le a \le 1$. Here $p'' \ge 318$, while $p(9/25) > 1/5$ and $|p'(9/25)| < 3$. The quadratic tangent bound gives $p > 1/5 - 9/636 > 0$.

If a child triangle had nonpositive real product, Theorem 25 would instead give $\gamma_i < 137/100$ for that child. The cap on its coordinate and the other three would imply

$$1 = \sum_i q_i < 3\,\frac{14240}{56681} + \frac{13700}{56141} < 1,$$

since $4\lambda(\lambda - 1) \ge 42441/10000$. Thus every triangle product has positive real part. Make the entries from one coordinate positive real by diagonal phases. For every remaining pair, the triangle through the chosen coordinate then gives positive real part of the corresponding entry. □

59 Reduction in order four

Put

$$P_4(t) = 1 + 6t^2 + 8t^3 + 9t^4, \quad M_4(t) = 1 + 3t + 9t^2 + 11t^3, \quad F(t) = M_4(t)/P_4(t).$$

These are the scalar permanent, top cofactor eigenvalue, and their quotient for positive equicorrelation. On $[0,1]$ we have $1 \le F(t) < 2$ by subtracting the polynomials. Differentiation gives $F' = -3E/P_4^2$, with

$$E(t) = 33t^6 + 54t^5 + 29t^4 + 28t^3 + 3t^2 - 2t - 1.$$

The single sign change and the opposite signs at zero and one give a unique positive root $t_4 \in (0,1)$ and the unique scalar maximum.

Theorem 27. *On order-four complex correlations invariant under a transitive group of ordinary coordinate permutations, the exact maximum is $F(t_4)$. Equality holds precisely on the positive scalar optimizer's diagonal unitary orbit within this class.*

Direct evaluation gives $329/200 < F(5/16)$, so Theorem 26 applies to every maximizer, including in the unrestricted class.

A transitive subgroup of S_4 has order $4, 8, 12$, or 24. At order four it is regular and contains a 4-cycle or is the regular Klein four subgroup. An order-eight subgroup is a Sylow subgroup containing a 4-cycle. At order twelve the subgroup is A_4 and contains the regular Klein four subgroup. The full group contains both. We analyze the resulting two families.

60 Klein four symmetry

Translation invariance under the regular Klein group and Hermitian symmetry give three real matching activities x, y, z. Their product is positive, so a diagonal character makes them all positive. The cofactor has nonnegative entries and constant row sum, which is its largest eigenvalue. The permanent with variable diagonal d_0 is

$$d_0^4 + 2d_0^2(x^2 + y^2 + z^2) + 8d_0xyz + (x^2 + y^2 + z^2)^2.$$

Differentiating along $J/4$ gives the row sum $M = 1 + s_1 + s_1^2 + s_1s_2 + 2p$, where $s_1 = x + y + z,\ s_2 = x^2 + y^2 + z^2,\ p = xyz$. Put $t = s_1/3$, $h_i = x_i - t$, $U = \sum h_i^2$, and $V = h_1h_2h_3$. Then

$$\begin{aligned} P &= P_4(t) + (2 - 4t + 6t^2)U + U^2 + 8V, \\ M &= M_4(t) + 2tU + 2V. \end{aligned}$$

The cubic with roots h_i has discriminant $U^3/2 - 27V^2$. Thus $|V| \le U^{3/2}/(3\sqrt{6})$. If $U > 0$, square completion in $\sqrt{U}$ gives, writing $F = F(t)$,

$$\frac{FP - M}{U} \ge F(2 - 4t + 6t^2) - 2t - \frac{(8F - 2)^2}{216F} > 0.$$

Indeed the first two terms are at least $2 - 6t + 6t^2 \ge 1/2$ and the final term is at most $49/108$ for $1 \le F \le 2$. Thus only the scalar family reaches its own mean's quotient.

61 Complex cyclic symmetry

The first row of a Hermitian circulant is $1, x+iy, b, x-iy$, with b real. A diagonal character interchanges x and y up to signs and negates b for an odd character. Positive triangle product $b(x^2-y^2)$ therefore allows a gauge with $b>0$ and $x>|y|$. Let $v=y^2$. The circulant eigenvalues $1+b\pm 2x, 1-b\pm 2y$ give $x,b\in[0,1]$, $v\in[0,1/4]$.

Cycle enumeration for variable diagonal d_0 gives

$$P(d_0)=d_0^4+d_0^2(4(x^2+v)+2b^2)+8d_0b(x^2-v)+4(x^2-v)^2+4b^2(x^2+v)+b^4.$$

Differentiating along $J/4$ at $d_0=1$ gives the all-ones cofactor eigenvalue

$$M=1+2x+b+4x^2+b^2+4bx+4x^3+4bx^2+2b^2x+b^3-4xv.$$

This is the largest Fourier cofactor eigenvalue in that gauge. The opposite mode replaces x by $-x$, lowering the result by $4x((1+b)^2+2x^2-2v)>0$. For $u=|y|<x$ the other two replace (x,b,v) by $(\pm u,-b,x^2)$. In comparison with the positive mode at u, the mixed cubic difference is $4xu(x-u)$; in comparison with the negative mode at $-u$, the potentially signed cubic terms combine to $4(x+u)(x-u)^2$. All other terms in each difference are nonnegative, including a $2b$ from the linear terms.

For $0\le x,b\le 1/2$ put $t=(2x+b)/3$ and $h=x-t$. Then $b=t-2h$ and direct expansion gives

$$\begin{aligned}P(1)&=P_4(t)+4h^2(3-6t+9t^2-4h+9h^2)+4v((1-b)^2-2x^2+v),\\ M&=M_4(t)+4h^2(3t-h)-4xv.\end{aligned}$$

The $4h^2$ coefficient in $FP-M$ is at least $3-9t+9t^2-3h+9h^2\ge 1/2$. The $4v$ coefficient is at least $F(1/4-2x^2)+x$, strictly positive at both endpoints $0,1/2$ and hence on the interval by concavity. Therefore $M/P(1)\le F(t)$ with equality exactly at $h=0=v$.

The remaining domain has $x\ge 1/2$ or $b\ge 1/2$. On the following boxes with $v\in[0,1/4]$, the indicated minimum tensor Bernstein coefficient of $41P(1)-25M$ is obtained at degree $(4,4,2)$:

x-interval	b-interval	Minimum
$[1/2,1]$	$[0,1]$	$1/4$
$[0,1/4]$	$[1/2,1]$	$117/64$
$[1/4,1/2]$	$[1/2,1]$	$71/64$

For these exact numbers, expand the displayed polynomial after the affine mapping of the box onto the unit cube, then use $s^j=\sum_{i=j}^n \binom{i}{j}B_{i,n}(s)/\binom{n}{j}$ in each coordinate. The supplied rational checker reconstructs the coefficients using this formula. Positivity of the basis gives a strict outer bound $M/P(1)<41/25<F(5/16)$. Together with the Klein comparison and uniqueness of t_4, this proves Theorem 27.

62 The exact maximum of the permanental cofactor ratio in order four

For a complex correlation matrix of order four, we determine the largest possible ratio of the top permanental cofactor eigenvalue to the permanent. Positive equicorrelation attains the maximum, approximately 1.6452266199, and every maximizer is obtained from it by diagonal unitary conjugation. The proof combines two stationary bordering inequalities to recover the scalar child quotient after averaging over the four principal deletions. This average controls the magnitude and phase terms in the full permanent at the exact scalar threshold.

For a matrix A, write $C(A)_{ij} = \operatorname{per} A(i \mid j)$ for its permanental cofactors, with $A(i \mid j)$ obtained by deleting row i and column j. For a correlation matrix, meaning a Hermitian positive semidefinite matrix with diagonal one, put

$$H(A) = C(A)^{\mathsf{T}} / \operatorname{per} A, \qquad \Lambda(A) = \lambda_{\max}(H(A)).$$

The deleted-product Gram representation gives $H(A) \succeq 0$, and the permanent is positive. Lemma 71 gives the leading large-order behavior. The exact positive equicorrelation theorem of Sections 36–41 applies in sufficiently large orders. Sections 56–61 prove the order-four result under a transitive input symmetry up to diagonal phases and acuteness of triangle products on a high order-four level, and Theorem 73 gives the unrestricted order-three result. The theorem here removes the symmetry condition in order four.

Theorem 28. *For every complex correlation matrix A of order four,*

$$\Lambda(A) \le F_4(b) = \frac{1 + 3b + 9b^2 + 11b^3}{1 + 6b^2 + 8b^3 + 9b^4}, \tag{62}$$

where b is the unique positive root of

$$R(b) = 33b^6 + 54b^5 + 29b^4 + 28b^3 + 3b^2 - 2b - 1 = 0. \tag{63}$$

Equality holds precisely on the diagonal-unitary orbit of $(1-b)I + bJ$, where J is the all-ones matrix.

The proof first excludes the boundary, then uses stationarity to bound the mean child quotient at the scalar value. The degree-four term needs a product of child triangle contributions, obtained from concavity in a logarithmic coordinate and an identity relating the four triangle phases to the three Hamiltonian phases. Equality in the final permanent comparison forces all magnitudes and phases.

62.1 Stationary bordering at a maximizer

Put $L = F_4(b)$ for the quotient in the theorem. For $t > 0$, the scalar cofactor has a positive constant off-diagonal, so its top eigenvector is constant. The scalar quotient is $F_4 = M_4/P_4$, where M_4, P_4 are the numerator and denominator in (62). Its derivative is $-3R/P_4^2$, where R is the polynomial in (63). Descartes' rule of signs and the sign change on $(.3131277889, .3131277890)$ isolate the unique positive root and give the scalar maximum there. The boundary estimate in Section 62.6 applies since $L \ge F_4(5/16) > 4113/2500$. Compactness gives a global maximizer A, whose quotient λ is at least L.

For each principal deletion A_i, define $P = \operatorname{per} A$, $x_i = \operatorname{per} A_i / P$, $\lambda_i = \lambda_{\max}(H_i)$, and $q_i = |e_i|^2$, where $H_i = C(A_i)^{\mathsf{T}} / \operatorname{per} A_i$ and e is any top unit cofactor eigenvector of A. At this maximizer we have

$$x_i e_{-i}^* H_i e_{-i} = \lambda(x_i - q_i), \qquad (\lambda - x_i)^2 q_i \le \lambda(1 - x_i)(x_i - q_i).$$

To see the first identity, hold e fixed. The Rayleigh quotient $e^* H(A) e$ is stationary under off-diagonal variations because it is bounded above by $\Lambda(A)$ and equals it at the maximizer. A positive diagonal congruence $A \mapsto DAD$ sends $H(A)$ to $D^{-1} H(A) D^{-1}$. Vary the i-th diagonal factor at speed $1/2$, so the diagonal entry of A changes at speed one. The off-diagonal derivative vanishes. The congruence derivative of the Rayleigh quotient is $-\lambda q_i$. Direct differentiation in A_{ii} differentiates the permanent to $\operatorname{per} A_i$ and each cofactor to its child cofactor on the remaining indices, with zeros in row and column i. The quotient rule proves the first identity.

For the second identity, the exact expansion with the appended correlation column $z_i = A_{-i,i}$ gives $H(A)_{-i,i} = x_i H_i z_i$ and $x_i z_i^* H_i z_i = 1 - x_i$. The eigenvector equation and Cauchy–Schwarz with respect to $H_i \succeq 0$ then give the result. The same expansion gives the row projection bound of Lemma 66, namely $q_i \le \lambda_i/(\lambda_i + 4\lambda(\lambda - 1))$.

Use $g = 1.42285 > \gamma_3$: the cubic $q(a) = 1.42285\, D(a) - N(a)$, with N, D as in Theorem 73, has $q'' > 2.53$ on $[0,1]$, $q(0.35772) > 5 \cdot 10^{-7}$ and $|q'(0.35772)| < 10^{-5}$, so $q > 0$ on $[0,1]$, and Theorem 73 gives $\gamma_3 = \max_a F_3(a) < g$. Set $Q = g/(g + 4L(L-1))$ and $q_0 = 1 - 3Q$. Then $q_0 \le q_i \le Q$, and the diagonal identity and the child upper bound give

$$x_i \le \frac{\lambda q_i}{\lambda - \lambda_i(1 - q_i)} \le \frac{LQ}{L - g(1-Q)} < .714.$$

Here the fraction increases in q_i and λ_i and decreases in λ. Permanent monotonicity under correlation bordering gives $x_i \in [0,1]$. The Cauchy bound gives $q_i \le \psi(x_i)$, where

$$\psi(x) = \frac{Lx(1-x)}{(L-x)^2 + L(1-x)}.$$

In using $\lambda \ge L$, the relevant denominator divided by λ is $\lambda + x^2/\lambda + 1 - 3x$, increasing for $\lambda > 1$ and $x \in [0,1]$. The function ψ increases on $[0, .714]$, and $\psi(.6879) < q_0$. Thus $x_i \in (.6879, .714)$. On this shorter interval ψ is strictly concave. The function

$$h(x) = \frac{L(L-x)^2}{(L-x)^2 + L(1-x)^2}$$

is increasing and strictly convex there. The elementary rational signs are recorded in Section 62.4.

Let $P_* = 1 + 6b^2 + 8b^3 + 9b^4$ and $x_* = (1 + 3b^2 + 2b^3)/P_*$. At the scalar matrix, the normalized cofactor's top eigenvector is $\mathbf{1}/2$, and the quotient is stationary in all off-diagonal directions by permutation invariance, conjugation invariance and $F_4'(b) = 0$. Both displayed stationary relations are equalities there, since the appended column is proportional to e_{-i}. Consequently

$$\psi(x_*) = \tfrac{1}{4}, \qquad h(x_*) = K = F_3(b), \qquad F_3(a) = \frac{1 + 2a + 3a^2}{1 + 3a^2 + 2a^3}.$$

The same bounds $x_* \in (.6879, .714)$ apply at this stationary point, using each of its hypotheses above.

62.2 Two averaged bounds at the scalar threshold

Since the q_i sum to one, Jensen gives $\psi(\bar{x}) \ge \frac{1}{4}\sum_i \psi(x_i) \ge 1/4$. Monotonicity gives $\bar{x} \ge x_*$.

For each q_i, let $y = \psi^{-1}(q_i)$ on $[.6879, .714]$. This exists because $\psi(.6879) < q_0 \le q_i \le \psi(x_i)$. The diagonal identity and Rayleigh's upper bound give

$$\lambda_i \ge \frac{\lambda(x_i - q_i)}{x_i(1 - q_i)} \ge \frac{L(y - q_i)}{y(1 - q_i)} = h(y).$$

The middle fraction increases with x_i at fixed q_i, and $y - q_i \ge 0$ since $\psi(y) \le y$. The inverse of an increasing concave function is convex. Since h is increasing and convex on the required interval, $h \circ \psi^{-1}$ is convex. It follows that

$$\bar{\lambda}_i \ge K, \qquad \lambda_i > h(.6879) > E = 1.4003.$$

Only these two bounds, and $\bar{x} \ge x_*$, are needed below.

62.3 Compatibility of the full permanent

In child i, let s_i be its sum of squared edge moduli and τ_i its real triangle product. The inequality $\lambda \geq L > 329/200$ puts A in the range of the acute-phase theorem of Section 58. Replace the child by equal moduli at fixed s_i, τ_i as in Lemma 72. Its principal phase coordinate a_i and nonnegative excess ζ_i satisfy

$$s_i = 3(a_i^2 + \zeta_i), \qquad \tau_i = a_i^3 - 3a_i\zeta_i, \qquad \lambda_i \leq \frac{N(a_i) - \zeta_i}{D(a_i) + 3(1 - 2a_i)\zeta_i} \leq F_3(a_i),$$

where $N = 1 + 2a + 3a^2$ and $D = 1 + 3a^2 + 2a^3$. The replacement has common modulus $\sqrt{a_i^2 + \zeta_i}$ and principal triangle argument of magnitude $3\arctan(\sqrt{\zeta_i}/a_i)$. It keeps the real triangle product and increases its modulus. In particular $y_i = \zeta_i/a_i^2$ belongs to $[0, 1/3]$, and the magnitude of the replacement argument is at least that of the original. The high phase-coordinate branch $a_i \geq 4^{-1/3}$ has quotient at most $4/3$, so $\lambda_i > E$ excludes it. On the remaining branch the displayed upper bound by F_3 follows by comparing the replacement quotient at ζ_i and at zero. Unimodality and the tests $F_3(.274) < E$ and $F_3(.452) < E$ give $a_i \in (.274, .452)$. See Lemma 72 and Theorem 73 for the replacement and its two branches, and Section 58 for the acute phase bound.

On this interval both $a \mapsto F_3(a)$ and $t \mapsto F_3(e^t)$ are concave. With $F = F_3$,

$$F''(a) = \frac{24a(a^5 + 2a^4 + 5a^3 + 2a^2 - 2a - 2)}{D(a)^3},$$

and

$$(aF'(a))' = \frac{2(a-1)(6a^6 + 13a^5 + 49a^4 + 52a^3 + 26a^2 - a - 1)}{D(a)^3}.$$

The first parenthesis is negative since its positive terms sum to less than two. The second is positive already from $26(.274)^2 - .452 - 1 > 0$. Therefore the arithmetic and geometric means of the a_i are at least b, since F_3 is increasing below b and the mean child quotient is at least $F_3(b)$. In particular, if u is the squared moduli sum over the six edges, then $u = \frac{1}{2}\sum_i s_i \geq 6b^2$.

For the sum of real triangle products, put $k = 3b^2/F_3'(b)$, so $1.03 < k < 1.05$. If $a_i \leq .39$, rearrange the child quotient to obtain

$$3a_i\zeta_i \leq \frac{3a_iD(a_i)}{1 + 3E(1 - 2a_i)}(F_3(a_i) - \lambda_i) \leq k(F_3(a_i) - \lambda_i).$$

The multiplier increases in a_i and at .39 is less than .96. The function $a^3 - k(F_3(a) - K)$ is strictly convex and has derivative zero at b. Thus $\tau_i \geq b^3 + k(\lambda_i - K)$. If $a_i \geq .39$, clearing the same quotient gives $\tau_i > .04 > b^3 + k(g - K)$, by the polynomial in Section 62.4. The inequality for τ_i holds in both ranges. Hence $v = \sum_i \tau_i \geq 4b^3$.

62.3.1 The degree-four contribution

Let M denote the sum of the three disjoint-transposition weights and the Hamiltonian terms in the permanent. We first control the product of the four real triangle products. The deficit of the replacement quotient from $F_3(a)$ at $y = \zeta/a^2$ is

$$\frac{(4 - 16a^3)a^2y}{D(a)[D(a) + 3(1 - 2a)a^2y]}.$$

Bounding the second denominator with $y \le 1/3$ gives $F_3(a_i) - \lambda_i \ge 3y_i/20$, because the resulting inequality after clearing denominators reduces to positivity of

$$-\tfrac{344}{3}a^5 - 12a^4 - 2a^3 + \tfrac{59}{3}a^2 - 1.$$

Its derivative is $-2a(860a^3+72a^2+9a-59)/3$, so its minimum on the interval occurs at an endpoint. Both endpoint evaluations are positive. For the geometric mean a_{g} of the a_i this implies

$$\prod_i \tau_i = a_{\mathrm{g}}^{12} \prod_i (1 - 3y_i) \ge a_{\mathrm{g}}^{12}\big[1 - 80(F_3(a_{\mathrm{g}}) - K)\big] \ge b^{12}.$$

For the first inequality use $\prod_i(1 - 3y_i) \ge 1 - 3\sum_i y_i$, the child deficit bound and Jensen in the logarithmic coordinate. For the last inequality the bracket is positive by $F_3 \le g$. The derivative of the preceding expression has the sign of

$$Q(a) = 12(1 + 80K)D(a)^2 - 960N(a)D(a) - 80a(N'D - ND')(a).$$

Every Bernstein coefficient of Q is positive on $[.3131277889, .452]$ at $K = 1.41669$, a rational lower bound for the actual K. In fact, the coefficients all exceed four. The checker described in Section 62.4 computes them by expanding $Q(\ell + (r - \ell)t) = \sum_j u_j t^j$ and taking $\sum_{j=0}^{i} u_j \binom{i}{j} / \binom{6}{j}$. Increasing K increases Q by a positive multiple of D^2. Since $a_{\mathrm{g}} \ge b$, the value is at least the value b^{12} at b.

Orient the four triangles as $(123), (124), (134), (234)$, and let $\phi_1, \dots, \phi_4$ be their principal arguments. They lie in $(-\pi/2, \pi/2)$, so their alternating signed sum, a multiple of 2π, is zero. Gauge the phases at vertex 1 to zero and take the remaining edge angles from the triangles through vertex 1. The real curls of these angles are the four principal arguments. Adding a vertex gradient, if desired, makes the angle row sums zero without changing any curl. Let p, q, r be the real circulations along $(1234), (1243), (1324)$. Direct addition of oriented edges gives the four triangle circulations

$$(p + q - r)/2, \quad (p + q + r)/2, \quad (p - q + r)/2, \quad (p - q - r)/2.$$

Thus $p^2 + q^2 + r^2 = \sum_i \phi_i^2$. Each squared triangle argument is at most $9y_i$. Therefore $p^2 + q^2 + r^2 \le 240(g - K) < 2$, all three Hamiltonian cosines are positive, and their product is at least $1 - 120(g - K) > .55^3$. With $X = \cos p,\ Y = \cos q,\ Z = \cos r,\ w = (XYZ)^{1/3}$, the product T of the four triangle cosines satisfies

$$T = \tfrac{1}{4}(X^2 + Y^2 + Z^2 + 2XYZ - 1) \le \tfrac{1}{4}(1 + w^3)^2, \qquad 1 + 2w \ge 3T^{1/3}.$$

To obtain the identity, pair the two products of cosines with sums of arguments p and use the addition formula. The bound on T uses $X^2 + Y^2 + Z^2 \le 2 + (XYZ)^2$, which follows from $1 - \prod(1 - v_j) \le \sum v_j$ for $v_j \in [0, 1]$. The last inequality follows on $[.55, 1]$ from

$$4(1 + 2w)^3 - 27(1 + w^3)^2 = (1 - w)(27w^5 + 27w^4 + 27w^3 + 49w^2 + w - 23) \ge 0.$$

The second factor increases and is positive at .55. If W is the product of the six edge moduli, the product of the three disjoint-transposition weights is W^2, as is the product of the three Hamiltonian magnitudes. Also $\prod_i \tau_i = W^2 T$. Applying the arithmetic–geometric mean inequality gives

$$M \ge 3W^{2/3}(1 + 2w) \ge 9(\textstyle\prod_i \tau_i)^{1/3} \ge 9b^4.$$

62.3.2 Saturation and equality

Now $P = 1 + u + 2v + M$ and $\sum_i \operatorname{per} A_i = 4 + 2u + 2v$ by expanding the permuted cycles. Write $S = 4x_* \in (2, 4)$. Since $\sum_i x_i \geq S$, one has

$$4 - S \;\geq\; (S-2)u + (2S-2)v + SM \;\geq\; (S-2)6b^2 + (2S-2)4b^3 + 9Sb^4 = 4 - S.$$

The last equality is the same expansion at the scalar matrix. Thus all coefficients have positive weights and all three bounds on u, v, M are equalities. In particular $u = 6b^2$ forces equality in $u = \frac{3}{2}\sum_i (a_i^2 + \zeta_i) \geq 6\bar{a}^2 \geq 6b^2$. All $a_i = b$ and all $\zeta_i = 0$. Each child has squared-moduli sum $3b^2$ and real triangle product b^3, forcing each modulus to equal b and each triangle phase to vanish. Diagonal phases making the edges at one vertex positive then make all edges positive. This identifies every maximizer and proves the theorem.

62.4 Exact rational bounds

All terminating decimals are rational. The scalar polynomial has opposite signs at .3131277889 and .3131277890. Enclosing the root in this interval gives

$$1.64522 < L < 1.64523, \quad F_3'(b) > 0, \quad K > 1.41669, \quad 1.03 < k < 1.05, \quad b^3 + k(g - K) < .04.$$

Let $J = (L-x)^2 + L(1-x)$ and $B = L^2 + L - 2L(L+1)x + (3L-1)x^2$. Then $\psi' = LB/J^2$ and the sign of ψ'' is the sign of $B'J - 2BJ'$. On $[0, .714]$, $B' < -3.08$ and $B(.714) > .143$. On $[.6879, .714]$, $J > 1.337,\ B < .227,\ -J' < 3.56$, so $B'J - 2BJ' < 0$. For $D_h = (L-x)^2 + L(1-x)^2$, the positive derivative $h' = 2L^2(L-x)(L-1)(1-x)/D_h^2$ satisfies

$$\frac{h''}{h'} = -\frac{1}{L-x} - \frac{1}{1-x} + \frac{4[(L-x) + L(1-x)]}{D_h} > .63$$

on the same shorter interval by taking the adverse endpoints in each term. The bounds on $\psi(.6879)$ and $LQ/[L - g(1-Q)]$ can be checked at $L = 1.64522$, as both decrease in L and q_0 increases in L. The lower bound for $h(.6879)$ is also checked there since h increases in L at fixed x.

For the high a_i triangle estimate let $z = a_i - .39$ and $D_E = 1 + 3E(1 - 2a_i)$. The polynomial $a_i^3 D_E - 3a_i(N(a_i) - ED(a_i)) - .04D_E$ has coefficients, from the constant term upward,

$$.0010038284, \quad .87405468, \quad 4.300212, \quad 8.8036.$$

It is positive for $z \geq 0$. The rational verifier in the source package checks these signs, the phase-polynomial endpoint tests and the Bernstein coefficients using the Python standard library.

62.5 A sharp permanent derivative bound

For a positive semidefinite matrix M of order four with positive diagonal and any positive semidefinite B, Theorem 28 gives

$$\frac{d}{dt} \operatorname{per}(M + tB)\Big|_{t=0} \leq F_4(b) \operatorname{per} M \sum_{i=1}^{4} \frac{B_{ii}}{M_{ii}}.$$

Indeed, with $D = \operatorname{diag}(\sqrt{M_{ii}})$ and $A = D^{-1}MD^{-1}$, the diagonal congruence rule gives $H(M) = D^{-1}H(A)D^{-1}$. The normalized derivative is $\operatorname{tr}(H(M)B)$, which is the trace of $H(A)$ against the positive semidefinite matrix $D^{-1}BD^{-1}$. For $B \neq 0$, equality requires that A lie on the maximizing orbit and that $D^{-1}BD^{-1}$ have range in its top cofactor eigenspace. These conditions also suffice. For the scalar maximizer this eigenspace consists of the constant vectors, so a rank-one B already attains the sharp constant.

62.6 A positive-definiteness bound

Lemma 29. *If an order-four correlation matrix satisfies $\Lambda(A) \geq 1.6452$, then*

$$A \succ .07I.$$

62.6.1 A row with small bordering loss

Let $g = \gamma_3$ denote the exact order-three maximum (Theorem 73), and put

$$U = \frac{1 + \sqrt{1 + 3g}}{2}, \qquad X = \frac{U}{2U - 1}.$$

The exact order-three scalar formula

$$F(a) = \frac{1 + 2a + 3a^2}{1 + 3a^2 + 2a^3}$$

gives $g \geq F(5/14)$ and $g < 1.423$. For the upper bound, the positive root of $3a^4 + 4a^3 + 6a^2 = 1$ lies in $(.35771, .35773)$. Bounding the numerator of F at the right endpoint and its denominator at the left gives a quotient below 1.423. It follows that $1.646 < U < 1.6478$ and $.717 < X < .719$.

Choose a row whose absolute sum in $H(A)$ is at least $\Lambda(A)$, and denote its principal deletion by T. Write

$$B = H(T), \quad \lambda = \lambda_{\max}(B), \quad x = \frac{\operatorname{per} T}{\operatorname{per} A}, \quad r = xBz, \quad D = \lambda x(1 - x) - \|r\|^2,$$

where z is the appended correlation column. The exact isolated-row identities also give $z^* Bz = (1 - x)/x$ and $D \geq 0$. The row's absolute sum is $x + \|r\|_1$.

The scalar upper envelope $f(x) = x + \sqrt{3g\, x(1 - x)}$ is maximized at X, with

$$U - f(x) \geq \sqrt{3g}\,(x - X)^2.$$

The entire row loss is at most $U - 1.6452 < .0026$, which gives $|x - X| < .036$ and thus $.681 < x < .755$. Its child-gap contribution is at least

$$\sqrt{3x(1 - x)}\,(\sqrt{g} - \sqrt{\lambda}) \geq \frac{\sqrt{3(.755)(.245)}}{2\sqrt{1.423}}(g - \lambda) \geq .312(g - \lambda).$$

Consequently

$$\lambda > F(5/14) - .0026/.312 > 1.4145.$$

The further contribution $\sqrt{3}(\sqrt{\lambda x(1 - x)} - \|r\|)$ bounds $\sqrt{3}\, D/(2\sqrt{\lambda x(1 - x)})$. Hence

$$D < .0026\sqrt{1.423/3} < .001792.$$

The exact identities are those of Lemma 70.

62.6.2 Spectral bounds on the child

Let $s = \sum_{i<j} |T_{ij}|^2$, $\tau = \mathrm{Re}(T_{12}T_{23}T_{31})$, and $R = s/3$. Order-three equal-modulus replacement at fixed s, τ can only increase the cofactor quotient (Lemma 72). It preserves $\operatorname{per} T = 1 + s + 2\tau$ and the original spectrum of T, since its trace, second characteristic coefficient and determinant are $3, 3 - s, 1 - s + 2\tau$. For the replacement write a for the phase coordinate in the exact order-three comparison. The replacement quotient and the smallest original eigenvalue are

$$G(a, R) = \frac{1 - R + 2a + 4a^2}{1 + 8a^3 + 3(1 - 2a)R}, \qquad \lambda_{\min}(T) = 1 - a - \sqrt{3(R - a^2)}.$$

The high phase branch $a \geq 4^{-1/3}$ has quotient at most $4/3$ (Theorem 73). In the remaining branch $G(a, R) \leq F(a)$. Since $G(a, R) > 1.4145$ and both $F(.30)$ and $F(.42)$ are below 1.4145, scalar unimodality gives $.30 < a < .42$. Its permanent therefore satisfies

$$\operatorname{per} T < \frac{1 + 2(.42) + 4(.42)^2}{1.4145} < 1.8.$$

Along the segment with a fixed from $R = a^2$ to the actual R, this denominator increases, and

$$-\partial_R G \geq \frac{4 - 16(.42)^3}{1.8^2} > .868.$$

As $F(a) - G(a, R) < 1.423 - 1.4145$, one has $R - a^2 < .01$. For a stronger spectral estimate, put $L = 1.4145$ and $\zeta = R - a^2$. The inequality $G(a, R) > L$ gives

$$\zeta < \frac{1 + 2a + 3a^2 - L(1 + 3a^2 + 2a^3)}{1 + 3L(1 - 2a)}.$$

The polynomial

$$(.51 - a)^2[1 + 3L(1 - 2a)] - 3[1 + 2a + 3a^2 - L(1 + 3a^2 + 2a^3)]$$

is the quadratic

$$\frac{881537}{50000}a^2 - \frac{135558387}{10^7}a + \frac{52146687}{2 \cdot 10^7}.$$

It is strictly positive, since its leading coefficient is positive and $4AC - B^2 = 11617321125831/10^{14} > 0$ for its coefficients A, B, C. As $a < .42$, it follows that $a + \sqrt{3\zeta} < .51$, or $T \succ .49I$.

Also $s > .27$ and $\tau = a^3 - 3a(R - a^2) > .018$, using monotonicity of $a^3 - .03a$ for $a > .30$. Thus

$$\operatorname{tr} B = \frac{3 + s}{1 + s + 2\tau} < \frac{3.27}{1.306} < 2.504.$$

The order-three comparison is from Sections 57–58.

For completeness, if a correlation matrix T has $T \succeq a_0 I$, its normalized cofactor satisfies $B \succeq a_0 \operatorname{diag}(B_{jj}^2)$. To see this in the Gaussian product representation, write the linear forms as $m_j + \sqrt{a_0} Z_j$ with independent standard complex coordinates Z_j and an independent Gaussian mean vector m. Conditional on m, put $w_j = a_0 + |m_j|^2$. The deleted-product Gram matrix divided by the conditional permanent is the sum of a rank-one matrix with coordinates m_j/w_j (or their conjugates according to the Gram convention) and the diagonal matrix with entries a_0/w_j^2. Under the probability tilted by $\prod_j w_j$, its mean is the normalized cofactor. Its diagonal entries are the tilted means of $1/w_j$, and Jensen gives the stated bound. Here each $B_{jj} \geq 1/\operatorname{per} T$, so $B \succ .151I$. Subtracting λ and the smallest eigenvalue from the trace bound gives $\lambda - \lambda_2(B) > .476$.

62.6.3 Schur-complement conclusion

Let $z = \alpha e + v$ with e a top unit eigenvector of B and $v \perp e$. The first bordering identity gives

$$|\alpha|^2 \le \frac{1-x}{x\lambda} < .332.$$

The reserve is exactly $D = x^2 z^* B(\lambda I - B)z$. It bounds

$$\|v\|^2 \le \frac{D}{x^2(.151)(.476)} < .054.$$

Thus $\|z\|^2 < .386$. In $A - .07I$ the principal block $T - .07I$ is greater than $.42I$, so its Schur complement is positive since

$$z^*(T - .07I)^{-1}z < .386/.42 < .93.$$

All terminating decimals in the proof denote exact rationals.

63 Permanent inequalities near weak-entry correlation matrices

We prove the Bapat–Sunder multiplier inequality for every complex correlation matrix in a dimension-independent Frobenius neighborhood of a correlation matrix whose off-diagonal entries are bounded by a fixed multiple of the reciprocal of the order. Equality requires the other factor to be the identity. Near a sharp permanental cofactor maximizer we prove the reverse fixed-factor comparison as well, by comparing its normalized Schur power to a mixture of subgroup projections for equicorrelation. The second normalized eigenvalue tends to one half along asymptotically sharp sequences. We also prove a strong self-conjugate inequality with exponential reserve when a principal block nearly maximizes the cofactor ratio and the complementary block has sufficiently small relative size.

63.1 Introduction

A correlation matrix is a Hermitian positive semidefinite matrix with diagonal one. We write per for the permanent, $\circ$ for entrywise multiplication, $|A|^2 = A \circ \bar{A}$ for the entrywise square of the moduli, and $C(A)_{ij} = \operatorname{per} A(i \mid j)$ for the raw cofactor matrix. Define $H(A) = C(A)^{\mathsf{T}}/\operatorname{per} A$ and $\Lambda(A) = \lambda_{\max}(H(A))$. The recursive upper bound of Lemma 71 is $c_1 = 1$, $c_n = (1 + \sqrt{1 + (n-1)c_{n-1}})/2$.

The Bapat–Sunder comparison bounds $\operatorname{per}(A \circ B)$ by the permanent of one input [28]. Counterexamples of Drury and Shchesnovich show that structural hypotheses are needed for both the inequality and permanent-on-top [29, 30]. There are positive results for structured supports and for multipliers satisfying a scaled diagonal dominance condition [31]. Sa-nguansin and Rodtes obtain self-conjugate Chollet inequalities under annular constraints on the entries [32]. For exact positive equicorrelation the comparison retaining $\operatorname{per} B$ can be read off from the convex-combination identity: the product is $(1-b)I+bB$, and its permanent is a convex combination of principal permanents of B, each at most $\operatorname{per} B$ by Lieb's inequality [1]. For related Hadamard-product inequalities see [33]. We consider arbitrary weak-entry correlation backgrounds and their full complex Frobenius neighborhoods.

Section 63.2 gives an inequality retaining $\operatorname{per} B$ for an arbitrary B when A is a weak-entry multiplier or a small perturbation. Section 63.3 retains $\operatorname{per} A$ in a smaller neighborhood of the equicorrelation cofactor optimizer and also controls the full Schur power. Section 63.4 determines the sharp exponential partner quotient in this neighborhood. Section 63.5 uses cofactor information from a principal block to control the self-conjugate comparison for the full matrix, giving a strong form of Chollet's inequality [34] on a broad class.

63.2 Strong uniform one-factor comparison

Let A, B, S be complex correlation matrices of order n. For every $K \geq 1$ there are $\delta_K > 0$ and N_K such that if, in some diagonal-unitary gauge,

$$\max_{i \neq j} |S_{ij}| \leq K/n, \qquad \|A - S\|_{\mathrm{F}} \leq \delta_K, \qquad n \geq N_K,$$

then

$$\operatorname{per}(A \circ B) \leq \operatorname{per} B$$

for every B, with equality precisely when $B = I$. One may take $\delta_K = 1/(512K^2)$ and $N_K = \lceil 10240K^4 \rceil$. In particular $\max_{i \neq j} |A_{ij}| \leq n^{-3/4}/11$ suffices for $n \geq 11^4$: take $S = A$ and $K = n^{1/4}/11 \geq 1$, so that $\|A - S\|_{\mathrm{F}} = 0 \leq \delta_K$ and $N_K = \lceil 10240\, n/11^4 \rceil \leq n$; the proof uses only these explicit values of δ_K and N_K. This proves the Bapat–Sunder inequality on a class of multipliers including the cofactor optimizer $S_n = (1-b_n)I + b_n J$, $b_n = 2/(n+1) + O(e^{-an})$. It has $\rho(|S_n - I|) \to 2$, outside the scaled diagonally dominant criterion. The unrestricted cofactor localization theorem of Sections 36–41 puts all sharp-scale near-maximizers, and every maximizer in large order, in this class.

63.2.1 Perturbing a weak-entry multiplier

For any correlation matrix B, write $p = \operatorname{per} B$, $P = S \circ B$, and $D = (A - S) \circ B$. In the expansion of $\operatorname{per}(P + D)$, the choices of D-entries are partial row-to-column injections. For each, the complementary minor of P has absolute permanent at most $\operatorname{per} P$. Indeed, the permanent of a cross submatrix is the inner product of two equal-degree Fischer products, so Cauchy–Schwarz bounds it by the geometric mean of the two principal permanents. Each is at most $\operatorname{per} P$ by Lieb's inequality. Dropping injectivity in the sum of the absolute values of the partial injections therefore gives, for $\delta = \|A - S\|_{\mathrm{F}}$,

$$\operatorname{per}(A \circ B) \leq (\operatorname{per} P) \prod_i (1 + \textstyle\sum_j |D_{ij}|) \leq (\operatorname{per} P) e^{n\delta}. \tag{W1}$$

Diagonal phase conjugation of A only conjugates the product $A \circ B$, preserving its permanent. We assume the displayed gauge throughout.

To bound $\operatorname{per} P$, entrywise absolute expansion gives an upper bound by $\operatorname{per}((1 - K/n)I + (K/n)|B|)$. For every principal k-set U, permutation-monomial Cauchy–Schwarz gives

$$\operatorname{per} |B[U]| \leq \big(k!\, \operatorname{per}(|B[U]|^2)\big)^{1/2}.$$

For any order-k correlation matrix G, let E be a diagonal PSD majorant in Loewner order. By the Schur product theorem $G \circ \bar{G} \preceq E \circ \bar{G} = E$. The permanent is Loewner monotone on PSD matrices, by its symmetric-tensor representation. Applying the diagonal-relaxation theorem of Anari, Gurvits, Oveis Gharan and Saberi [35, Theorem 1] therefore gives $\operatorname{per}(|G|^2) \leq c^k \operatorname{per} G$ with $c = e^{1+\gamma_{\mathrm{E}}} < 5$. Here γ_{E} is Euler's constant. Expanding the identity contribution in the entrywise upper bound and using $\operatorname{per} B[U] \leq p$ yields

$$\operatorname{per} P \leq \sqrt{p} \sum_{k=0}^{n} \frac{(K\sqrt{c})^k}{\sqrt{k!}} \leq \sqrt{2p}\, e^{cK^2}. \tag{W2}$$

For the last inequality use Cauchy–Schwarz with factors $2^{-k/2}$ and $(\sqrt{2c}K)^k/\sqrt{k!}$.

63.2.2 Completing the strong comparison

Take $\sigma = 1/(128K^2)$, and first suppose $p \le e^{\sigma n}$. Partition the edges of the complete graph into at most $2n$ matchings by greedy edge coloring. For every matching M, Lieb's inequality [1] gives $p \ge \prod_{ij\in M}(1 + |B_{ij}|^2)$. Since $\log(1 + x) \ge x/2$ on $[0, 1]$, summing over the matching partition implies

$$\|B - I\|_{\mathrm{F}}^2 \le 8n \log p, \qquad \max_{i\ne j} |B_{ij}|^2 \le 2 \log p.$$

Put $C = |\operatorname{off}(A \circ B)|$, with the absolute value entrywise. Then

$$\|C\|_{\mathrm{op}} \le (K/n)\|B - I\|_{\mathrm{F}} + \delta \le K\sqrt{8\sigma} + \delta < 1/2$$

for $\delta \le \sigma/4$. Expanding $\operatorname{per}(I + C)$ over vertex-disjoint directed cycles, bounding $\prod_c(1 + w_c) \le \exp \sum_c w_c$, and noting that the directed j-cycles have total weight at most $\operatorname{tr}(C^j)/j$ because $C \ge 0$ entrywise, gives $\log \operatorname{per}(I+C) \le \sum_{j\ge 2} \operatorname{tr}(C^j)/j$: one allows overlapping directed cycles and bounds their weights by the closed walk sum. The operator norm bound makes this at most $\|C\|_{\mathrm{F}}^2$ by the spectral theorem. The two matching estimates yield

$$\log \operatorname{per}(A \circ B) \le \|C\|_{\mathrm{F}}^2 \le (16K^2/n + 4\delta^2) \log p \le \tfrac{1}{2} \log p \tag{W3}$$

when $n \ge N_K$. If $B \ne I$, a pair principal block gives $p > 1$, so the comparison is strict.

If $p \ge e^{\sigma n}$, (W1) and (W2) give

$$\operatorname{per}(A \circ B)/p \le \sqrt{2}\, \exp(cK^2 - \sigma n/4) \le e^{-\sigma n/8},$$

since $\sigma n/8 = n/(1024K^2) \ge 10K^2$. This completes the proof, including equality at $B = I$.

Since $\operatorname{per} A \ge 1$ for every correlation matrix A by Lieb's inequality [1], the comparison gives $\operatorname{per}(A \circ B) \le \operatorname{per} B \le \operatorname{per} A \operatorname{per} B$. Thus Chollet's inequality holds for every correlation matrix B whenever A satisfies the hypotheses above.

63.3 The reverse fixed-factor inequality and the Schur power gap

Let $Q = (1 - b_n)I + b_n J$ be the positive equicorrelation optimizer. Its Schur power has a simple top eigenvalue by entrywise positivity. The next argument quantifies its spectral gap and controls complex perturbations. Suppose that in a diagonal phase gauge

$$\|A - Q\|_{\mathrm{F}} + \|(A - Q)\mathbf{1}\|_2 \le \omega/\sqrt{n}. \tag{W4}$$

For a sufficiently small absolute $\omega > 0$ and all sufficiently large n, the normalized Schur power $\Pi(A)/\operatorname{per} A$ has top eigenvalue 1 of multiplicity one and all other eigenvalues bounded above by $1 - c_0$ for an absolute $c_0 > 0$. Here $\Pi(A)_{\sigma,\tau} = \prod_i A_{\sigma(i),\tau(i)}$ is indexed by $\mathfrak{S}_n$. If the left side of (W4) is $o(n^{-1/2})$, the second eigenvalue converges to $1/2$. In particular

$$\operatorname{per}(A \circ B) \le \operatorname{per} A \left(1 - c_0 + c_0 \operatorname{per} B/n!\right)$$

for every correlation matrix B, with equality in $\operatorname{per}(A \circ B) \le \operatorname{per} A$ precisely when B has rank one. The cofactor localization theorem of Sections 36–41 makes (W4) hold on a fixed small deficit band $\gamma_n - \Lambda(A) \le \eta/n$, where $\gamma_n = \max_A \Lambda(A)$. It makes the left side $o(n^{-1/2})$ on sequences with $\gamma_n - \Lambda(A) = o(n^{-1})$.

63.3.1 Comparing permutation kernels

Write $a_\pi(A) = \prod_i A_{i,\pi(i)}$ and $\nu_Q(\pi) = a_\pi(Q)/\operatorname{per} Q$. Since $b_n > 0$, this is a probability on all permutations. Put $E = A - Q$, $R = Q + 2\operatorname{Re} E$, and let G have diagonal one and off-diagonal entries $G_{ij} = |A_{ij}|^2/b_n$. Then R is a correlation matrix by (W4) and $\lambda_{\min}(Q) \to 1$. Also $G = R+W$, where W has diagonal zero and $\sum_{ij} |W_{ij}| = \|E\|_{\mathrm{F}}^2/b_n = O(\omega^2)$. The segment from R to G is positive definite for ω sufficiently small. The cofactor derivative and $|H_{ij}| \le 1$ show that the absolute change of log permanent along this segment is $O(\omega^2)$.

The uniform local holomorphic chart also gives

$$\left|\log \frac{\operatorname{per} A}{\operatorname{per} Q}\right| + \left|\log \frac{\operatorname{per} R}{\operatorname{per} Q}\right| = O(\omega). \tag{W5}$$

For this consequence use the phase fit and real longitudinal fit in the local proof on each of A, R individually. They bound the longitudinal perturbation parameter t and the centered zero-diagonal transverse scale $\alpha = s^2$ by $O(\omega/\sqrt{n})$. The chart divides the permanent by $(\operatorname{per} Q)\exp(a(t))$, with the other anchor identically zero because the transverse direction has zero diagonal. At the radial anchor $\tau = 1/2$, the log ratio $a(t)$ of the two products of diagonal factors is bounded by $C|t|\sqrt{n}$, from $\sum_i |h_i| \le \sqrt{n}$ for its unit longitudinal direction. The normalized permanent is holomorphic, uniformly bounded and uniformly zero-free on a fixed smaller polydisc, with value one at the origin and even dependence on s. Cauchy's estimate gives deviation $O(|t| + |s|^2)$ there. For the real covariance evaluations it follows that its log is $O(|t| + \alpha)$, yielding (W5). These facts use the permanent normalization in the uniform local proof Sections 21–22, rather than an inference about the permanent from the objective instability alone.

Set $r_\pi = a_\pi(A)/a_\pi(Q)$. The entrywise definition of G gives the exact identity

$$\mathbb{E}_{\nu_Q} |r_\pi - 1|^2 = \frac{\operatorname{per} G}{\operatorname{per} Q} - 2\frac{\operatorname{per} A}{\operatorname{per} Q} + 1 = O(\omega).$$

Thus the ℓ^1 distance of the normalized permutation coefficients of A and Q is $O(\sqrt{\omega})$. Indeed Cauchy–Schwarz gives the bound with $a_\pi(A)/\operatorname{per} Q$, and (W5) gives the bound on normalizing by $\operatorname{per} A$. Every Fourier operator of the normalized coefficient difference has norm at most this ℓ^1 distance, since it is a linear combination of unitary permutation actions. In particular it holds for the regular representation, which yields the normalized Schur powers.

63.3.2 The scalar kernel and its gap

Writing $t_n = (1 - b_n)/b_n$, the law ν_Q can be sampled by first choosing a subset U with weight $|U|! b_n^{|U|}(1 - b_n)^{n-|U|}/\operatorname{per} Q$, then permuting U uniformly and fixing its complement. To verify the law at π, sum over U containing the moved set of π. The total is exactly $b_n^{|\operatorname{supp}\pi|}/\operatorname{per} Q$. The subset size k is $n - N$, where N is Poisson of mean t_n, conditioned on $N \le n$. In particular $k/n \to 1/2$ in probability.

For fixed $1 \le k \le n$, the uniform average of the subgroup projections onto S_U-invariant vectors is central in every unitary S_n representation. By Young branching [36], on the irreducible of shape λ its eigenvalue is the fraction of standard tableaux of shape λ in which $1, \dots, k$ all lie in the first row. This assertion follows by iterating the restriction rule down to S_k, where the trivial irreducible has shape (k).

For a nontrivial shape let j be the first label outside the first row. Necessarily it is in cell $(2, 1)$. The number of tableaux with first outside label j is nonincreasing in $j = 2, \dots, n$. For $j \ge 3$ swap j and $j - 1$ in cells $(2, 1)$ and $(1, j - 1)$. The result is a standard tableau with first outside label

$j-1$, and the map is injective. It follows that the fraction with $j>k$ is at most $(n-k)/(n-1)$, with equality in the shape $(n-1,1)$. For $k=0$ the subgroup projection is the identity. Therefore the largest nontrivial eigenvalue of the scalar kernel is

$$\mathbb{E}\min\{1,(n-k)/(n-1)\}\longrightarrow 1/2$$

(compare [37, Theorem 1.8], which gives this gap for mean-field subset shuffles, attained on the standard representation). Combining this with the ℓ^1 perturbation proves the claimed uniform gap for ω small. It proves convergence of the second eigenvalue if $\omega\to 0$, by Weyl's inequality. The constant vector remains a top eigenvector because the exact row sum is $\operatorname{per} A$.

There is a corresponding limit for each fixed number of boxes outside the first row. Fix a partition μ of r and write $\lambda=(n-r,\mu)$ for large n. For $k\ge r$, the number of tableaux in the S_k-invariant count is $f^\mu\binom{n-k}{r}$, where f^μ counts standard tableaux of shape μ. Choose the r lower-row labels above k, fill the lower shape in any standard order, and put all other labels increasingly in the first row. The first r positions in that row contain labels at most k, so the conditions across the rows are automatic. The full dimension f^λ is at most $f^\mu\binom{n}{r}$ and at least $f^\mu\binom{n-r}{r}$ by imposing that the lower-row labels avoid $1,\dots,r$. Thus the normalized scalar eigenvalue on this block tends to 2^{-r} after averaging over $k/n\to 1/2$. The event $k<r$ has probability tending to zero. The same limit holds for all eigenvalues in this block for matrices A with $\omega\to 0$, by the uniform Fourier operator comparison.

63.3.3 Testing a product Gram vector

Take unit Gram vectors $v_1,\dots,v_n$ for B and the product tensor $w=v_1\otimes\cdots\otimes v_n$. With the permutation action convention in which its matrix coefficient is $a_\pi(B)$, the quadratic form of $\sum_\pi a_\pi(A)\rho(\pi)$ on w is $\operatorname{per}(A\circ B)$. Equivalently use conjugate Gram vectors if the action gives the inverse coefficient. The full symmetrization of w has squared norm $\operatorname{per} B/n!$. The Fourier bound holds in every representation, so splitting off the invariant component proves the displayed inequality. If B has rank larger than one, some pair of its Gram vectors are nonproportional and the corresponding transposition does not fix w. Thus its symmetrized squared norm is less than one. If B has rank one, $A\circ B$ is a diagonal phase conjugate of A.

The normalized spectral gap is a permanent-on-top result for this band. In conjunction with the first section it gives both $\operatorname{per}(A\circ B)\le\operatorname{per} B$ and $\operatorname{per}(A\circ B)\le\operatorname{per} A$ when the first factor is near a sharp cofactor maximizer. These two inequalities have different equality cases. The natural permutation block is the entry-weighted cofactor $A\circ C(A)$, up to transpose, by permanent expansion. Thus its nonconstant spectrum is also bounded above by $1-c_0$ after normalizing by $\operatorname{per} A$. Along asymptotically sharp sequences its nonconstant eigenvalues converge uniformly to $1/2$, since the conjugation-invariant comparison acts by the same scalar on the entire standard representation. By [2, Theorem 1], a positive semidefinite matrix A satisfying the 1985 Bapat–Sunder inequality $\operatorname{per}(A\circ B)\le\operatorname{per} A\prod_i b_{ii}$ for every positive semidefinite B also satisfies the 1986 Bapat–Sunder inequality $\lambda_{\max}(F_A)\le\operatorname{per} A$, where $F_A=[a_{ij}\operatorname{per} A(i,j)]$; both conjectures fail in general. The displayed comparison gives the first inequality for every A in band (W4), after diagonal scaling of B to a correlation matrix, so their theorem gives the second for these A, in agreement with the spectral bound just stated.

Permanental dominance on the band. For a subgroup $H\le\mathfrak{S}_n$ and a character χ of H put $d_\chi^H(A)=\sum_{\sigma\in H}\chi(\sigma)\prod_i A_{i,\sigma(i)}$. Lieb's permanental dominance conjecture [1] asserts $d_\chi^H(A)\le\chi(e)\operatorname{per} A$ for every positive semidefinite A; see [38, Section 4, (10)] for its statement and status. Soules observed that permanent-on-top implies it [39], and Merris recorded the implication for

all characters [40, p. 216]. Consequently, for all sufficiently large n, every correlation matrix A satisfying (W4) in some diagonal phase gauge, and every DAD with D positive diagonal, satisfies $d_\chi^H(A) \le \chi(e) \operatorname{per} A$ for every subgroup $H \le \mathfrak{S}_n$ and every character χ of H. Indeed, for irreducible χ put $x_\sigma = \chi(\sigma)$ for $\sigma \in H$ and $x_\sigma = 0$ otherwise. Since $\Pi(A)_{\sigma,\tau} = a_{\tau\sigma^{-1}}(A)$, the orthogonality relation $\sum_{h\in H} \chi(gh)\chi(h^{-1}) = |H|\chi(g)/\chi(e)$ gives $x^*\Pi(A)x = |H|\, d_\chi^H(A)/\chi(e)$, while $x^*x = |H|$, and the Rayleigh quotient of the Hermitian matrix $\Pi(A)$ is at most its top eigenvalue $\operatorname{per} A$. A reducible character is a nonnegative integral combination of irreducible ones; both sides are invariant under diagonal unitary conjugation and scale by $\prod_i D_{ii}^2$ under $A \mapsto DAD$. For entrywise nonnegative A the inequality is immediate, since $|\chi(\sigma)| \le \chi(e)$. Band (W4) also contains matrices with a negative triangle product $A_{12}A_{23}A_{31}$, which no diagonal gauge makes nonnegative: one symmetric pair of entries may move by $\omega/(2\sqrt{2n})$, which exceeds b_n once $n > 32/\omega^2$.

63.4 Exact uniform exponential rate for Chollet's quotient

If $\|A - S_n\|_{\mathrm{F}} = o(1)$, then

$$\left(\max_B \frac{\operatorname{per}(A \circ B)}{\operatorname{per} A \operatorname{per} B}\right)^{1/n} \longrightarrow \frac{\sqrt{e}}{2}.$$

In fact for sufficiently large n the maximum over partners is exactly $1/\operatorname{per} A$, attained only by $B = I$ by the strong comparison above. Since the least eigenvalue of S_n tends to one, Loewner pinching gives $(1 - o(1))S_n \preceq A \preceq (1 + o(1))S_n$. The permanent is monotone for PSD matrices in Loewner order, by its symmetric-tensor representation and telescoping tensor powers. By homogeneity it follows that $(\operatorname{per} A)^{1/n}$ and $(\operatorname{per} S_n)^{1/n}$ have the same limit, $2/\sqrt{e}$. The scalar limit follows from $n!b_n^n \sum_{k=0}^n t_n^k/k!$, where $t_n = (1 - b_n)/b_n \sim n/2$. The truncated sum is e^{t_n} times a Poisson lower tail probability tending to one.

63.5 Strong self-permanent inequality from a broad-band principal block

Suppose T is a correlation matrix of order n satisfying $n(c_n - \Lambda(T)) \le M$, with M fixed. There are constants $a_M > 0$ and K_M such that if T is a principal block of a correlation matrix

$$A = \begin{pmatrix} T & X \\ X^* & G \end{pmatrix}$$

with $k \le a_M n$ additional indices, then

$$\frac{\operatorname{per}(A \circ \bar{A})}{\operatorname{per} A} \le K_M n^{-1/2} e^{-a_M n}$$

for all sufficiently large n. The adjoining rows and the block G are arbitrary subject to positive semidefiniteness. This yields the strong self-conjugate permanent inequality on the whole class. Permutation-monomial Cauchy–Schwarz also gives an exponential reserve in Chollet's inequality when both factors meet this condition.

In the cofactor gauge, write $T\mathbf{1} = 3\mathbf{1} + r$. The two-deletion comparison and the weighted Laplace identity of Sections 36–41 and Lemma 70 give $\|T - I\|_{\mathrm{F}} = O_M(1)$, $\|r\|_2 = O_M(1)$, and $\sum_i r_i = O_M(1)$. On the subset S with $|r_i| \le 1$ all but $O_M(1)$ indices remain. With $u = \mathbf{1}/\sqrt{n}$, PSD Cauchy–Schwarz gives $T \succeq (Tu)(Tu)^*/(u^*Tu)$ and $u^*Tu = 3 + O_M(n^{-1})$. Restrict to S. There $\sum_{i\in S} r_i = O_M(1)$ and $\sum_{i\in S} |r_i|^2 = O_M(1)$, so Taylor's bound on $\log|1 + r_i/3|$ gives $\prod_{i\in S} |3 + r_i|^2 = 9^{|S|} e^{O_M(1)}$. Principal monotonicity and Loewner monotonicity give

$$\operatorname{per} T \ge |S|! \prod_{i\in S} \frac{|3 + r_i|^2}{n(u^*Tu)} \ge c_M n!(3/n)^n \ge c'_M \sqrt{n}(3/e)^n. \tag{W6}$$

The comparison of factorials uses $n - |S| = O_M(1)$.

Let $q(H) = \operatorname{per}(|H|^2)$. Positivity gives $\|X_{\cdot j}\|_2^2 \le \|T\|_{\mathrm{op}} \le K'_M$. In the permanent of $|A|^2$, first sum over injections of the old rows into all columns. For a fixed injection the remaining rows are the k new rows. Select a subset J of these rows that use old columns, bounding their total absolute weights by $\prod_{j\in J} \|X_{\cdot j}\|_2^2$ by dropping injectivity. The remaining new rows and remaining new columns form a cross minor of $|G|^2$, whose permanent is at most $q(G)$ by Fischer Cauchy–Schwarz and the principal bound. Summing over J gives the factor $(1 + K'_M)^k q(G)$. The sum over old-row injections is at most $\exp(\|T - I\|_{\mathrm{F}}^2 + \|X\|_{\mathrm{F}}^2)$ by the row-product bound. Thus

$$q(A) \le e^{O_M(1)+O_M(k)} q(G).$$

The diagonal-relaxation consequence used in (W2) gives $q(G) \le c^k \operatorname{per} G$. Lieb gives $\operatorname{per} A \ge \operatorname{per} T \operatorname{per} G$. Combine with (W6) and choose a_M sufficiently small. The empty additional block gives the corresponding statement for T itself.

64 Sharp near-null bounds for permanental cofactors

A principal minor's permanental cofactor matrix can have very small eigenvalues, so ordinary single-row bordering estimates do not directly recover the appended correlations. We prove a sharp fixed-order comparison between the minor and a fractional power of its normalized cofactor matrix. For a minor of order m, every feasible appended column has projection of order at most $\varepsilon^{1/[2(m-1)]}$ onto cofactor eigenvalues at most ε. A rank-two family built from roots of unity shows that the exponent cannot be increased. The upper bound uses projective Gram coordinates and polynomial interpolation after grouping close complex nodes.

Let B be a complex correlation matrix, meaning a Hermitian positive semidefinite matrix with diagonal one. Write $C(B)_{ij} = \operatorname{per} B(i \mid j)$ for its permanental cofactor matrix and put $H = C(B)^{\mathsf{T}} / \operatorname{per} B$. The deleted-product representation gives $H \succeq 0$. In the exact bordering identities, a feasible appended column z enters through Hz and $z^* Hz$ (Lemma 70). The question is to control z in directions of very small cofactor eigenvalue, and we answer it for each fixed order.

Theorem 30. *For a correlation matrix B of order $m \ge 2$ and unit vector c,*

$$\|B^{1/2} c\| \le K_m (c^* H c)^{1/[2(m-1)]}. \tag{64}$$

Here $K_m = 480 m^6$. It follows that $B \preceq m K_m^2 H^{1/(m-1)}$. If $\left(\begin{smallmatrix} B & z \\ z^* & 1 \end{smallmatrix}\right) \succeq 0$ and E_ε is the spectral projector of H onto eigenvalues at most $\varepsilon \ge 0$, then

$$\|E_\varepsilon z\| \le K_m \varepsilon^{1/[2(m-1)]}. \tag{65}$$

The exponents in all three bounds are optimal for fixed m.

Theorem 30 is uniform over the correlations and allows singular minors. The constant is polynomial in m and the exponent depends on the order. The published large-order alignment theorem uses a different argument, based on two columns appended to a common deletion in Sections 36–41. The present result quantifies the near-null part of the single-row method without changing that large-order conclusion.

64.1 Deleted products and projective coordinates

Represent $B_{ij} = \mathbb{E}[L_i(G)\overline{L_j(G)}]$ by unit complex linear forms L_i and a standard complex Gaussian vector G of dimension $d = \operatorname{rank} B$. Let $P = \prod_i L_i$ and $F = \sum_i c_i \prod_{j \neq i} L_j$. In the Gaussian norm,

$$\|F\|^2 = \operatorname{per} B \, c^* H c, \qquad \left\| \sum_i c_i \overline{L_i} \right\|_{\text{coeff}} = \|B^{1/2} c\|.$$

Here the second norm is the ordinary Euclidean norm of the coefficients of an antilinear form. These identities follow by Gaussian expansion and explain the transpose convention in H. These are the identities of Lemma 70.

For $d = 1$, one has $H = B/m$ and (64) follows. Otherwise choose a unit vector w_0 with $|L_i(w_0)| \geq 1/m$ for every i. For a unit linear form on the complex unit sphere in dimension d, the squared modulus has distribution $\mathrm{Beta}(1, d-1)$. In particular the probability of modulus below α is at most $(d-1)\alpha^2$. A union bound with $\alpha = 1/m$ gives the choice. Use w_0 as first coordinate and gauge the phases of the forms so the first coefficients α_i are positive. The same diagonal unitary acts on B, H, and the test vector, so both sides of (64) are unchanged. Write

$$L_i(w_1, w') = \alpha_i (w_1 + v_i \cdot w'), \qquad \|v_i\| \leq m, \qquad d_i = \frac{c_i}{\alpha_i}.$$

Choose another unit vector g in the remaining coordinates so that the complex numbers $t_i = v_i \cdot g/(2m)$ obey

$$|t_i| \leq \tfrac{1}{2}, \qquad |t_i - t_j| \geq \frac{\|v_i - v_j\|}{2m^{5/2}}.$$

For $d - 1 = 1$ this holds for any unit g. In higher dimension apply the same small-ball bound with threshold $m^{-3/2}$ to each nonzero normalized difference and take a union bound over the at most $m(m-1)/2$ pairs. Equal projective forms may share a node. Distinct projective forms have distinct nodes.

The vector whose norm is sought is now $\sum_i d_i f_i$ with

$$f_i = \frac{(1, \overline{v_i})}{1 + \|v_i\|^2}.$$

These vectors have norm at most one and satisfy $\|f_i - f_j\| \leq L|t_i - t_j|$ with $L = 2m^{5/2}$. Indeed, Euclidean inversion obeys $\|u/\|u\|^2 - v/\|v\|^2\| = \|u - v\|/(\|u\|\|v\|)$, as follows by squaring. Apply it to $(1, \overline{v}_i)$ and $(1, \overline{v}_j)$, of norms at least one.

64.2 Holomorphic moments from the cofactor norm

On the complex circle $|z| = 1$, evaluate the quotient at $w = w_0 + zg/(2m)$. Then

$$\frac{F(w)}{P(w)} = \sum_i \frac{d_i}{1 + z t_i} = \sum_{k \geq 0} (-1)^k M_k z^k, \qquad M_k = \sum_i d_i t_i^k.$$

Each $|L_i(w)| \geq 1/(2m)$. For a homogeneous polynomial of degree $m - 1$, Cauchy–Schwarz in the Gaussian monomial basis gives

$$|F(w)| \leq \|F\| \frac{\|w\|^{m-1}}{\sqrt{(m-1)!}} \leq 2\sqrt{m}\, \sqrt{c^* H c},$$

using $\operatorname{per} B \leq m!$ and $(1 + 1/(2m))^{m-1} < 2$. Cauchy's coefficient formula therefore gives

$$|M_k| \leq 2\sqrt{m} (2m)^m \sqrt{c^* H c}.$$

64.3 Interpolation of the target coefficients

The flattening and Newton interpolation step is related to the real-node argument for Gaussian mixtures by Wu and Yang. Here the nodes and weights can be complex, and the bounded holomorphic moments come from deleted Gram products. See [41, arXiv:1807.07237, Proposition 10, third proof].

Fix $0 < s \le 1$. Join distinct nodes at distance below s and take connected components. A component has diameter at most $(m-1)s$, and nodes in different components have distance at least s. Replace each f_i by the value at a representative of its component, at error at most $L(m-1)s$. The replaced data stay Lipschitz on the nodes with constant $(2m-1)L$. Within a component the values are constant. Across components the representative errors sum to at most $2L(m-1)s$, which is at most $2L(m-1)$ times the node separation.

Order the distinct nodes by consecutive components and interpolate the replaced vector data by a Newton polynomial Q of degree at most $m-1$. The first divided differences have norm at most $(2m-1)L$. Every positive-order divided difference with all its nodes in one component is zero. For an interval of nodes meeting more than one component, its first and last nodes have distance at least s. The divided-difference recursion therefore gives an upper bound $2^{k-1}(2m-1)L\,s^{-(k-1)}$ at order $k \ge 1$. The coefficient norm sum of the Newton basis product at order k is at most $(3/2)^k$. Hence the coefficient norm sum of Q is bounded by

$$5m^{9/2}2^{m-2}(3/2)^{m-1}s^{-(m-2)}.$$

Repeated nodes have identical target values and need only be interpolated once.

Using $\sum_i |d_i| \le m^{3/2}$ and the moment bounds gives

$$\left\| \sum_i d_i f_i \right\| \le 2m^5\big(s + B_m s^{-(m-2)}\sqrt{c^*Hc}\big), \qquad B_m = 5(2m)^m 2^{m-2}(3/2)^{m-1}.$$

Choose $s = (B_m\sqrt{c^*Hc})^{1/(m-1)}$ if this is at most one, and use $\|B^{1/2}\| \le \sqrt{m}$ otherwise. At zero cofactor energy let s decrease to zero. In both cases the norm is at most $4m^5 B_m^{1/(m-1)}(c^*Hc)^{1/[2(m-1)]}$. Since $(2m)^{1/(m-1)} \le 4$, one has $B_m^{1/(m-1)} \le 120m$.

64.3.1 Appended columns and the matrix inequality

If the bordering matrix with blocks $B, z, z^*, 1$ is positive semidefinite, $z = B^{1/2}v$ for some $\|v\| \le 1$. This follows from the Schur complement on the range of B, and remains valid when B is singular. For a unit c in the range of E_ε, $|c^*z| \le \|B^{1/2}c\|$, and its cofactor energy is at most ε. Taking the supremum proves (65).

For the matrix inequality, apply (64) to an orthonormal eigenbasis e_i of H, with eigenvalues λ_i. For any coefficients s_i,

$$\left\| B^{1/2}\sum_i s_i e_i \right\| \le K_m \sum_i |s_i|\lambda_i^{1/[2(m-1)]} \le \sqrt{m}K_m \left(\sum_i |s_i|^2 \lambda_i^{1/(m-1)} \right)^{1/2}.$$

Squaring gives the stated Loewner comparison.

64.4 Separated groups of Gram directions

The same argument gives an improvement when the normalized projective forms have separated groups. Suppose the nodes t_i are partitioned into groups of at most r distinct nodes, with $1 \le r \le m$

and separation at least $\Delta \in (0,1]$ between groups. For $r \geq 2$, use $s \leq \Delta$ and order first by the separated groups and then by the fine components. In a Newton divided difference of order at least r, the end nodes belong to different groups, so the denominator is at least Δ. The same recursion bounds the polynomial coefficient sum by a constant times $s^{-(r-2)}\Delta^{-(m-r)}$. Writing $e = c^*Hc$, the two terms in the norm estimate are now bounded by a constant depending only on m times

$$s + \sqrt{e}\, s^{-(r-2)}\Delta^{-(m-r)}.$$

If $s = (\sqrt{e}\,\Delta^{-(m-r)})^{1/(r-1)} \leq \Delta$, the norm is

$$O_m\Big(e^{1/[2(r-1)]}\Delta^{-(m-r)/(r-1)}\Big).$$

If instead $s > \Delta$, equivalently $\sqrt{e} > \Delta^{m-1}$, then $\Delta^{-(m-r)/(r-1)} > e^{-(m-r)/[2(m-1)(r-1)]}$, so the displayed expression exceeds $e^{1/[2(m-1)]}$, and Theorem 30 already gives a bound of the same or better order. At $e = 0$ take the limit. If $r = 1$, interpolate the original data directly to get $O_m(\sqrt{e}\,\Delta^{-(m-2)})$. Separation of projective coordinate vectors gives node separation up to the factor $1/(2m^{5/2})$.

64.5 The cofactor kernel

At zero cofactor energy the polynomial F vanishes. Distinct projective forms give distinct nodes, so the first m moments of the weights d_i vanish. The Vandermonde matrix on distinct nodes shows that the weights sum to zero within each proportionality class. Conversely, those cancellations make F zero. Thus the rank of the cofactor is exactly the number of distinct projective Gram forms.

64.6 Sharpness and scope

Let ζ_i run over the m-th roots of unity and take

$$L_i(w_1, w_2) = \frac{w_1 + a\zeta_i w_2}{\sqrt{1+a^2}}, \qquad c_i = \frac{\zeta_i}{\sqrt{m}}, \qquad 0 < a < 1.$$

Then $Bc = [ma^2/(1+a^2)]c$. The normalized cofactor has c as an eigenvector as well, with eigenvalue

$$\varepsilon_a = \frac{a^{2m-2}(1+a^2)}{1+a^{2m}}.$$

To verify this, the full product is proportional to $w_1^m - (-aw_2)^m$, with Gaussian squared norm $m!(1+a^{2m})/(1+a^2)^m$. The i-th deleted product before normalization is $\sum_{k=0}^{m-1}(-a\zeta_i)^k w_1^{m-1-k}w_2^k$. Root-of-unity orthogonality diagonalizes the deleted-product Gram. For the vector c, only $k = m-1$ contributes, giving the stated eigenvalue. Thus $\|B^{1/2}c\|$ is of order a, while c^*Hc is of order a^{2m-2}. The feasible bordering column $z = B^{1/2}c$ lies in the corresponding near-null eigenspace, so any larger exponent fails for (65) too. Testing any proposed stronger Loewner exponent on the same c gives the matching obstruction.

The exponent bound concerns all fixed-order Gram minors, connecting a nearly null cofactor direction to the original Gram form. The example also rules out a uniform linear comparison of H against B in orders $m \geq 3$. The large-order alignment rests instead on the two-row deletion argument of Sections 36–41.

65 Cutoff and Rényi profiles for permanental subset shuffles

We study permutation walks that choose a uniformly random subset of positions at each step and uniformly shuffle that subset; the Schur power of positive equicorrelation is such a walk. When the untouched fraction concentrates near $\rho \in (0,1)$, total variation, relative entropy, Hellinger distance and every Rényi order below one have Poisson cutoff profiles at $\log_{1/\rho} n$. For the permanent law the density in every L^α, $1 < \alpha < \infty$, is a Poisson fixed-point tilt plus the identity spike of the all-identity paths, which determines the exact Rényi profile at every finite order and the critical order at which the cutoff location moves; a uniform two-term formula holds at order infinity. Observing a fraction ζ of positions scales both terms by ζ on the saturated branch and gives the sharp projected cutoff at every order above one.

65.1 The shuffle and its limits

Start a walk at the identity permutation in S_n. At each step draw J_n, choose a uniform subset of size $n - J_n$, and apply a uniform permutation supported on that subset. The steps are independent. Denote the one-step law by μ_n.

Theorem 31. *Suppose $\rho_n \to \rho \in (0,1)$ and*

$$\Pr(|J_n/n - \rho_n| > \delta_n) = o(1/\log n), \qquad \delta_n \log n \to 0.$$

For integer times t_n with $n\rho_n^{t_n} \to \theta \in (0,\infty)$,

$$d_{\mathrm{TV}}(\mu_n^{*t_n}, \mathrm{Unif}(S_n)) \longrightarrow d_{\mathrm{TV}}(\mathrm{Pois}(1+\theta), \mathrm{Pois}(1)).$$

There is total variation cutoff with a bounded window at $\log n/\log(1/\rho_n)$.

The permanent law arises from equicorrelation $A_b = (1-b)I + bJ$ with $0 < b < 1$. Its Schur-power coefficients are $b^{|\operatorname{supp}\pi|}$, where the support of a permutation is its set of moved positions. Expand

$$b^{|\operatorname{supp}\pi|} = \sum_{S \supseteq \operatorname{supp}\pi} b^{|S|}(1-b)^{n-|S|}.$$

The normalized coefficients give the subset shuffle with probability $|S|!b^{|S|}(1-b)^{n-|S|}/P_n$ of choosing S, where $P_n = \operatorname{per} A_b$. The untouched size $J_n = n - |S|$ has the Poisson law of mean $\eta_n = (1-b_n)/b_n$ conditioned to be at most n.

Theorem 32. *For the permanent law assume $\rho_n = \eta_n/n \to \rho \in (0,1)$ and put $I_\rho = -\log\rho - 1 + \rho$. At $t = \lfloor s\log n \rfloor$, the unconditioned chi-square divergence tends to zero above*

$$s_\rho = \max\left\{\frac{1}{\log(1/\rho)}, \frac{1}{2I_\rho}\right\}$$

and tends to infinity below it. More precisely, for $t = O(\log n)$ such that $\Theta_n = n\rho_n^t$ and $Q_n = n!P_n^{-2t}$ both stay bounded, the divergence is $e^{\Theta_n^2} - 1 + Q_n + o(1)$.

At the scalar optimum of Sections 1–5 the published cofactor optimization gives $\rho_n = \eta_n/n = 1/2 + O(1/n)$, so $\rho_n^t = 2^{-t}(1+o(1))$ for $t = O(\log n)$. There Theorem 31 centers at $\log_2 n$, whereas Theorem 32 centers to leading order at $\log n/[2(\log 2 - 1/2)]$.

65.2 Rényi, relative-entropy and uniform profiles

Write U_n for uniform measure on S_n, $f_t = \mu_n^{*t}/U_n$ for the time-t density, F for the fixed-point count and e for the identity. For $0 < \alpha < \infty$, $\alpha \neq 1$, the Rényi divergence is

$$D_\alpha(\mu_n^{*t} \| U_n) = \frac{\log \int f_t^\alpha \, dU_n}{\alpha - 1};$$

at $\alpha = 1$ it is relative entropy and at $\alpha = \infty$ it is $\log \|f_t\|_\infty$. Put $\ell = \log(1/\rho)$ and

$$I_\rho(y) = \rho - y + y \log(y/\rho), \qquad I_\rho(1) = I_\rho,$$

so that I_ρ is the rate of Theorem 32.

Theorem 33 (Finite orders above one)**.** *For the permanent law let $\rho_n = \eta_n/n \to \rho \in (0,1)$, fix $1 < \alpha < \infty$ and put $\beta = 1 - 1/\alpha$. For integer $t = O(\log n)$ define*

$$\Theta = n\rho_n^t, \qquad Q_\alpha = (n!)^\beta P_n^{-t}.$$

If Θ and Q_α are bounded, then

$$f_t = e^{-\Theta}(1+\Theta)^F + n! P_n^{-t} \mathbf{1}_{\{e\}} + o_{L^\alpha(U_n)}(1), \tag{66}$$

$$\int f_t^\alpha \, dU_n = \exp\big((1+\Theta)^\alpha - 1 - \alpha\Theta\big) + Q_\alpha^\alpha + o(1). \tag{67}$$

At $t = \lfloor s \log n \rfloor$, D_α tends to zero for $s > s_\alpha$ and to infinity for $0 \le s < s_\alpha$, where

$$s_\alpha = \max\left\{ \frac{1}{\ell}, \frac{\beta}{I_\rho} \right\}.$$

The one-step kernel μ_n is the normalized Schur power $\Pi(A_{b_n})/\operatorname{per} A_{b_n}$ of Section 63.3, acting by convolution, so its eigenvalues are the multipliers a_λ. At the cofactor optimizer the uniform spectral gap of Section 63.3 is a one-step statement, and its second eigenvalue, tending to $1/2$, is the standard multiplier $a_{(n-1,1)} = \rho_n + O(1/n)$ that drives the lower bounds at $\log n/\ell$ and, at order infinity, at $2\log n/\ell$. The two terms of (66) are a Poisson fixed-point tilt and a spike carried by the paths whose every step is the identity. At $\alpha = 2$, $Q_2^2 = n! P_n^{-2t}$ and (67) is the chi-square profile of Theorem 32. The critical order is

$$\alpha_\rho = \frac{\ell}{\ell - I_\rho}, \qquad \text{equivalently } \beta = I_\rho/\ell.$$

For $1 < \alpha < \alpha_\rho$ the Rényi divergence has the Poisson profile $((1+\theta)^\alpha - 1 - \alpha\theta)/(\alpha - 1)$ at the total variation window $\Theta \to \theta$, and for $\alpha > \alpha_\rho$ it diverges there. At the scalar cofactor optimizer, $\rho = 1/2$ and $\ell - I_\rho = 1/2$, so $\alpha_{1/2} = 2\log 2 \approx 1.386$. Every order $1 < \alpha < 2\log 2$ cuts off at $\log_2 n$ with the Poisson profile, while every order $\alpha > 2\log 2$ cuts off later, at $\beta \log n/(\log 2 - 1/2)$.

Theorem 34 (Relative entropy and orders below one)**.** *Under the size concentration of Theorem 31, let $t = O(\log n)$ and $n\rho_n^t \to \theta \in [0, \infty)$. Then*

$$D_1(\mu_n^{*t} \| U_n) \longrightarrow (1+\theta)\log(1+\theta) - \theta,$$

$$D_\alpha(\mu_n^{*t} \| U_n) \longrightarrow \frac{(1+\theta)^\alpha - 1 - \alpha\theta}{\alpha - 1} \qquad (0 < \alpha < 1),$$

and the squared Hellinger distance $1 - \int \sqrt{f_t} \, dU_n$ tends to $1 - \exp(-(\sqrt{1+\theta} - 1)^2/2)$.

Thus relative entropy, Hellinger distance and every order below one share the total variation cutoff location $\log n/\log(1/\rho)$, each with the Poisson profile of the corresponding divergence between $\mathrm{Pois}(1+\theta)$ and $\mathrm{Pois}(1)$.

Theorem 35 (Order infinity)**.** *For the permanent law with $\rho_n = \eta_n/n \to \rho \in (0,1)$, let $t = O(\log n)$ and suppose $Z_n = n^2\rho_n^t$ and $Q_\infty = n!P_n^{-t}$ are bounded. Then, uniformly on S_n,*

$$f_t = \exp(n\rho_n^t F) + Q_\infty \mathbf{1}_{\{e\}} + o(1).$$

Hence $\|f_t - 1\|_\infty = e^{Z_n} - 1 + Q_\infty + o(1)$, $D_\infty(\mu_n^{*t}\|U_n) = \log(e^{Z_n} + Q_\infty) + o(1)$, and the separation distance $1 - \min_\sigma f_t(\sigma)$ tends to zero. At $t = \lfloor s\log n\rfloor$ the order-infinity divergence tends to zero for $s > \max\{2/\ell, 1/I_\rho\}$ and to infinity for $0 \le s < \max\{2/\ell, 1/I_\rho\}$.

Proposition 36 (The transition at the critical order)**.** *Let $\beta = I_\rho/\ell$, $(\rho_n - \rho)\log n \to 0$ and $\Theta \to \theta \in (0,\infty)$. Then $\log Q_\alpha = \beta n(\log\theta - 1) + o(n)$. Consequently the Poisson Rényi profile of Theorem 33 holds for $\theta < e$, and $D_\alpha \to \infty$ for $\theta > e$. If $\rho_n = \rho + O(1/n)$, as at the scalar cofactor optimizer, then*

$$\log Q_\alpha = \beta n(\log\Theta - 1) - \frac{(\log n)^2}{2\ell} + O(\log n),$$

so, along any sequence of n for which $\Theta = e + o((\log n)^2/n)$, the identity term vanishes; and along any sequence for which Q_α is bounded and bounded away from zero, necessarily $\log\Theta - 1 = (\log n)^2/(2\beta\ell n) + O(\log n/n)$. Since Θ ranges over a lattice of ratio ρ_n, these windows are met along subsequences of n.

65.3 Partial observation

Observe only the ordered images of the first k_n positions. Let V_n be the uniform law on ordered injective k_n-tuples, F_k the number of observed fixed points, x_0 the initial tuple and f_t^{proj} the density of the observed tuple with respect to V_n. Assume $k_n/n \to \zeta \in (0,1]$ and put $m = n - k_n$. Write $P_m(b)$ for the permanent of the order-m equicorrelation matrix with off-diagonal entry b, with $P_0 = 1$, and

$$R_\rho(x) = \int_{1-x}^{1} \max\{\log(u/\rho), 0\}\, du = \begin{cases} -x\log\rho - x - (1-x)\log(1-x), & 0 \le x \le 1-\rho, \\ I_\rho, & 1-\rho \le x \le 1. \end{cases}$$

The function R_ρ is increasing and concave, and $R_\rho(1) = I_\rho$. The range $\zeta > 1 - \rho$ is the saturated branch.

Theorem 37 (Projected total variation, entropy and Hellinger profiles)**.** *Under the size concentration of Theorem 31, let $k_n/n \to \zeta \in [0,1]$, $t = O(\log n)$ and $\Theta = n\rho_n^t \to \theta \in [0,\infty)$. Then*

$$d_{\mathrm{TV}}(\textit{observed tuple}, V_n) \longrightarrow d_{\mathrm{TV}}\big(\mathrm{Pois}(\zeta(1+\theta)), \mathrm{Pois}(\zeta)\big),$$
$$D_1(\textit{observed tuple}\|V_n) \longrightarrow \zeta\big((1+\theta)\log(1+\theta) - \theta\big),$$

and the squared Hellinger distance tends to $1 - \exp\big(-\frac{\zeta}{2}(\sqrt{1+\theta} - 1)^2\big)$. The limiting density is $e^{-\zeta\theta}(1+\theta)^Z$ with $Z \sim \mathrm{Pois}(\zeta)$, the likelihood ratio of $\mathrm{Pois}(\zeta(1+\theta))$ to $\mathrm{Pois}(\zeta)$.

Theorem 38 (Sharp projected cutoff)**.** *For the permanent law with $\rho_n = \eta_n/n \to \rho \in (0,1)$ and every fixed $1 < \alpha < \infty$, the projected Rényi divergence of order α at $t = \lfloor s\log n\rfloor$ tends to zero for $s > s_{\alpha,\zeta}$ and to infinity for $0 \le s < s_{\alpha,\zeta}$, where*

$$s_{\alpha,\zeta} = \max\left\{\frac{1}{\ell}, \frac{\beta\zeta}{R_\rho(\zeta)}\right\}.$$

At order infinity the same holds with $s_{\infty,\zeta} = \max\{2/\ell, \zeta/R_\rho(\zeta)\}$. *When* $\beta\zeta/R_\rho(\zeta) < 1/\ell$, *the projected divergence of order* α *has the Poisson profile*

$$D_\alpha \longrightarrow \frac{\zeta[(1+\theta)^\alpha - 1 - \alpha\theta]}{\alpha - 1}$$

at the total variation window $\Theta = n\rho_n^t \to \theta \in (0, \infty)$.

At $\zeta = 1$ this is Theorem 33. If $\zeta \le 1 - \rho$ and $1 < \alpha \le 2$, the Poisson scale governs, since then $\beta\zeta/R_\rho(\zeta) < 1/\ell$ (Section 65.15).

Theorem 39 (Exact projected profiles on the saturated branch)**.** *For the permanent law with* $\rho_n = \eta_n/n \to \rho \in (0,1)$ *and* $k_n/n \to \zeta$, *let* $1 < \alpha < \infty$ *be fixed,* $\zeta > 1 - \rho$, $t = O(\log n)$, *and suppose* $\Theta = n\rho_n^t$ *and*

$$Q_{\alpha,k} = (n)_{k_n}^{\beta} \big(P_m(b_n)/P_n(b_n)\big)^t$$

are bounded. Then

$$f_t^{\mathrm{proj}} = e^{-\zeta\Theta}(1+\Theta)^{F_k} + (n)_{k_n}\big(P_m/P_n\big)^t \mathbf{1}_{\{x_0\}} + o_{L^\alpha(V_n)}(1),$$

$$\int (f_t^{\mathrm{proj}})^\alpha \, dV_n = \exp\big(\zeta[(1+\Theta)^\alpha - 1 - \alpha\Theta]\big) + Q_{\alpha,k}^\alpha + o(1).$$

For $1 < \alpha \le 2$ *the same two formulas hold for* $0 < \zeta \le 1 - \rho$, *with* $Q_{\alpha,k} \to 0$ *in these windows.*

Theorem 40 (Projected uniform window)**.** *For the permanent law with* $\rho_n = \eta_n/n \to \rho \in (0,1)$ *and* $k_n/n \to \zeta \in (0,1]$, *let* $t = O(\log n)$, *and suppose* $W = n^2\rho_n^t$ *and* $Q_{\infty,k} = (n)_{k_n}(P_m/P_n)^t$ *are bounded. Then, uniformly over tuples,*

$$f_t^{\mathrm{proj}} = \exp(n\rho_n^t F_k) + Q_{\infty,k}\mathbf{1}_{\{x_0\}} + o(1).$$

In particular the peak density is $\exp(\zeta W) + Q_{\infty,k} + o(1)$, *attained at* x_0, *and the projected separation distance tends to zero.*

The one-step probability that every tracked position is fixed is exactly $P_m(b_n)/P_n(b_n)$, and $\log(P_n/P_m) = nR_\rho(\zeta) + o(n)$. The identity spike of the full walk therefore projects to the atom of x_0, and the rate $R_\rho(\zeta)$ of that atom replaces I_ρ. On the saturated branch $R_\rho(\zeta) = I_\rho$: observing a fraction ζ of the positions multiplies both the identity scale and the fixed-point intensity of the Poisson tilt by ζ, and keeps the total variation location $\log n/\ell$.

65.4 Relation to the literature

Cutoff was isolated by Diaconis [42], with random transpositions and riffle shuffles among the first examples. For random walks on random regular graphs Lubetzky and Sly proved total variation cutoff at $\frac{d}{d-2}\log_{d-1} n$ with a $\sqrt{\log n}$ window for simple random walk, and cutoff at $\log_{d-1}(dn)$ with a bounded window for the non-backtracking walk [43, Theorems 1–2]; the subset shuffle likewise has a bounded window. For reversible chains Chen and Saloff-Coste showed that max-L^p cutoff for $1 < p \le \infty$ is equivalent to the product condition [44]; their framework allows the cutoff time to vary with p, and for a reversible walk on a group $\|f_{2t} - 1\|_\infty = \|f_t - 1\|_2^2$, so the L^2 and L^∞ locations always differ by the familiar factor two (here $s_\infty = 2s_2$ when the identity term governs). The new content of Theorem 33 is the explicit critical order $\alpha_\rho = \ell/(\ell - I_\rho) > 1$ (equal to $2\log 2$ at $\rho = 1/2$, and exceeding 2 for small ρ): every order $1 < \alpha < \alpha_\rho$ shares the total variation location $\log n/\ell$, and above α_ρ the location moves continuously as $\beta \log n/I_\rho$.

Salez proved entropic cutoff criteria through varentropy for non-negatively curved chains [45], and Bristiel and Caputo determine the spectral gap and entropy factorization for mean-field block shuffles [37, Theorems 1.8–1.9]; Theorem 34 gives the full relative-entropy profile of the subset shuffle. Nestoridi and Olesker-Taylor extended Teyssier's spectral method to limit profiles of general reversible chains and random walks on homogeneous spaces [46]; the partial-observation results of Section 65.3 are limit profiles on the homogeneous space S_n/S_m. Teyssier obtains the Poisson profile for transpositions, individual conjugacy classes with a macroscopic fixed-point count, and a specific random involution mixture [47, 48]. A uniform subset shuffle mixes many conjugacy classes; its Fourier eigenvalues are subgroup-invariant dimensions, which control the fixed-point tilt, and the explicit one-step identity probability $1/P_n$ gives the second term of (67).

65.5 Truncation and all representation levels

Condition a step on $|J_n/n - \rho_n| \le \delta_n$. At $t = O(\log n)$ this changes the walk's law by $o(1)$ in total variation, by coupling and a union bound. Denote the conditioned kernel by T_n. Its measure is central, and its Fourier eigenvalues a_λ are nonnegative because the kernel is a mixture of subgroup averages.

Write $L = \lambda_1 = n - j$, and let $\nu = (\lambda_2, \lambda_3, \dots)$ be the diagram below the top row, of size j. If the selected size is $k = n - J_n$, its Fourier eigenvalue on the irreducible of shape λ is $f^{\lambda/(k)}/f^\lambda$, the proportion of standard tableaux having their first k entries in the top row. This is zero if $k > L$. If $k \le L$, there are at most $\binom{n-k}{j} f^\nu$ such tableaux. Choose the j entries below the top row and fill them in a relative standard tableau, ignoring any additional top-row inequalities.

The hook-length formula gives

$$f^\lambda = \binom{n}{j} f^\nu / h_\lambda, \qquad h_\lambda = \prod_{\ell=1}^{L} \left(1 + \frac{\lambda'_\ell - 1}{L - \ell + 1}\right).$$

The column heights $\lambda'_\ell - 1$ are decreasing, and the reciprocal denominators are increasing. The reversed Chebyshev sum inequality therefore gives

$$\log h_\lambda \le \sum_{\ell=1}^{L} \frac{\lambda'_\ell - 1}{L - \ell + 1} \le \frac{j H_L}{L}, \qquad H_L = \sum_{\ell=1}^{L} \ell^{-1}.$$

For $a_\lambda \ne 0$, the conditioning implies $L \ge n(1 - \rho_n - \delta_n)$, a fixed positive fraction of n. Also $\binom{n-k}{j}/\binom{n}{j} \le ((n-k)/n)^j$. Consequently

$$0 \le a_\lambda \le \left[(\rho_n + \delta_n) \exp(C \log n/n)\right]^j$$

with C depending only on a fixed compact subinterval containing the ρ_n for large n. Summing squared dimensions over each level j gives at most

$$\binom{n}{j}^2 \sum_{\nu \vdash j} (f^\nu)^2 = \binom{n}{j}^2 j! \le n^{2j}/j!.$$

Thus, at bounded $n\rho_n^t$ and $t = O(\log n)$, the sum of the squared Fourier contributions beyond level M is bounded by

$$\sum_{j > M} \frac{1}{j!} \left[n^2 \big((\rho_n + \delta_n) \exp(C \log n/n)\big)^{2t}\right]^j.$$

The bracket tends to θ^2 at the asserted times, since $\delta_n \log n \to 0$. The tail tends to zero as $M \to \infty$, uniformly at times with the bracket bounded. For $n\rho_n^t \to 0$ and $t = O(\log n)$, the entire nontrivial square sum tends to zero.

65.6 Fixed levels and the fixed-point density

At a fixed j and large n, the conditioned block sizes k are between j and L. Removing the first k top-row boxes separates the remaining top-row chain from the lower diagram ν. Hence there are exactly $\binom{n-k}{j} f^\nu$ skew tableaux. Also $h_\lambda = 1 + O_j(1/n)$, since the first j factors contain all the lower boxes. Uniformly for the conditioned sizes, their eigenvalues are therefore

$$\rho_n^j \left(1 + O_j(\delta_n + 1/n)\right).$$

Taking the mixture and then raising to $t = O(\log n)$ gives $a_\lambda^t = (1 + o(1))\rho_n^{jt}$. The coefficient $f^\lambda a_\lambda^t$ of the irreducible character in the density relative to uniform converges to $f^\nu \theta^j / j!$.

Let $F(\sigma)$ count fixed points, and set $G_j = \sum_{\lambda:\lambda_1 = n-j} f^\nu \chi_\lambda$. The falling factorial $(F)_k$ is the character of the action on ordered distinct k-tuples. By the branching rule it is the induced trivial character from S_{n-k}. When $n \geq 2k$, the corresponding skew shape for level $j \leq k$ again separates a top row from ν, giving multiplicity $\binom{k}{j} f^\nu$. Thus

$$(F)_k = \sum_{j=0}^{k} \binom{k}{j} G_j, \qquad G_j(F) = \sum_{k=0}^{j} (-1)^{j-k} \binom{j}{k} (F)_k.$$

These are the fixed-point polynomials used for the Poisson profile.

For each fixed M, the density truncated at level M is within $o(1)$ in L^2 of $\sum_{j=0}^{M} \theta^j G_j(F)/j!$. Under uniform measure on S_n, F converges to Pois(1). The expectations of its fixed-degree polynomials, and their absolute deviations from one, converge because its falling factorial moments of every fixed order are one for all large n. The preceding square-sum estimate controls the discarded levels.

For $Z \sim \mathrm{Pois}(1)$, the generating series is

$$\sum_{j \geq 0} \frac{\theta^j}{j!} G_j(Z) = e^{-\theta} (1 + \theta)^Z.$$

The series converges in L^2: its polynomials are orthogonal with squared norm $j!$, as follows by expanding $\mathbb{E}[e^{-u-v}(1+u)^Z(1+v)^Z] = e^{uv}$. The same approximation holds in $L^2(U_n)$ with the finite-n tilt $e^{-\Theta}(1+\Theta)^F$, uniformly for bounded $\Theta = n\rho_n^t$ and $t = O(\log n)$: the conditioned Fourier tail has the factorial bound above, and the binomial series of the tilt is absolutely summable in L^2, uniformly for bounded Θ, since $\|(F)_k\|_2 \leq C 2^{k/2} \sqrt{k!}$ by $\Pr(F = k + l) \leq 1/(k+l)!$ and $\binom{k+l}{k} \leq 2^{k+l}$; multiplication by the exponential series and binomial inversion give the same fixed Charlier coefficients. Taking the L^1 deviation from one proves the profile after $M \to \infty$. The limiting density is exactly the likelihood ratio of $\mathrm{Pois}(1+\theta)$ with respect to $\mathrm{Pois}(1)$.

65.7 Early times and the permanent case

At $t = O(\log n)$ with $n\rho_n^t \to \infty$, in the conditioned walk let U count positions that were never selected. Its expectation is np_n^t, where $p_n = \mathbb{E}[J_n]/n = \rho_n + O(\delta_n)$. Its two-position survival probability uses $q_n = \mathbb{E}[J_n(J_n - 1)]/[n(n-1)]$ in place of p_n^2. Concentration in the conditioned interval gives $q_n/p_n^2 = 1 + O(\delta_n^2 + 1/n)$. Thus $\operatorname{Var} U/(\mathbb{E}U)^2 \to 0$. The number of fixed points is at least U, whereas under uniform measure its mean is one. This distinguishes the laws at early times. At late times the square-sum upper bound applies and contraction extends it beyond $O(\log n)$. Together with the profile, this yields cutoff with a bounded window.

For positive equicorrelation the weight of a complement of size j is proportional to $\eta_n^j/j!$, with $\eta_n = (1-b_n)/b_n$. If $\eta_n/n \to \rho \in (0,1)$, a standard Poisson tail bound gives the hypothesis with $\delta_n = n^{-1/4}$. At the scalar cofactor optimizer the published Poisson formula gives $\rho_n = \eta_n/n = 1/2+O(1/n)$, so $\rho_n^t = 2^{-t}(1+o(1))$ for $t = O(\log n)$. Therefore at integer times for which $n2^{-t_n} \to \theta$, the profile above applies with the same θ. The result uses the subset-shuffle structure. Discrete times near cutoff can have several lattice subsequential limits.

65.8 Approximation by a uniform shuffle outside the untouched set

The Fourier comparison gives a more direct statement for varying block sizes. Fix a deterministic size schedule $K_1, \dots, K_t$, with each $K_s \geq cn$ for a fixed $c > 0$, and $1 \leq t = O(\log n)$. At step s choose an independent uniform K_s-subset and uniformly permute it. Let U be the random set of positions never selected and put $J_s = n - K_s$. Assume $\Theta_n = n\prod_s (J_s/n)$ is bounded. The law of the resulting permutation differs by $o(1)$ in total variation from the law that first draws U, fixes it, and uniformly permutes its complement. This approximation concerns the law averaged over the selected subsets for the given size schedule.

To prove it, conditional on U's size the set U is uniform. For fixed j with n large, every active set in either law has size large enough to separate the first row from the lower diagram after skewing. The level-j eigenvalue of the true schedule is

$$h_\lambda^t \prod_{s=1}^{t} \frac{(J_s)_j}{(n)_j},$$

including zero when some $J_s < j$. The surrogate eigenvalue is

$$h_\lambda \frac{\mathbb{E}[(|U|)_j]}{(n)_j} = h_\lambda \prod_{s=1}^{t} \frac{(J_s)_j}{(n)_j},$$

by independence of the chosen sets. At fixed j, $h_\lambda^{t-1} = 1 + O_j(t/n)$. The high levels for both kernels are controlled by the previous hook upper bound, with $n\rho_n^t$ replaced by Θ_n. In particular, the squared tails have factorial bounds uniformly when Θ_n is bounded. Comparing fixed levels and then taking the tail cutoff to infinity proves convergence of the difference of densities in L^2.

The surrogate density relative to uniform is $\mathbb{E}[(F)_{|U|}]$, a function only of the fixed count F. This follows by counting the $|U|$-subsets of the fixed points of any given permutation. In the surrogate the fixed count has law $|U| + \mathrm{Fix}(\mathrm{Unif}_{n-|U|})$, with the second term conditionally uniform on the complement. Its law is within $o(1)$ in total variation of $|U| + Z$, where $Z \sim \mathrm{Pois}(1)$ is independent, because $n - |U| \geq cn$. Therefore if $|U|$ converges in law to an integer-valued W, the total variation profile is the distance of $W + Z$ from Z. The approximation also passes to random size schedules when the hypotheses hold with high probability and Θ_n is tight, by conditioning and truncation. In the concentrated setting above, the exact factorial moments of $|U|$ converge to θ^j, recovering the same Poisson profile.

65.9 The unconditioned L^2 scale for the permanent kernel

The rare small-block steps can affect the unconditioned square distance. For the permanent law with $\rho_n = \eta_n/n \to \rho \in (0,1)$, define

$$I_\rho = -\log\rho - 1 + \rho, \qquad s_\rho = \max\left\{\frac{1}{\log(1/\rho)}, \frac{1}{2I_\rho}\right\}.$$

At $t = \lfloor s \log n \rfloor$, the unconditioned chi-square divergence tends to zero for every $s > s_\rho$, and tends to infinity for $0 \le s < s_\rho$. In particular, at the cofactor optimizer $s_{1/2} = 1/[2(\log 2 - 1/2)]$, strictly later than the total variation location $1/\log 2$.

For the upper bound, let $F_\eta(k) = \Pr(\mathrm{Pois}(\eta) \le k)$. The exact truncated Poisson factorial moment and the previous hook bound give

$$a_\lambda \le h_\lambda \frac{\eta_n^j}{(n)_j} \frac{F_{\eta_n}(n-j)}{F_{\eta_n}(n)} \quad \text{for } j = n - \lambda_1.$$

Indeed, $\mathbb{E}[(J_n)_j] = \eta_n^j F_{\eta_n}(n-j)/F_{\eta_n}(n)$. Uniformly over shapes, $\log h_\lambda = O(\sqrt{n} \log n)$. If $L \le \sqrt{n}$, bound each factor by $n+1$. Otherwise use the reversed Chebyshev bound from the first part.

Write $x = j/n$. Stirling's factorial estimate, uniformly in j, and the Poisson lower-tail Chernoff bound give $\log a_\lambda \le -n R_{\rho_n}(x) + O(\sqrt{n} \log n)$, where

$$R_\rho(x) = \begin{cases} -x \log \rho - x - (1-x)\log(1-x), & 0 \le x \le 1-\rho, \\ I_\rho, & 1-\rho \le x \le 1. \end{cases}$$

For the second branch, subtract the lower-tail rate $\rho - (1-x) + (1-x)\log((1-x)/\rho)$ from the logarithm of $\eta_n^j/(n)_j$. Cancellation gives the constant I_ρ. The function R_ρ is concave, since its initial derivative is $\log((1-x)/\rho)$ and decreases to zero. Hence $R_\rho(x) \ge x I_\rho$. For any fixed $\epsilon > 0$, the sum of squared representation dimensions at level $j \ge \epsilon n$ is at most $\binom{n}{j}^2 j! = \exp(j \log n + O(n))$ as an upper bound. Therefore these levels contribute $o(1)$ when $2sI_\rho > 1$, since the negative exponent then dominates the uniform $O(t\sqrt{n} \log n)$ error.

For $1 \le j < \epsilon n$, use $\log h_\lambda \le O(j \log n/n)$, $(n)_j \ge n^j(1-\epsilon)^j$, and $F_{\eta_n}(n) = 1 - O(e^{-cn})$. This bounds a_λ by the j-th power of $\rho_n(1+o(1))/(1-\epsilon)$. Since $s > 1/\log(1/\rho)$, choose $\epsilon > 0$ small enough to make $n^2[\rho_n(1+o(1))/(1-\epsilon)]^{2t} \to 0$. The factorial square sum then tends to zero also over those levels.

The standard Fourier block contributes $(n-1)^2(\rho_n + O(1/n))^{2t}$ to chi-square. This proves the first lower bound for $s < 1/\log(1/\rho)$. For the other, the identity has single-step probability P_n^{-1}, where $P_n = \operatorname{per} A_{b_n}$. The Poisson formula and Stirling give $\log P_n = n I_{\rho_n} + O(\log n)$. Thus the t-step probability at the identity is at least P_n^{-t}, by taking only identity increments. Its contribution to chi-square is at least $n!(P_n^{-t} - 1/n!)^2$ once $P_n^{-t} \ge 1/n!$. This goes to infinity for $s < 1/(2I_\rho)$.

65.10 Two contributions to the square-distance profile

With $t = O(\log n)$, set $\Theta_n = n\rho_n^t$ and $Q_n = n! P_n^{-2t}$. Whenever both sequences are bounded, the unconditioned chi-square divergence satisfies

$$\chi^2(\mu_n^{*t}, \mathrm{Unif}(S_n)) = e^{\Theta_n^2} - 1 + Q_n + o(1).$$

This refinement uses the normalized character estimate of Féray and Śniady. See [49, Theorem 1].

For the low levels, fixed j gives the same Fourier asymptotics as in the conditioned proof. Indeed the removed probability is exponentially small in $\sqrt{n}$, while the eigenvalues at fixed j stay bounded away from zero. Thus the level-j square sum equals $\Theta_n^{2j}/j! + o(1)$ for every fixed $j \ge 1$. Up to $j \le \sqrt{n}$, the factorial moment bound dominates these levels by a uniform factorial tail, since $(n)_j \ge [n(1-n^{-1/2})]^j$ and $h_\lambda \le \exp(O(j \log n/n))$. For $\sqrt{n} < j \le \epsilon n$ and small fixed $\epsilon > 0$, the same bound gives

$$\sum_{\lambda : \lambda_1 = n-j} (f^\lambda)^2 a_\lambda^{2t} \le \frac{\Theta_n^{2j}}{j!} \exp\bigl(O(\epsilon j t + j \log^2 n/n)\bigr).$$

Since $t = O(\log n)$ and $j! \geq (j/e)^j$, choosing ϵ sufficiently small makes these levels vanish in sum. This proves the stated contribution $e^{\Theta_n^2} - 1$ from levels at most ϵn.

For the remaining modes let D be the larger of the first row and first column lengths. Put $M_\lambda = P_n a_\lambda$. In the unnormalized subgroup expansion for M_λ, discard all active sizes $k > L = \lambda_1$, because they have zero invariant dimension. The identity coefficient of the remaining sum is $\Pr(\mathrm{Bin}(n, b_n) \leq L)$. Each other permutation has support $m \leq L$ and coefficient at most b_n^m. By the character bound, for $D \leq \epsilon_0 n$ with sufficiently small fixed $\epsilon_0 > 0$, its normalized character is at most $(CD/n)^{m/2}$ in absolute value. Its transposition length is between $m/2$ and m. There are at most n^m permutations with support m, and nb_n is bounded. The sum over $m \geq 2$ shows

$$|M_\lambda - \Pr(\mathrm{Bin}(n, b_n) \leq L)| = O(D/n) \quad (D \leq \epsilon_0 n).$$

Under Plancherel weights $(f^\lambda)^2/n!$, the event $D \geq n^{3/4}$ has exponentially small probability on the scale $n^{3/4} \log n$. To see this for rows, the dimension bound gives mass at most $\binom{n}{L}/L!$ at any first-row length L, and $\binom{n}{L}/L! \leq (e^2 n/L^2)^L$. Apply conjugation of shapes for columns. On the remaining Plancherel mass, $L \geq n^{1/4}$ since every box lies in a rectangle of sides at most D. Thus the binomial probability is $1 - o(n^{-1/4})$, and $M_\lambda = 1 + O(n^{-1/4})$. Raising to $2t$ leaves $1 + o(1)$. These levels are disjoint from the low levels and contribute $Q_n(1 + o(1))$. On $n^{3/4} < D \leq \epsilon_0 n$ the bound $M_\lambda = O(1)$ and the Plancherel tail make the contribution $o(1)$.

If $D > \epsilon_0 n$ and $L \leq \rho_n n$, the rate bound above gives $\log M_\lambda \leq O(\sqrt{n} \log n)$. The macroscopic Plancherel tail is exponentially small on the scale $n \log n$, so these terms contribute $o(1)$ when Q_n is bounded. Finally, for $L > \rho_n n$ and $j > \epsilon n$, the concavity bound on the first branch strengthens to $R_{\rho_n}(j/n) \geq (j/n) I_{\rho_n}/(1 - \rho_n)$. Boundedness of Q_n forces $t \geq (1/(2I_{\rho_n}) - o(1)) \log n$, so the strengthened rate kills these levels by the dimension bound. This accounts for all irreducibles.

In the regime $1/(2I_\rho) > 1/\log(1/\rho)$, the Θ_n term vanishes near the later square-distance transition. The Q_n term measures the residual contribution of the probability of the identity. The two parameters retain lattice effects and can be used directly without replacing either transition by a continuous-time window.

65.11 Proof of the finite-order profiles

Throughout this section $1 < \alpha < \infty$ is fixed and Θ, Q_α are bounded as in Theorem 33. The permanent formula gives

$$P_n = n! b_n^n \sum_{j=0}^{n} \frac{\eta_n^j}{j!}, \qquad \log P_n = n I_{\rho_n} + O(\log n). \tag{68}$$

The logarithmic estimate follows from Stirling's formula and the exponential upper-tail bound for $\mathrm{Pois}(\eta_n)$. In particular any subsequential limit s_* of $t/\log n$ satisfies $s_* \geq 1/\ell$ and $s_* \geq \beta/I_\rho$. We split the density into retained schedules, the all-identity event, and the remaining paths.

Moment bounds for retained schedules. For a deterministic schedule of untouched sizes $J_1, \ldots, J_t$, put

$$y = \left(\prod_i J_i/n \right)^{1/t}, \qquad \Theta' = n y^t \exp(Ct \log n/n).$$

Consider schedules with $y \leq y_0 < 1$. At least one selected size is at least $(1 - y_0)n$, so any shape with nonzero eigenproduct has $\lambda_1 \geq (1 - y_0)n$. The tableau and hook estimates of Section 65.5 bound

its character coefficient at level $j = n - \lambda_1$ by $f^\nu(\Theta')^j/j!$, taking C large enough. The character $(F)_j$ contains this irreducible with multiplicity f^ν, because removing all the first-row boxes leaves ν. Thus the expansion of $(1+\Theta')^F = \sum_j (\Theta')^j (F)_j/j!$ dominates all character coefficients of the schedule density. Both expansions have nonnegative coefficients. Tensor product multiplicities are nonnegative, so for every positive integer r,

$$\|f_{\mathrm{sched}}\|_r^r \le \mathbb{E}_{U_n}(1+\Theta')^{rF} \le \exp\big((1+\Theta')^r - 1\big). \tag{69}$$

The last bound uses $\mathbb{E}_{U_n}(F)_j = 1$ for $j \le n$.

Retain schedules for which every $|J_i/n - \rho_n| \le n^{-1/4}$. Their discarded probability is $O(te^{-c\sqrt{n}})$. On the retained event Θ' stays bounded, and (69) and Minkowski's inequality give uniform bounds on all fixed integer moments. The L^2 approximation in Section 65.6 therefore gives L^α approximation by $e^{-\Theta}(1+\Theta)^F$, by interpolation against a higher integer moment if needed. The tilt itself has bounded higher moments by the fixed-point factorial identities. The same approximation holds for the subprobability on retained paths since their probability tends to one.

Paths with few actual moves. The actual one-step probability of a permutation with support size m is $P_n^{-1} b_n^m$. Write m_i for each actual support size and $M = \sum_i m_i$. The event $M = 0$ gives the identity density $n! P_n^{-t} \mathbf{1}_{\{e\}}$. Conditional on an identity step its latent selected size is $\mathrm{Bin}(n, b_n)$, since a subset of size k produces the identity with probability $1/k!$. The binomial has bounded mean, and the retained sizes are macroscopic. The contribution from all-identity paths on the retained event therefore has L^α norm $o(Q_\alpha) = o(1)$.

For $0 < M \le \delta n$, condition on the actual support sizes. The supports are independent uniform sets, and the size sequence probability is at most $P_n^{-t}(nb_n)^M$. Expose the selected positions one at a time, sampling without replacement within each step. The number Δ of repetitions is M minus the size of the union. Every exposed position repeats with conditional probability at most $M/(n-M)$, hence

$$\Pr(\Delta \ge M/4 \mid (m_i)) \le 2^M \left(\frac{M}{n-M}\right)^{M/4}. \tag{70}$$

On the complementary event more than $M/2$ positions are touched exactly once. Such a position is moved by the final permutation because its own item changes at its sole step and never changes thereafter. The resulting law is conjugation invariant. A conjugacy class with k moved positions has size at least $(n)_k/k^{k/2}$, since its nontrivial cycle multiplicities a_j satisfy $\prod_{j\ge 2} j^{a_j} a_j! \le k^{\sum a_j} \le k^{k/2}$. For $M/2 \le k \le M \le \delta n$ this is at least $\exp(cM \log n)$ for an absolute $c > 0$ and sufficiently large n. Thus this part's norm in counting ℓ^α is at most $\exp(-c\beta M \log n)$. The norm on the repetition event is bounded by its probability.

If $M \le \sqrt{n}$, there are at most t^M size sequences, and the right side of (70) is at most $(Cn^{-1/8})^M$. Both parts are summably small after multiplying by $t^M (nb_n)^M$. If $\sqrt{n} < M \le \delta n$, the sequence count is at most $(M+t)^t = \exp(O((\log n)^2))$. Since $nb_n = O(1)$, choose $\delta > 0$ small enough that the repeat bound times $(nb_n)^M$ decays exponentially in M. The nonrepeat bound also decays after the same multiplication. Conversion to density norms multiplies the sum by $Q_\alpha = (n!)^\beta P_n^{-t} = O(1)$, so the subprobability with $0 < M \le \delta n$ has density of norm $o(1)$.

Exceptional schedules with many moves. It remains to handle nonretained paths with $M > \delta n$. Their total latent selected size is at least M, so $u := 1 - y \ge \delta/t$ by arithmetic–geometric mean. We can drop the constraint on actual moves while keeping this latent constraint, because all subprobability densities are nonnegative.

For $y \le \rho_n + (\log n)^{-2}$, Θ' stays bounded. The exceptional probability tends to zero, so (69) gives a vanishing contribution. Along a subsequence assume $t/\log n \to s_*$. Choose an integer $r > \alpha$ and $y_0 > \rho$ sufficiently close to ρ that $r(1 + s_* \log y_0) < 1$. For $\rho_n + (\log n)^{-2} < y \le y_0$ the norm is at most $\exp(O(n^a))$ for some $a < 1$. The probability is at most $\exp(-cn/(\log n)^3)$. In fact $y \ge x$ implies $\sum_i J_i \ge ntx$. Before conditioning the sizes at n, this sum is Poisson of mean $nt\rho_n$. The conditioning changes the probability bound by a factor $1 + o(1)$; the resulting estimate is

$$\Pr(y \ge x) \le (1 + o(1))e^{-ntI_{\rho_n}(x)} \quad (x > \rho_n). \tag{71}$$

The quadratic rate near ρ_n gives the asserted middle-range bound. It dominates the schedule norms.

For $y \ge y_0$, nonzero eigenproducts are bounded by $h_\lambda^t y^{tj}$ and require $j \le \min_i J_i \le ny$. Uniformly in shape, $\log h_\lambda = O(\sqrt{n}\log n)$, by bounding the hook factors individually if $\lambda_1 \le \sqrt{n}$ and otherwise using Section 65.5. If $\alpha \le 2$, interpolate L^1 and L^2, so the Plancherel sum for the schedule is raised to power β. If $\alpha \ge 2$, inverse Fourier interpolation from the weighted ℓ^1 and ℓ^2 norms, with weights $(f^\lambda)^2$, bounds $\|f_{\mathrm{sched}}\|_\alpha$ by $\big(\sum_\lambda (f^\lambda)^2 a_{\lambda,\mathrm{sched}}^{1/\beta}\big)^\beta$. The endpoint bounds follow from the character bound and orthogonality. Thus set $\kappa = 2$ for $\alpha \le 2$ and $\kappa = 1/\beta$ otherwise. The dimension bound at level j gives

$$\log \|f_{\mathrm{sched}}\|_\alpha \le \beta ny[\log n + \kappa t \log y]_+ + O(n). \tag{72}$$

We compare (72) with (71). Where $1 + \kappa s_* \log y > 0$, put $a = \log(y/\rho)$ and $d = \ell - 1/(\kappa s_*)$. The ratio

$$\frac{s_* I_\rho(y)}{y(1 + \kappa s_* \log y)} = \frac{a - 1 + e^{-a}}{\kappa(a - d)}$$

has derivative with sign $1 - d + e^{-a}(d - a - 1)$. This last expression increases with a in the stated range. At $a = \ell$, it is at most $1 - \rho - \ell(1 + \rho)/2 < 0$ for $\alpha \le 2$, using $s_* \ge 1/\ell$ and $\tanh(\ell/2) < \ell/2$. For $\alpha \ge 2$ it is at most $\rho(1 - \rho - \ell) < 0$, since $s_* \ge \beta/I_\rho$ gives $d \ge 1 - \rho$. Thus the ratio decreases strictly to $s_* I_\rho$ at $y = 1$. Where the positive part is zero the probability rate is strictly positive. If $s_* > \beta/I_\rho$, there is a uniform negative exponent on $[y_0, 1]$ at scale $n \log n$.

In the remaining case $s_* = \beta/I_\rho$, the same comparison has a negative linear margin in $u = 1 - y$ near zero, relative to its endpoint exponent at $y = 1$. Indeed the ratio has strictly negative derivative at one, so the normalized combined exponent has strictly positive derivative there. For the actual parameters $t/\log n, \rho_n$ this derivative remains bounded below near one. Away from one the limiting exponent is strictly negative. Also bounded Q_α gives

$$\beta \log(n!) - ntI_{\rho_n} = \log Q_\alpha + O((\log n)^2) \le O((\log n)^2). \tag{73}$$

Thus (72) suffices down to $u = C/\log n$ with C large, by absorbing the $O(n)$ error.

For the last range $u \le C/\log n$, use the first-row dimension bound under Plancherel measure,

$$\Pr_{\mathrm{Pl}}(\lambda_1 = L) \le \binom{n}{L}/L! \le (e^2 n/L^2)^L.$$

For $L \ge n/(\log n)^2$ its log is at most $-L(\log n - 4\log\log n - 2)$; below that use the trivial bound one, which adds only $o(n)$ to the same estimate. Every contributing shape satisfies $L \ge nu$. The coefficient of L in the combined dimension and Fourier exponent is negative for large n. Summing over L thus bounds the logarithm of the norm by

$$\beta\{\log(n!) + \kappa t n \log y - nu(\log n + \kappa t \log y - 4\log\log n - 2)\} + o(n).$$

Combining with (71) and expanding at $u = 0$ gives

$$\beta \log(n!) - ntI_{\rho_n} + nu\{t\ell_n - \beta(\log n + \kappa t)\} + o(n), \qquad \ell_n = \log(1/\rho_n).$$

For $\alpha \le 2$, the braces divided by $\log n$ tend to $s_*(1-\rho-2\beta) < 0$, since $\beta = s_* I_\rho \ge I_\rho/\ell > (1-\rho)/2$. For $\alpha \ge 2$, the limit is $s_*(\ell-1)-\beta = -s_*\rho < 0$. Now $u \ge \delta/t$, so (73) gives a negative bound of order n. In each high range one can sum the probability-times-norm bounds over bins of length $(\log n)^{-2}$, using the probability bound at the lower endpoint of a bin and the norm estimate maximized over it. The exponent variation in a bin is $o(n)$, by the uniform $O(n \log n)$ Lipschitz bounds for the displayed leading expressions. There are $O((\log n)^2)$ bins. Their total tends to zero.

The power profile and the mixing scale. Minkowski's inequality now gives (66) by combining the retained paths, the all-identity event and the two negligible pieces. The tilt's contribution to the α-moment is

$$e^{-\alpha\Theta} \sum_{k=0}^{n} \frac{((1+\Theta)^\alpha - 1)^k}{k!} = \exp((1+\Theta)^\alpha - 1 - \alpha\Theta) + o(1),$$

uniformly for bounded Θ. Its own L^α mass at e tends to zero, since $g(e) \le (1 + \sup \Theta)^n$ and $U_n(e) = 1/n!$. The spike gives Q_α^α, proving (67).

For $s > s_\alpha$, both parameters vanish, so the divergence tends to zero. Below β/I_ρ, $Q_\alpha \to \infty$, and the all-identity event gives a diverging lower bound. Below $1/\ell$, the untouched count from Section 65.7 diverges in probability. There is a diverging fixed-point threshold whose walk probability tends to one and whose uniform probability tends to zero. Binary data processing gives divergence of relative entropy, and monotonicity in the order gives the Rényi lower bound.

65.12 Proof of the entropy and lower-order profiles

Let g_n be the retained density of Section 65.5. By Section 65.6 it converges in distribution under U_n to $G_\theta = e^{-\theta}(1+\theta)^Z$, $Z \sim \mathrm{Pois}(1)$, and has uniformly bounded second moments. The function $x \log x$, extended continuously to zero, has bounded negative part and is $o(x^2)$ at infinity. Consequently $\int g_n \log g_n \, dU_n \to \mathbb{E} G_\theta \log G_\theta$. The same argument applies to g_n^α for $0 < \alpha < 1$. Since G_θ is the Poisson likelihood ratio, direct evaluation gives

$$\mathbb{E} G_\theta \log G_\theta = (1+\theta)\log(1+\theta) - \theta, \qquad \mathbb{E} G_\theta^\alpha = \exp((1+\theta)^\alpha - 1 - \alpha\theta).$$

To remove conditioning for relative entropy, let ν be the one-step retained law, ω the exceptional law, and ϵ_n the exceptional probability. Put $q_n = (\rho_n + \delta_n)\exp(C \log n/n)$. The all-level square bound gives

$$D_1(\nu^{*s} \| U_n) \le \log(1 + \chi^2(\nu^{*s}, U_n)) \le n^2 q_n^{2s}.$$

This also holds at $s = 0$ since $\log(n!) \le n^2$. Conditional on the number V of exceptional steps, the final law is $\nu^{*(t-V)} * \omega^{*V}$. Centrality makes the order of these factors immaterial. Convolution contracts relative entropy. Since nq_n^t is bounded and $q_n \ge q_- > 0$ for large n, the weighted entropies from $V \ge 1$ sum to at most a constant times

$$(1 - \epsilon_n + \epsilon_n q_-^{-2})^t - (1-\epsilon_n)^t = o(1).$$

Here V is binomial and $t\epsilon_n \to 0$. Convexity gives the upper bound for the full entropy. For a lower bound, the entropy of the mixture is at least the weighted entropy of the all-retained component

minus the binary entropy of $1-(1-\epsilon_n)^t$, by the chain rule and the mutual information bound for the indicator of an exceptional path. This proves the relative entropy profile. For $0<\alpha<1$, the inequality $|x^\alpha-y^\alpha|\le|x-y|^\alpha$ and Jensen's inequality transfer the affinity using the $o(1)$ total variation difference. The case $\alpha=1/2$ gives the stated Hellinger profile. This proves Theorem 34.

65.13 Proof of the uniform window

Let a_λ be the nonnegative unconditioned Fourier multipliers. The character bound implies

$$\|f_t-1\|_\infty = f_t(e)-1 = \sum_{\lambda\neq(n)} (f^\lambda)^2 a_\lambda^t.$$

We analyze this sum and the corresponding character expansion when Z_n, Q_∞ are bounded as in Theorem 35. The truncated Poisson factorial moment gives, for $j=n-\lambda_1$,

$$a_\lambda \le h_\lambda \frac{\eta_n^j}{(n)_j}\frac{F_{\eta_n}(n-j)}{F_{\eta_n}(n)}, \qquad F_\eta(k)=\Pr(\mathrm{Pois}(\eta)\le k).$$

Here the exact factorial moment is $\mathbb{E}(J_n)_j=\eta_n^j F_{\eta_n}(n-j)/F_{\eta_n}(n)$. For $x=j/n$, Stirling and the Poisson lower-tail estimate give $\log a_\lambda \le -nR_{\rho_n}(x)+O(\sqrt{n}\log n)$, where

$$R_\rho(x)=\begin{cases} -x\log\rho - x - (1-x)\log(1-x), & 0\le x\le 1-\rho,\\ I_\rho, & 1-\rho\le x\le 1.\end{cases}$$

Indeed the factorial ratio gives the first branch. The additional lower-tail exponent when $1-x<\rho$ cancels its x-dependence. The first branch is concave and joins the constant second branch with derivative zero.

For each fixed j, $f^\lambda a_\lambda^t$ at level j is $f^\nu(n\rho_n^t)^j(1+o(1))/j!$, uniformly over its fixed shapes, by the fixed-level argument of Section 65.6. The exponentially small size truncation error has no effect there. The ordered-tuple identity in that section shows

$$\sum_{\lambda:\lambda_1=n-j} f^\nu\chi_\lambda = (F)_j + O_j(n^{j-1})$$

uniformly for fixed $j\ge1$: the omitted levels have bounded multiplicities and dimensions $O_j(n^{j-1})$. The level therefore contributes $(n\rho_n^tF)^j/j!+o(1)$ uniformly. For $j\le\varepsilon n$, the factorial moment bound and the level dimension bound give the absolute contribution at most

$$\frac{Z_n^j}{j!}\exp\big(O(tj^2/n+tj\log n/n)\big).$$

Up to $j\le\sqrt{n}$ this gives a uniform factorial tail. For $\sqrt{n}<j\le\varepsilon n$, choose ε small enough that the factorial absorbs the error with decay of order $j\log n$. Thus the sum of the low levels tends uniformly to $\exp(n\rho_n^tF)$, including the trivial character.

For the other levels set $D=\max(\lambda_1,\lambda_1')$, $L=\lambda_1$, and $M_\lambda=P_na_\lambda$. In the subgroup expansion of the unnormalized multiplier, omit active sizes larger than L, whose invariant dimensions are zero. The remaining identity coefficient is $\Pr(\mathrm{Bin}(n,b_n)\le L)$. Each other permutation has support $m\le L$ and coefficient at most b_n^m. For $D\le\varepsilon_0 n$ with fixed small $\varepsilon_0>0$, the character estimate of Féray and Śniady [49, Theorem 1] bounds its normalized character by $(CD/n)^{m/2}$. Indeed its

transposition length lies between $m/2$ and $m \le D$. Summing using at most n^m permutations and $nb_n = O(1)$ gives

$$\left|M_\lambda - \Pr(\mathrm{Bin}(n, b_n) \le L)\right| = O(D/n) \qquad (D \le \varepsilon_0 n).$$

The Plancherel row bound in Section 65.11, also applied to conjugate shapes for columns, makes $D > n^{3/4}$ have exponentially small probability on the $n^{3/4} \log n$ scale. If $D \le n^{3/4}$ then $L \ge n^{1/4}$, giving $M_\lambda = 1 + O(n^{-1/4})$. Consequently $M_\lambda^t = 1 + o(1)$ uniformly on these shapes. On the remaining shapes with $D \le \varepsilon_0 n$, the bounded M_λ and the Plancherel tail make the Fourier contribution $o(Q_\infty)$. If $D > \varepsilon_0 n$ and $L \le \rho_n n$, then $\log M_\lambda = O(\sqrt{n} \log n)$ from the rate bound, which is absorbed by the macroscopic Plancherel tail. If $L > \rho_n n$ and $j > \varepsilon n$, concavity on the first rate branch gives $R_{\rho_n}(j/n) \ge (j/n) I_{\rho_n}/(1 - \rho_n)$. Bounded Q_∞ forces $t \ge (1/I_{\rho_n} - o(1)) \log n$. The squared-dimension level bound $\exp(j \log n + O(n))$ then makes those contributions vanish. All these estimates hold for uniform absolute character sums by $|\chi_\lambda| \le f^\lambda$.

Replacing the Plancherel-typical multipliers by P_n^{-t} costs $o(1)$, and completing their character sum to the full regular character costs at most $Q_\infty \Pr_{\mathrm{Pl}}(D > n^{3/4}) = o(1)$. The regular character is $n! \mathbf{1}_{\{e\}}$. Combining this contribution with the low levels proves the uniform approximation. Its value at e gives the asserted norm. It is at least one everywhere before the uniform error, and equals one on derangements, so the separation assertion follows in this window. Above the stated order-infinity scale both parameters vanish. Below it either the identity lower bound diverges or the standard Fourier term $(n-1)^2(\rho_n + O(1/n))^t$ diverges. The standard multiplier follows from one-box branching and the Poisson tail. This completes Theorem 35.

65.14 Proof of the critical-order transition

We prove Proposition 36. At the finite critical order $\beta = I_\rho/\ell$, suppose $(\rho_n - \rho) \log n \to 0$ and $\Theta \to \theta \in (0, \infty)$. Stirling's formula in (68) gives

$$\log Q_\alpha = \beta n(\log \theta - 1) + o(n).$$

For $\theta < e$ the Poisson Rényi profile follows from Theorem 33. For $\theta > e$ the identity lower bound diverges. At $\theta = e$ the exact value of Q_α decides, as follows.

For the rate $\rho_n = \rho + O(1/n)$ at the scalar cofactor optimizer, the permanent expansion sharpens to

$$\log P_n = n I_{\rho_n} + \tfrac{1}{2} \log(2\pi n) - 1/\rho_n + O(1/n).$$

This uses $b_n = (1 + n\rho_n)^{-1}$ and the exponentially small upper Poisson tail in the exact permanent formula. For $\Theta \to e$ this yields

$$\log Q_\alpha = \beta n(\log \Theta - 1) - \frac{(\log n)^2}{2\ell} + O(\log n).$$

In particular, if $\Theta = e + o((\log n)^2/n)$, the identity contribution vanishes. A nonzero bounded identity contribution requires $\log \Theta - 1 = (\log n)^2/(2\beta\ell n) + O(\log n/n)$. The defining formula for Q_α carries the exact integer-time dependence.

65.15 Proofs for partial observation

Quotient Fourier analysis. The ordered tuple space is the quotient S_n/S_m. The irreducible of shape λ has multiplicity $b_\lambda = f^{\lambda/(m)}$ in its permutation representation, so its isotypic component has dimension $d_\lambda = f^\lambda b_\lambda$. The projected density started at x_0 is the sum of the isotypic projector kernels at x_0, multiplied by the same central multipliers a_λ^t as in the full walk. On a transitive

space the projector kernel has constant diagonal d_λ; by Cauchy–Schwarz it is bounded in absolute value by d_λ, and its squared $L^2(V_n)$ norm is d_λ. The kernels of different shapes are orthogonal. Interpolation therefore bounds the inverse transform from coefficients weighted by d_λ in ℓ^p to $L^\alpha(V_n)$, with $p = 1/\beta$ for $2 \le \alpha < \infty$ and $p = 1$ at infinity, and for $1 < \alpha \le 2$ the interpolation between L^1 and L^2 of Section 65.11 applies. At level $j = n - \lambda_1$ the multiplicity vanishes for $j > k_n$, and skew tableaux give $b_\lambda \le \binom{k_n}{j} f^\nu$, so

$$\sum_{\lambda:\lambda_1=n-j} d_\lambda \le \binom{n}{j}\binom{k_n}{j} j! \le \exp(j \log n + O(n)).$$

The projected tilt. On the uniform space let $F_{\mathrm{rest}} = F - F_k$. Given the observed tuple, the remaining map is uniform between the m remaining positions and the m unused labels; if A counts the remaining positions whose own labels are unused, then $\mathbb{E}[(F_{\mathrm{rest}})_j \mid \text{tuple}] = (A)_j/(m)_j$. The image set of a uniform tuple is a uniform k_n-subset, so $\mathbb{E}A = m^2/n$ and $\operatorname{Var} A \le m$. Uniformly for bounded Θ,

$$\mathbb{E}\big[(1+\Theta)^{F_{\mathrm{rest}}} \mid \text{tuple}\big] = \sum_{j=0}^{m} \frac{\Theta^j}{j!}\frac{(A)_j}{(m)_j} = e^{(1-\zeta)\Theta} + o(1)$$

in probability, with the deterministic bound $e^{\sup\Theta}$; when m stays bounded, $\zeta = 1$ and $\Pr(A > 0) \le m^2/n$. Since $\mathbb{E}_{V_n}(F_k)_j = (k_n)_j/(n)_j \le 1$, all fixed exponential moments of F_k are bounded. Conditional expectation is contractive in every L^α, so it carries the retained full-space approximation of Section 65.11 to the projected tilt $e^{-\zeta\Theta}(1+\Theta)^{F_k}$. Its α-th moment is $\exp(\zeta[(1+\Theta)^\alpha - 1 - \alpha\Theta]) + o(1)$ by the factorial-moment expansion. At infinity the full-space low-level expansion $\exp(n\rho_n^t F) + o(1)$ of Section 65.13 applies to retained schedules; with $c_n = n\rho_n^t = O(1/n)$ the conditional factor of unobserved fixed points lies between 1 and $\exp(e^{c_n} - 1) = 1 + O(1/n)$, which gives the uniform projected tilt $\exp(n\rho_n^t F_k)$.

Proof of Theorem 37. For $\zeta = 0$ the factor $e^{-\zeta\Theta}(1+\Theta)^{F_k}$ tends to one in every $L^r(V_n)$ since $\mathbb{E}_{V_n}(F_k)_j = (k_n)_j/(n)_j \to 0$ for $j \ge 1$; the argument below covers both cases. The full retained density of Section 65.5 is approximated in $L^2(U_n)$ by $e^{-\Theta}(1+\Theta)^F$ (Section 65.6), with uniformly bounded integer moments by (69). Conditional expectation onto the observed tuple is contractive, so by the projected tilt computation above the retained projected density is approximated in $L^2(V_n)$ by $e^{-\zeta\Theta}(1+\Theta)^{F_k}$, with bounded second moments. Under V_n, F_k converges in law to $\mathrm{Pois}(\zeta)$ with bounded exponential moments, since its factorial moments are $(k_n)_j/(n)_j \to \zeta^j$. The discarded schedules have total mass $o(1)$, which gives the total variation limit and, by $|x^{1/2} - y^{1/2}| \le |x-y|^{1/2}$ and Jensen, the Hellinger limit. For relative entropy, $x \log x$ has bounded negative part and is $o(x^2)$, so the retained part converges to $\mathbb{E}G \log G$ for the Poisson likelihood ratio G; data processing and convexity bound the discarded part by the weighted full-space entropies of Section 65.12, which are $o(1)$, and the binary entropy of the exceptional indicator gives the matching lower bound. The Poisson divergences evaluate to the stated limits.

Upper bounds at finite orders. We prove the upper assertions of Theorem 38 at every fixed $1 < \alpha < \infty$ by decomposing the walk according to its deterministic size schedule; no truncation by Fourier level is used. Fix a schedule $K_1, \dots, K_t$, let $J_i = n - K_i$, let $y = \sum_i J_i/(nt)$ be its arithmetic retained fraction and let f_{sched} be the full density of the walk with this schedule. Its character coefficients are nonnegative, and the projected density of the walk is the mixture over schedules of the projections of f_{sched}. By the triangle inequality and contractivity of projection it

suffices to bound, over three ranges of y, the schedule probability times $\|f_{\mathrm{sched}}-1\|_{L^\alpha(U_n)}$ on the first range, times $\|f_{\mathrm{sched}}\|_{L^\alpha(U_n)}$ on the second, and times the projected norm $\|f^{\mathrm{proj}}_{\mathrm{sched}}\|_{L^\alpha(V_n)}$ on the third. Along a subsequence let $t/\log n\to s$, and assume $s>s_{\alpha,\zeta}$.

Range $y\le \rho_n+(\log n)^{-2}$. The geometric retained fraction is at most y, so the comparison behind (69) applies with $\Theta'=ny^te^{o(1)}=n\rho_n^te^{O(1/\log n)}\to 0$, since $s>1/\ell$. Let $g=(1+\Theta')^F$. The nontrivial coefficients of f_{sched} are dominated by those of g and the trivial coefficient is 1 for f_{sched} and $\mathbb{E}_{U_n}g\ge 1$ for g, so by Plancherel

$$\|f_{\mathrm{sched}}-1\|_2^2\le \|g\|_2^2-1\le \exp\bigl((1+\Theta')^2-1\bigr)-1\longrightarrow 0,$$

while $\|f_{\mathrm{sched}}\|_r^r\le \exp((1+\Theta')^r-1)$ stays bounded for a fixed integer $r>\alpha$. Hölder interpolation between L^2 and L^r gives $\|f_{\mathrm{sched}}-1\|_\alpha\to 0$ uniformly over this range.

Range $\rho_n+(\log n)^{-2}<y\le y_0$. Fix an integer $r>\alpha$ and $y_0>\rho$ with $r(1+s\log y_0)<1$. Then $\Theta'\le n^{a'}$ with $ra'<1$, so $\|f_{\mathrm{sched}}\|_\alpha\le\|f_{\mathrm{sched}}\|_r\le\exp(O(n^a))$, $a<1$, while by (71) the total probability of the range is at most $\exp(-cn/(\log n)^3)$.

Range $y>y_0$. At level $j=nx$, branching gives $a_{\lambda,K_i}\le h_\lambda(J_i)_j/(n)_j$, and terms vanish unless every $J_i\ge j$. Concavity of $\log(J-\ell')$ for each $0\le\ell'<j$ bounds the product over steps by its value at the arithmetic mean, so the logarithm of the product is at most $ntg(x,y)+O(t\log n)$, where

$$g(x,y)=y\log y-(y-x)\log(y-x)+(1-x)\log(1-x)\le 0,\qquad 0\le x\le y.$$

The hook factors contribute $t\cdot O(\sqrt{n}\log n)$, and the quotient dimensions at level $j\le k_n$ sum to $\exp(j\log n+O(n))$. For $2\le\alpha<\infty$ the quotient interpolation with weights d_λ in $\ell^{1/\beta}$ gives $\|f^{\mathrm{proj}}_{\mathrm{sched}}\|_\alpha\le(\sum_\lambda d_\lambda a_{\lambda,\mathrm{sched}}^{t/\beta})^\beta$; for $1<\alpha<2$, interpolation between L^1 and L^2 raises the quotient Plancherel sum $\sum_\lambda d_\lambda a_{\lambda,\mathrm{sched}}^{2t}$, with weights d_λ and hence only levels $j\le k_n$, to the power β. With $\kappa=1/\beta$ for $\alpha\ge 2$ and $\kappa=2$ for $\alpha<2$, and with the tail cost $\exp(-ntI_{\rho_n}(y))$ of the arithmetic retained fraction (from (71)), the logarithm of probability times norm is at most $n\log n$ times

$$E(x,y)=\beta x+s\{\beta\kappa\, g(x,y)-I_\rho(y)\}+o(1),$$

maximized over $0\le x\le\min(y,k_n/n)$. We show $\sup E<0$ uniformly on $y\ge y_0$.

For $2\le\alpha<\infty$, $\beta\kappa=1$. Since $\partial_y I_\rho(y)=\log(y/\rho)$ and $\partial_y g(x,y)=\log(y/(y-x))$, the function $y\mapsto I_\rho(y)-g(x,y)$ has increasing derivative $\log((y-x)/\rho)$ and is minimized on $[x,1]$ at $y=\min(\rho+x,1)$. Direct substitution gives $I_\rho(\rho+x)-g(x,\rho+x)=-x-x\log\rho-(1-x)\log(1-x)$ for $x\le 1-\rho$, and $I_\rho(1)-g(x,1)=I_\rho$ for $x\ge 1-\rho$. Thus

$$\min_{x\le y\le 1}\{I_\rho(y)-g(x,y)\}=R_\rho(x).$$

Hence $E(x,y)\le\beta x-sR_\rho(x)+o(1)$. Since $R_\rho(x)/x$ is nonincreasing, for $x\le\min\{\zeta+\epsilon,1\}$ we get $\beta x-sR_\rho(x)\le x(\beta-sR_\rho(\zeta+\epsilon)/(\zeta+\epsilon))\le -c_1x$ with $c_1>0$, for ϵ small, because $s>\beta\zeta/R_\rho(\zeta)$. For $x\le x_1$ use instead $g\le 0$: $E\le\beta x_1-sI_\rho(y_0)+o(1)$, which is negative for x_1 small, since $I_\rho(y_0)>0$. Thus $\sup E\le-\min\{c_1x_1,sI_\rho(y_0)-\beta x_1\}<0$.

For $1<\alpha<2$, $\kappa=2$ and $\beta<1/2$. Let $V(x)=\min_{x\le y\le 1}\{I_\rho(y)-2\beta g(x,y)\}$, so that $E\le\beta x-sV(x)+o(1)$. The expression is strictly convex in y; its minimizer solves $y^{1-2\beta}(y-x)^{2\beta}=\rho$ until it reaches 1, at $x_*=1-\rho^{1/(2\beta)}$, and $V(x)=I_\rho$ beyond. At an interior minimizer

$$V'(x)=-2\beta\log\frac{y-x}{1-x},\qquad \operatorname{sign}V''(x)=\operatorname{sign}(1-2\beta-y),$$

by differentiating the implicit equation, and the minimizer increases with x. Hence $xV'(x)-V(x)$, which vanishes at zero and has derivative $xV''(x)$, first increases and then decreases, or only

decreases, and the minimum of $V(x)/x$ on $(0,\zeta]$ is attained at an endpoint, with left limit $2\beta\ell$. If $\zeta \geq x_*$ the right endpoint value is I_ρ/ζ; otherwise use the endpoints of $(0, x_*]$. Both exceed $\beta\ell$, because $\beta(1-\rho^{1/(2\beta)}) \leq (1-\rho)/2 < I_\rho/\ell$: the first inequality follows by differentiating $b \mapsto b(1-e^{-\ell/(2b)})$, the second is $\tanh(\ell/2) < \ell/2$. Since $s > 1/\ell$ and $s > \beta\zeta/I_\rho$, we get $sV(x)/x \geq \beta + c_1$ on $[x_1, \min\{\zeta+\epsilon, 1\}]$, and small x is handled by $g \leq 0$ as before. For $\zeta \leq 1-\rho$, concavity of R_ρ and $R_\rho(1-\rho) = I_\rho$ give $\beta\zeta/R_\rho(\zeta) \leq \beta(1-\rho)/I_\rho < 1/\ell$, the inequality used after Theorem 38.

In all cases the uniform negative exponent on $y \geq y_0$ is of order $n\log n$, and the error terms are $o(n\log n)$. Summing over bins of y of width $(\log n)^{-2}$, with the upper endpoint for the norm and the lower one for the probability, keeps the margin, since I_ρ is Lipschitz on $[y_0, 1]$. The three ranges together give $\|f_t^{\mathrm{proj}} - 1\|_{L^\alpha(V_n)} \to 0$, which proves the upper assertions at every finite order.

The Poisson profile at the total variation window follows from the same decomposition with $s = 1/\ell$ and $\Theta \to \theta$. On the first range Θ' stays bounded; the non-retained schedules there have probability $o(1)$ and bounded norm, and the retained schedules give the projected tilt $e^{-\zeta\Theta}(1+\Theta)^{F_k}$ in $L^\alpha(V_n)$, by the tilt computation above and the bounded integer moments. The second range is unchanged. On the third, the hypothesis $\beta\zeta/R_\rho(\zeta) < 1/\ell$ (for $\alpha \geq 2$), or the endpoint bound $V(x)/x > \beta\ell$ together with $\beta\zeta/I_\rho < 1/\ell$ when $\zeta \geq x_*$ (for $\alpha < 2$), gives the strict margin at $s = 1/\ell$.

Upper bound at order infinity. Here $\|f_t^{\mathrm{proj}} - 1\|_\infty \leq \sum_{\lambda \neq (n)} d_\lambda a_\lambda^t$, a sum with nonnegative terms. The eigenvalue bound $\log a_\lambda \leq -nR_{\rho_n}(j/n) + O(\sqrt{n}\log n)$ of Section 65.13 and the quotient dimension bound make the levels $j \geq \epsilon n$ vanish when $t/\log n$ stays strictly above $\zeta/R_\rho(\zeta)$, because $R_\rho(x)/x$ is nonincreasing. For $1 \leq j < \epsilon n$ the level sum is at most $(n^2\rho_n^t)^j \exp(O(tj^2/n + tj\log n/n))/j!$, as in Section 65.13, which tends to zero in sum when $n^2\rho_n^t \to 0$.

Lower bounds. Condition on retained step sizes. The number of tracked positions untouched by all steps has mean $k_n \prod_s (J_s/n)$, which diverges for $s < 1/\ell$, and its pairwise untouched probabilities are at most products of marginals, so it diverges in probability. The observed fixed-point count dominates it, while its uniform mean is $k_n/n \leq 1$. Binary data processing for relative entropy and monotonicity in the order give divergence below $1/\ell$ at every order above one. At infinity the $(n-1,1)$ projector contributes $(n-1)\min(k_n, n-1)a_{(n-1,1)}^t$ at x_0, with $a_{(n-1,1)} = \rho_n + O(1/n)$, which diverges for $s < 2/\ell$.

In one step the probability of fixing every tracked position is exactly $P_m(b_n)/P_n(b_n)$: the permutations fixing the tracked positions are those supported in the untracked ones, with weights $b_n^{|\operatorname{supp}\pi|}$. Hence x_0 carries mass at least $(P_m/P_n)^t$. The truncated exponential formula $P_m = m! b_n^m \sum_{j \leq m} \eta_n^j/j!$ and Stirling give $\log(P_n/P_m) = nR_\rho(\zeta) + o(n)$ and $\log(n)_{k_n} = \zeta n\log n + o(n\log n)$. The atom has $L^\alpha(V_n)$ norm at least $(n)_{k_n}^\beta (P_m/P_n)^t$, with $\beta = 1$ at infinity, which diverges for $s < \beta\zeta/R_\rho(\zeta)$. This completes Theorem 38.

The saturated profile. We prove Theorem 39. Let $p_H = P_m/P_n$. Since $mb_n \to (1-\zeta)/\rho < 1$, $P_m = O(1)$. Call a tracked index changed in a step if its image changes. By centrality, conditional on the current tuple and the number r of changes, the set of changed indices is uniform, and the law of r does not depend on the tuple. Specifying the r changed tracked labels and summing over j moved untracked labels, the relative mass of $r \geq 1$ changes is at most

$$\frac{1}{p_H}\Pr(r) \leq \binom{k_n}{r} \sum_{j=0}^{m} \binom{m}{j} (r+j)!\, b_n^{r+j} \leq (k_n b_n)^r (1-mb_n)^{-r-1} \leq C^r,$$

using $P_m \geq 1$ and $\sum_{j \geq 0} \binom{r+j}{r} z^j = (1-z)^{-r-1}$.

For a total M of tracked changes with $0 < M \le \delta n$, the repeated-position bound (70), with k_n in place of n, applies. Off the event of more than $M/4$ repeats, at least $M/2$ tracked indices change exactly once and still differ at the end, so the final number r' of changed indices satisfies $M/2 \le r' \le M$. Simultaneous relabeling by S_{k_n} preserves the law and acts transitively on changed-index sets of size r', so each tuple lies in an orbit of size at least $\binom{k_n}{r'}$ and the counting L^α norm is at most $(M/k_n)^{\beta M/2}$. On the repeat event use its mass $2^M (M/(k_n - M))^{M/4}$. For $M \le \sqrt{n}$ both factors beat $C^M t^M$; for $\sqrt{n} < M \le \delta n$ a small δ beats C^M exponentially, over at most $\exp(O(\log^2 n))$ size sequences. After multiplication by bounded $Q_{\alpha,k}$, all paths of total tracked activity in $(0, \delta n]$ vanish in $L^\alpha(V_n)$. The paths with no tracked change have mass exactly p_H^t and end at x_0, which gives the atom $(n)_{k_n} p_H^t \mathbf{1}_{\{x_0\}}$. Conditional on this event the moved untracked labels have uniformly bounded mean, since $mb_n < 1 - \epsilon$, so the retained macroscopic schedules capture only $o(p_H^t)$ of it.

For paths with more than δn tracked changes, the total selected size $M_K = \sum_i K_i$ exceeds δn, and we bound their full nonnegative schedule density. Arithmetic retained fractions $y \le y_0$ are negligible as above. Along a subsequence let $t/\log n \to s$; bounded Θ and $Q_{\alpha,k}$ give $s \ge 1/\ell$ and $s \ge \beta\zeta/I_\rho$. If the second inequality is strict, the high-schedule estimates have a strict margin; for $\alpha \ge 2$ they use $\kappa = 1/\beta$ and $\min_y (I_\rho(y) - g(x,y)) = R_\rho(x)$, with minimizer $y = \min(\rho + x, 1)$. In the case $s = \beta\zeta/I_\rho$, put $\kappa = 2$ for $\alpha \le 2$ and $\kappa = 1/\beta$ for $\alpha \ge 2$. The exponent $\beta x + s\{\beta\kappa g(x,y) - I_\rho(y)\}$ is strictly negative on $y \ge y_0$ except at $x = \zeta, y = 1$: below order two by the strict endpoint comparison for $V(x)/x$, since its left value $2\beta\ell$ exceeds β/s and $\zeta > x_*$; from order two on by concavity of R_ρ, its positive initial slope and its constant branch. In both regimes $\ell + \beta\kappa \log(1 - \zeta) < 0$, with $1 - \zeta$ replaced by a fixed $1 - \zeta + \epsilon$ when $\zeta = 1$; below order two this follows from $\zeta > x_*$, and from order two on from $\zeta > 1 - \rho$.

Let $\Psi_n(y)$ be the maximum over $x \le \min(y, k_n/n)$ of the finite-n exponent. For y close to one every maximizer has $x \ge \zeta - \epsilon$, and there

$$\partial_y \ge \beta\kappa \log \frac{1}{1 - \zeta + \epsilon} - \log(1/\rho_n) > 0,$$

so $\Psi_n(y) \le \Psi_n(1) - c(1 - y)$, while bounded $Q_{\alpha,k}$ gives $\Psi_n(1) \le O(1/\log n)$. The rough bounds therefore suffice for $y \le 1 - C/\log n$. For the remaining bins write $u = 1 - y = M_K/(nt)$. Levels $x < \zeta - \epsilon$ lose dimension on the $n \log n$ scale. On the others,

$$\frac{(n - K_i)_j}{(n)_j} = \frac{(n - j)_{K_i}}{(n)_{K_i}} \le (1 - j/n)^{K_i} \le (1 - \zeta + \epsilon)^{K_i},$$

and summing quotient dimensions by their total $(n)_{k_n}$, the logarithm of norm times probability is at most

$$\beta \log(n)_{k_n} - nt I_{\rho_n} + M_K \big(\log(1/\rho_n) + \beta\kappa \log(1 - \zeta + \epsilon)\big) + o(n).$$

The first two terms are $\log Q_{\alpha,k} + O(\log^2 n)$ since P_m is bounded, and the coefficient of $M_K > \delta n$ is strictly negative, so this contribution vanishes exponentially. Bins of width $(\log n)^{-2}$ change the exponent by $o(n)$, as I_{ρ_n} is uniformly Lipschitz on $[y_0, 1]$. The retained density and the all-stay atom give the asserted approximation; the projected tilt has vanishing L^α mass at x_0, since an exponential in k_n is beaten by $(n)_{k_n}^{1/\alpha}$, so the power integral splits. For $1 < \alpha \le 2$ and $\zeta \le 1 - \rho$, the strict inequality $\beta\zeta/R_\rho(\zeta) < 1/\ell$ and the strict order-two upper bound above give the same formulas with $Q_{\alpha,k} \to 0$.

The projected uniform window. We prove Theorem 40. If $\zeta \le 1 - \rho$, concavity gives $R_\rho(\zeta)/\zeta \ge I_\rho/(1 - \rho) > \ell/2$, so the high quotient levels vanish by the strict margin at $t/\log n \ge 2/\ell$, the low

levels give the uniform tilt, and $Q_{\infty,k} \to 0$. On the saturated branch use the path decomposition above with $\beta = \kappa = 1$. For small or moderate arithmetic retained fractions the characterwise comparison of Section 65.11 gives the supremum bound $(1+\Theta')^n$, $\Theta' = ny^t e^{o(1)}$; choosing $y_0 > \rho$ with $2 + s \log y_0 < 1$, possible since $s \geq 2/\ell$, this is subexponential on the intermediate tail, and $W = O(1)$ bounds the norm near ρ_n. The sparse tracking bound holds in counting L^∞, and the large-activity estimate is the same with $\kappa = 1$. Projector kernels are nonnegative on the diagonal and bounded by it, so the peak is at x_0, and the uniform approximation is at least $1 - o(1)$ everywhere, so separation tends to zero.

65.16 Observation through colour tables

The partial observation of Section 65.3 records the cards at fixed positions. Here only colour counts are recorded. Partition the n positions into d blocks $B_1, \dots, B_{\mathsf{d}}$ with $|B_a| = n\mathsf{q}_{a,n}$, $\mathsf{q}_{a,n} \to \mathsf{q}_a > 0$, and colour the positions with r colours of fixed counts $n\mathsf{p}_{b,n}$, $\mathsf{p}_{b,n} \to \mathsf{p}_b > 0$, where $\mathsf{d}, \mathsf{r} \geq 2$ are fixed. A colouring c is moved by a permutation σ to $c \circ \sigma^{-1}$, and we observe the table $\mathsf{N}_{ab} = \#\{i \in B_a : c(i) = b\}$. Let π_n be the law of the table of a uniformly random colouring, and let $f_{t,n}$ be the density of the observed table at time t relative to π_n, for a deterministic initial colouring c^0 whose table satisfies $\mathsf{N}^0/n \to \Upsilon$. Put $\mathcal{D} = \Upsilon - \mathsf{q}\mathsf{p}^{\mathsf{T}}$,

$$\mathcal{V} = \sum_{a,b} \frac{\mathcal{D}_{ab}^2}{\mathsf{q}_a \mathsf{p}_b}, \qquad \mathcal{I}(S) = \sum_{a,b} \frac{\mathcal{D}_{ab}\,(S_{ab} - \mathsf{q}_a \mathsf{p}_b)}{\mathsf{q}_a \mathsf{p}_b},$$

for tables S of fractions with margins q, p, and let $\mathcal{I}_\pm$ be the extrema of $\mathcal{I}$ on the limiting transportation polytope of such S. Since the shuffles below commute with permutations inside the blocks, the table is again a Markov chain.

We treat three shuffles. In the *deterministic schedule*, step $s \leq t$ chooses a uniform K_s-subset and permutes it uniformly, with $K_s \geq cn$ for a fixed $c > 0$, $J_s = n - K_s$, and $w_n = \prod_s (J_s/n)$. The *permanental shuffle* is the subset shuffle of Section 65.1 with $\rho_n = \eta_n/n \to \rho \in (0,1)$ and $w_n = \rho_n^t$. In both, $t = O(\log n)$. In the *transposition shuffle* a uniform transposition is applied at the times of a rate-one Poisson process, $t = O(n \log n)$, and $w_n = e^{-2t/n}$.

Theorem 41 (Gaussian window)**.** *If $\sqrt{n}\, w_n \to \varkappa \in [0, \infty)$, then $f_{t,n}$ under π_n converges in distribution and in every L^b, $1 \leq b < \infty$, to $\exp(\varkappa\sqrt{\mathcal{V}}\, G - \varkappa^2 \mathcal{V}/2)$ with $G \sim N(0,1)$; more precisely*

$$\big\| f_{t,n} - \exp\big(\varkappa\sqrt{n}\, \mathcal{I}(\mathsf{N}/n) - \varkappa^2 \mathcal{V}/2\big) \big\|_{L^b(\pi_n)} \to 0.$$

Consequently the Rényi divergence of every fixed order $\alpha \in (0, \infty)$ tends to $\alpha\varkappa^2\mathcal{V}/2$ (relative entropy at $\alpha = 1$), total variation tends to $2\Phi(\varkappa\sqrt{\mathcal{V}}/2) - 1$, and one minus the Hellinger affinity tends to $1 - e^{-\varkappa^2\mathcal{V}/8}$.

Theorem 42 (uniform window)**.** *Let $nw_n \to \upsilon \in [0, \infty)$. For the permanental and transposition shuffles, $f_{t,n}(\mathsf{N}) = e^{\upsilon \mathcal{I}(\mathsf{N}/n)} + o(1)$ uniformly over feasible tables. For a deterministic schedule, let U be the number of positions never selected and $\mathcal{H}_n(z) = \mathbb{E}(1+z)^{\mathsf{U}}$; then $f_{t,n}(\mathsf{N}) = \mathcal{H}_n(\mathcal{I}(\mathsf{N}/n)) + o(1)$ uniformly. In particular the order-infinity divergence is $\upsilon\mathcal{I}_+ + o(1)$ and the separation distance is $1 - e^{\upsilon\mathcal{I}_-} + o(1)$ for the permanental and transposition shuffles, and $\log \mathcal{H}_n(\mathcal{I}_+) + o(1)$ and $1 - \mathcal{H}_n(\mathcal{I}_-) + o(1)$ for a deterministic schedule. If $\min_s J_s \to \infty$ then $\mathcal{H}_n(z) \to e^{\upsilon z}$; along a subsequence with $\min_s J_s = L$ fixed, $\mathcal{H}_n(z) \to (1 + \upsilon z/L)^L$.*

Theorem 43 (untouched-set surrogate)**.** *For a deterministic schedule, and for independent steps of the permanental law, compare the shuffle with the surrogate that fixes the random set of positions*

never selected and permutes its complement uniformly. After averaging over the selected sets, the two observed densities differ by $o(1)$ in every $L^b(\pi_n)$, $b < \infty$, when $\sqrt{n}\, w_n = O(1)$, and uniformly when $n w_n = O(1)$.

Remark 44 (observation of the initial colour classes). If the blocks are the colour classes of c^0, then $\Upsilon = \operatorname{diag}(\mathsf{p})$, and the law of the colouring is invariant under permutations inside the blocks, because c^0 is and the shuffles are conjugation invariant; so the table carries the full likelihood of the colouring. Here $\mathcal{V} = \mathcal{I}_+ = \mathsf{r} - 1$ and $\mathcal{I}_- = -\min\{1, (1 - \mathsf{p}_{\max})/\mathsf{p}_{\max}\}$. For the transposition shuffle the two windows are centred at $(n/4)\log n$ and $(n/2)\log n$ with windows of order n; for the permanental shuffle at $\log n/(2\log(1/\rho_n))$ and $\log n/\log(1/\rho_n)$ with bounded windows. The finite-n values for two colours in two equal blocks, computed exactly by uniformization in `verify_color_tables.py`, approach the limits: at $n = 6400$ the total variation is 0.3825 and 0.6818 for $\varkappa = 1, 2$ (limits 0.3829, 0.6827).

Diaconis and Shahshahani proved cutoff for the Bernoulli–Laplace diffusion model [50]; in the time scale used here, with n positions in two equal blocks, it occurs at $(n/4)\log n$. For random transpositions on repeated cards Shen obtained the Gaussian total variation profile in the window $(n/2)(\log n - \frac{1}{2}\log l + c)$ [51]; in the equal-population case of Remark 44 this corresponds to $\varkappa = e^{-c}/\sqrt{\mathsf{r}}$. Limit profiles for the two-urn model and for balanced urn exchanges are in [52, 53], and limit profiles for projections of random walks on groups in [54]. Bristiel and Caputo give the gap and entropy constants of the subset update [37, Theorems 1.8–1.9]. Theorems 41–43 add arbitrary initial tables and unequal margins, block observation, every finite L^b norm, the uniform window, and subset updates.

65.17 Proofs for colour tables

Let X be the set of colourings with the given colour counts, with uniform measure, and $L^2(X) = \bigoplus_\lambda W_\lambda$ its isotypic decomposition under $\mathfrak{S}_n$; write $W_{(j)} = \bigoplus_{n-\lambda_1=j} W_\lambda$. A central random permutation with Fourier eigenvalues a_λ acts on W_λ as multiplication by a_λ. Let P_{N} be conditional expectation onto functions of the table; it is the average over the block subgroup $\mathfrak{S}_{B_1} \times \cdots \times \mathfrak{S}_{B_{\mathsf{d}}}$, whose orbits are the tables, so it commutes with every isotypic projection. With g_0 the density of the point mass at c^0,

$$f_{t,n} = \sum_{j \ge 0} \sum_{n-\lambda_1=j} a_\lambda^t \, P_{\mathsf{N}} \Pi_\lambda g_0, \tag{74}$$

where Π_λ is the isotypic projection, and $\|\Pi_{(j)} g_0\|_2^2 = |\Pi_{(j)} g_0(c^0)| = \dim W_{(j)} \le \binom{n}{j} \mathsf{r}^j$, since the reproducing kernel of an invariant subspace is constant on the diagonal by transitivity and $W_{(j)}$ lies in the span of functions of j coordinates (Lemma 45). The same bound holds pointwise: $|\Pi_{(j)} g_0| \le \dim W_{(j)}$ by Cauchy–Schwarz, and P_{N} preserves it.

Lemma 45 (polynomial layers)**.** *Let $\mathcal{U}_j \subset L^2(X)$ be the span of functions of at most j coordinates, and $\mathcal{P}_j$ the space of functions of the table that are polynomials of degree at most j in its entries. Then $\mathcal{U}_j = \bigoplus_{i \le j} W_{(i)}$. If every block has at least j positions, then $P_{\mathsf{N}} \mathcal{U}_j = \mathcal{P}_j$, and $P_{\mathsf{N}} W_{(j)}$ is the π_n-orthogonal complement of $\mathcal{P}_{j-1}$ in $\mathcal{P}_j$.*

Proof. For a set A of positions, the subgroup $\mathfrak{S}_{A^c}$ acts transitively on the colourings that agree with a given colouring on A, so the $\mathfrak{S}_{A^c}$-invariant functions are exactly the functions of $c|_A$. By the branching rule, the restriction of the irreducible representation V_λ to $\mathfrak{S}_{n-|A|} \times \mathfrak{S}_{|A|}$ contains the trivial representation of $\mathfrak{S}_{n-|A|}$ if and only if $\lambda_1 \ge n - |A|$. Hence $\mathcal{U}_j$, the sum of the $\mathfrak{S}_{A^c}$-invariants over $|A| \le j$, lies in $\bigoplus_{i \le j} W_{(i)}$. Conversely, write $W_\lambda \cong V_\lambda \otimes \mathbb{C}^{m_\lambda}$ with $\mathfrak{S}_n$ acting on the first

factor; for $n-\lambda_1 \le j$ and $|A| = j$ the invariants are $V_\lambda^{\mathfrak{S}_{A^c}} \otimes \mathbb{C}^{m_\lambda} \neq 0$, and their $\mathfrak{S}_n$-translates span $V_\lambda \otimes \mathbb{C}^{m_\lambda}$ because V_λ is irreducible. Since $\mathcal{U}_j$ is $\mathfrak{S}_n$-invariant, it contains W_λ.

The functions of $c|_A$ are spanned by products $\prod_{i\in A} \mathbf{1}[c(i) = b_i]$. Averaging over the block subgroup, the positions of $A \cap B_a$ receive a uniformly random injective assignment of positions of B_a, so the average is $\prod_a \prod_b (\mathsf{N}_{ab})_{k_{ab}} / \prod_a (|B_a|)_{k_a}$, where k_{ab} counts $i \in A \cap B_a$ with $b_i = b$, $k_a = \sum_b k_{ab}$, and $(x)_k$ is the falling factorial. When every $|B_a| \ge j$, every exponent array with $\sum k_{ab} \le j$ occurs. The falling-factorial monomials of total degree at most j are a triangular basis of the polynomials of degree at most j, so $P_{\mathsf{N}}\mathcal{U}_j = \mathcal{P}_j$. Finally $P_{\mathsf{N}} W_{(i)} \subset W_{(i)}$, and these spaces are mutually orthogonal, so $\mathcal{P}_j = \bigoplus_{i\le j} P_{\mathsf{N}} W_{(i)}$ is an orthogonal decomposition, which gives the last claim. The script `verify_color_tables.py` checks, in exact arithmetic on a multislice with $n = 7$, that the transposition operator preserves each $\mathcal{P}_j$ and acts on the layers by the content eigenvalues of level j. □

By the first part of the lemma, $\dim W_{(j)} \le \dim \mathcal{U}_j \le \binom{n}{j} \mathsf{r}^j \le (\mathsf{r}n)^j/j!$.

Lemma 46 (multipliers)**.** *Fix j. In each of the three shuffles there is a common multiplier $m_{j,n}$ with $a_\lambda^t = m_{j,n}(1+o(1))$ uniformly over shapes of level j, and $m_{j,n} = (1+o(1))w_n^j$ for the permanental and transposition shuffles whenever $t = O(\log n)$, respectively $t = O(n \log n)$, and for the deterministic schedule when $\sqrt{n}\, w_n \to \varkappa > 0$. For the deterministic schedule $m_{j,n} = \prod_s (J_s)_j/(n)_j$. For all levels and both windows, $\sum_{j>M} \sup_{n-\lambda_1=j} |a_\lambda|^{2t} (\mathsf{r}n)^j/j! \to 0$ as $M \to \infty$, uniformly in large n, when $\sqrt{n}\, w_n = O(1)$; and $\sum_{j>M} \sup_{n-\lambda_1=j} |a_\lambda|^t (\mathsf{r}n)^j/j! \to 0$ in the same sense when $nw_n = O(1)$.*

Proof. Subset shuffles. For a single selected size $k \le \lambda_1$ the eigenvalue is $f^{\lambda/(k)}/f^\lambda \le h_\lambda (n-k)_j/(n)_j$, with equality for fixed j and $k \ge j$, and $h_\lambda = 1 + O_j(1/n)$; it vanishes for $k > \lambda_1$. Both statements and the bound $\log h_\lambda \le Cj \log n/n$ for $\lambda_1 \ge cn$ are in Sections 65.5–65.6. For the deterministic schedule the eigenvalue at level j is therefore $h_\lambda^t \prod_s (J_s)_j/(n)_j$ with $h_\lambda^t = 1 + O_j(t/n)$, and is at most $w_n^j e^{Ctj \log n/n}$ on every level. When $\sqrt{n}\, w_n \to \varkappa > 0$, $\min_s J_s \ge nw_n \asymp \sqrt{n}$, so $\prod_s (J_s)_j/J_s^j = 1 - O(j^2 t/\sqrt{n})$ and $n^j/(n)_j \to 1$, giving $m_{j,n} = (1+o(1))w_n^j$. For the permanental law, the exact factorial moment of Section 65.9 gives $a_\lambda = \rho_n^j(1 + O_j(1/n))$ at fixed j, the bound $a_\lambda \le \rho_n^j \exp(Cj(\log n + j)/n)$ for $j \le n/(\log n)^2$, and, after conditioning on $J/n \le \rho + \epsilon < 1$, the bound $a_\lambda \le 2e^{-c_1 j}$ for larger j; with $t = O(\log n)$ and $t \ge c_2 \log n$ in both windows, these give the tail bounds, since $(\mathsf{r}n)^j/j! \le e^{O(j \log\log n)}$ for $j > n/(\log n)^2$.

Transpositions. Here $a_\lambda^t = e^{-t\delta_\lambda}$ with $\delta_\lambda = 1 - \sum_{\square\in\lambda} c(\square)/\binom{n}{2}$, and Lemma 47 gives $\delta_\lambda = 2j/n + O_j(n^{-2})$, hence $a_\lambda^t = w_n^j(1 + O(\log n/n))$ at fixed j. For $j \le n/(\log n)^2$ the lower bound $\delta_\lambda \ge 2j(n-j+1)/(n(n-1))$ gives $a_\lambda^t \le w_n^j e^{O(j)}$, and for larger j the bound $\delta_\lambda \ge j/(n-1)$ gives $a_\lambda^t \le n^{-j/5}$ in both windows. With the dimension bound these give the tails. □

Lemma 47 (content)**.** *For a partition λ of n with $\lambda_1 = n-j$ and $\nu = (\lambda_2, \lambda_3, \dots)$, the transposition deficit is $\delta_\lambda = [j(2n-j+1) - 2\,\mathrm{cont}(\nu)]/(n(n-1))$, where* cont *is the sum of contents, and*

$$\frac{2j(n-j+1)}{n(n-1)} \le \delta_\lambda \le \frac{4j}{n-1}, \qquad \delta_\lambda \ge \frac{j}{n-1}.$$

Proof. The normalized character of a transposition is $\sum_\square c(\square)/\binom{n}{2}$. The first row contributes $\binom{n-j}{2}$, and a box of ν in row i and column k of ν has content $k-i-1$ in λ, which gives the formula. The content of ν is at most $\binom{j}{2}$, attained by a single row, which gives the first lower bound, and at least $-j(j-1)/2$, which gives $\delta_\lambda \le (2nj+j)/(n(n-1)) \le 4j/(n-1)$. Each row i of λ contributes at most $\binom{\lambda_i}{2}$, so the total content is at most $n(\lambda_1 - 1)/2$, which gives $\delta_\lambda \ge j/(n-1)$. The script `verify_color_tables.py` checks all three bounds for every partition of $n \le 40$. □

Lemma 48 (the null table)**.** *Under π_n, the free entries $(\mathsf{N}_{ab})_{a<\mathsf{d},\,b<\mathsf{r}}$, centred at $n\mathsf{q}_{a,n}\mathsf{p}_{b,n}$ and divided by $\sqrt{n}$, converge with all moments, and with uniformly sub-Gaussian tails, to a centred Gaussian vector Z of covariance $\Sigma = (\operatorname{diag}\mathsf{q} - \mathsf{q}\mathsf{q}^{\mathsf{T}})_{<\mathsf{d}} \otimes (\operatorname{diag}\mathsf{p} - \mathsf{p}\mathsf{p}^{\mathsf{T}})_{<\mathsf{r}}$. For every array Y with zero margins, $Y_{\mathrm{free}}^{\mathsf{T}}\Sigma^{-1}Y_{\mathrm{free}} = \sum_{a,b} Y_{ab}^2/(\mathsf{q}_a\mathsf{p}_b)$.*

Proof. The table of a uniform colouring has the multivariate hypergeometric law

$$\pi_n(\mathsf{N}) = \frac{\prod_a |B_a|! \prod_b (n\mathsf{p}_{b,n})!}{n! \prod_{a,b} \mathsf{N}_{ab}!}.$$

Each entry is hypergeometric, so Hoeffding's comparison of sampling without and with replacement [55] gives $\pi_n(|\mathsf{N}_{ab} - n\mathsf{q}_{a,n}\mathsf{p}_{b,n}| > u\sqrt{n}) \le 2e^{-2u^2}$. Stirling's formula, uniformly on bounded sets of the scaled free entries, gives the local limit with exponent $-\frac{1}{2}\sum_{a,b} Y_{ab}^2/(\mathsf{q}_a\mathsf{p}_b)$, where Y is the scaled table completed by its zero margins, and the free entries form a lattice of mesh $n^{-1/2}$. Riemann sums and the tail bound give convergence of all moments. The inverse of $(\operatorname{diag}\mathsf{q} - \mathsf{q}\mathsf{q}^{\mathsf{T}})_{<\mathsf{d}}$ is $\operatorname{diag}(1/\mathsf{q}_a)_{a<\mathsf{d}} + \mathbf{1}\mathbf{1}^{\mathsf{T}}/\mathsf{q}_{\mathsf{d}}$, and similarly for p; expanding the Kronecker product with $Y_{a\mathsf{r}} = -\sum_{b<\mathsf{r}} Y_{ab}$ and $Y_{\mathsf{d}b} = -\sum_{a<\mathsf{d}} Y_{ab}$ gives the quadratic form. □

Lemma 49 (hypercontractive layers)**.** *For every $b \in (2,\infty)$ there is C_b, depending only on b, r and $\min_b \mathsf{p}_b$, such that $\|h\|_{L^b(X)} \le e^{C_b j}\|h\|_{L^2(X)}$ for every $h \in \bigoplus_{i\le j} W_{(i)}$; in particular for every function of the table in $\mathcal{P}_j$.*

Proof. Let Q average over a uniform transposition, so $\mathcal{E}(f,f) = \langle (I-Q)f, f\rangle = \frac{1}{2}\mathbb{E}[(f(c)-f(c\circ\tau))^2]$. Filmus, O'Donnell and Wu prove $\mathbb{E}[(f(c)-f(c\circ\tau))^2] \ge \varrho_n \operatorname{Ent}(f^2)$ with $\varrho_n^{-1} \le n\sum_b \frac{1}{2}\log_2(4/\mathsf{p}_{b,n})$ [56, Theorem 1]. By Gross's theorem for finite Markov chains [57], the semigroup $P_s = e^{-s(I-Q)}$ satisfies $\|P_s f\|_b \le \|f\|_2$ for $s = \log(b-1)/(2\varrho_n) \le C_b' n$. On $W_{(i)}$, $i \le j$, the operator $I-Q$ acts by $\delta_\lambda \le 4j/(n-1)$ (Lemma 47). Hence $h = P_s\tilde{h}$ with $\tilde{h} = \sum_\lambda e^{s\delta_\lambda}\Pi_\lambda h$ and $\|\tilde{h}\|_2 \le e^{4C_b' nj/(n-1)}\|h\|_2$, so $\|h\|_b \le \|\tilde{h}\|_2 \le e^{8C_b' j}\|h\|_2$. For real h this is the claim; complex h follow by splitting into real and imaginary parts, both in the same space. A function of the table has the same L^b norms on X and under π_n. □

Proof of Theorem 41. By Lemma 45 and (74), the level-j part of $f_{t,n}$ is $\sum_\lambda a_\lambda^t P_{\mathsf{N}}\Pi_\lambda g_0$. Write it as $m_{j,n}\mathcal{K}_j + R_j$, where $\mathcal{K}_j = P_{\mathsf{N}}\Pi_{(j)}g_0$ is the reproducing kernel at N^0 of the layer $\mathcal{P}_j \ominus \mathcal{P}_{j-1}$ in $L^2(\pi_n)$, because $P_{\mathsf{N}}g_0$ is the density of the point mass at N^0. By Lemma 46, $\|R_j\|_2 \le o(1)\, w_n^j \dim W_{(j)}^{1/2} = o(1)$ at fixed j, and by Lemma 49 the same holds in L^b. The tails beyond a fixed level M are bounded in L^b by $\sum_{j>M} e^{C_b j}\sup|a_\lambda|^t((\mathsf{r}n)^j/j!)^{1/2}$, which tends to zero as $M\to\infty$ uniformly in n by Lemma 46, the factor $e^{C_b j}$ being absorbed by the factorial.

Let $\mathsf{w} = \Sigma^{-1/2}\mathsf{Z}_n$ be the whitened scaled table. Gram–Schmidt applied to the monomials in w in graded lexicographic order gives an orthonormal basis $\{\mathsf{P}_{\mathbf{k},n}\}$ of each $\mathcal{P}_j$ under π_n, the multi-indices of degree j spanning the layer. By Lemma 48 the Gram matrices of the monomials converge to their Gaussian values, which are positive definite, so the coefficients converge to those of the normalized Hermite polynomials $\mathrm{He}_{\mathbf{k}}/\sqrt{\mathbf{k}!}$. At the initial table $\mathsf{w} = \sqrt{n}(\mathsf{v} + o(1))$ with $\mathsf{v} = \Sigma^{-1/2}\mathcal{D}_{\mathrm{free}}$, so $n^{-j/2}\mathsf{P}_{\mathbf{k},n}(\mathsf{N}^0) \to \mathsf{v}^{\mathbf{k}}/\sqrt{\mathbf{k}!}$ and

$$m_{j,n}\mathcal{K}_j = \sum_{|\mathbf{k}|=j} m_{j,n}\mathsf{P}_{\mathbf{k},n}(\mathsf{N}^0)\mathsf{P}_{\mathbf{k},n} = \sum_{|\mathbf{k}|=j} \frac{\varkappa^j \mathsf{v}^{\mathbf{k}}}{\mathbf{k}!}\mathrm{He}_{\mathbf{k}}(\mathsf{w}) + o(1)$$

in every $L^b(\pi_n)$, using the uniform sub-Gaussian tails of Lemma 48. The generating function of the Hermite polynomials gives $\sum_{\mathbf{k}} \varkappa^{|\mathbf{k}|}\mathsf{v}^{\mathbf{k}}\mathrm{He}_{\mathbf{k}}(\mathsf{w})/\mathbf{k}! = \exp(\varkappa\mathsf{v}^{\mathsf{T}}\mathsf{w} - \varkappa^2|\mathsf{v}|^2/2)$, with L^b-convergent series

for Gaussian w by Gaussian hypercontractivity. Lemma 48 gives $|\mathsf{v}|^2 = \mathcal{V}$ and $\mathsf{v}^{\mathsf{T}}\mathsf{w} = \mathcal{D}_{\mathrm{free}}^{\mathsf{T}}\Sigma^{-1}\mathsf{Z}_n = \sqrt{n}\,\mathcal{I}(\mathsf{N}/n)$. Fixing a truncation, letting $n \to \infty$ with the uniform exponential moments of w, and then removing the truncation proves the displayed approximation and the limit in law.

The divergence limits follow by integrating against π_n: for $\alpha > 1$ from the L^α approximation, which gives $\mathbb{E}f^\alpha \to e^{\alpha(\alpha-1)\varkappa^2\mathcal{V}/2}$; for $\alpha < 1$, total variation and Hellinger from convergence in law and uniform integrability, since $f^\alpha \le 1 + f$ and f is bounded in L^2; and for relative entropy from $x \log x = O(x^2)$ at infinity. The limiting values are those of the Gaussian shift $N(\varkappa\sqrt{\mathcal{V}}, 1)$ against $N(0,1)$. □

Proof of Theorem 42. At $nw_n = O(1)$ use the pointwise bound $|P_{\mathsf{N}}\Pi_{(j)}g_0| \le \dim W_{(j)}$ and the second tail bound of Lemma 46; the levels beyond M are uniformly small. At fixed j, the coefficient convergence above gives $n^{-j}\mathcal{K}_j(\mathsf{N}) = (\mathsf{v}^{\mathsf{T}}\Sigma^{-1/2}(S - \mathsf{q}\mathsf{p}^{\mathsf{T}})_{\mathrm{free}})^j/j! + o(1) = \mathcal{I}(S)^j/j! + o(1)$ uniformly over feasible $S = \mathsf{N}/n$, by the multinomial theorem. For the permanental and transposition shuffles $n^j m_{j,n} \to \upsilon^j$, which sums to $e^{\upsilon\mathcal{I}(S)}$. For a deterministic schedule $n^j m_{j,n} = (n^j/(n)_j)\mathbb{E}[(\mathsf{U})_j]$, because $\mathbb{E}[(\mathsf{U})_j] = \prod_s (J_s)_j/(n)_j \cdot (n)_j$ by independence of the steps, and the series $\sum_j \mathbb{E}[(\mathsf{U})_j]z^j/j!$ is $\mathcal{H}_n(z)$; the factorial bound $\mathbb{E}(\mathsf{U})_j \le (nw_n)^j$ controls the tails. If $\min_s J_s \to \infty$, all factorial ratios may be replaced by powers; along a subsequence with smallest complement L, the other complements are at least $n\sqrt{w_n}$ and the limiting factorial moments are $(L)_j(\upsilon/L)^j$, which give $(1 + \upsilon z/L)^L$.

For the extremes, $1 + \mathcal{I}(S) = \sum_{a,b} \Upsilon_{ab}S_{ab}/(\mathsf{q}_a\mathsf{p}_b) \ge 0$, so $\mathcal{H}_n$ and $e^{\upsilon z}$ are increasing on the range of $\mathcal{I}$; every limiting table is approximated by normalized feasible tables (round a limiting coupling by an integral flow), and $\mathcal{I}$ is continuous, so the extreme values of $f_{t,n}$ converge to the stated quantities, with $\mathcal{I}_+ \ge 0 \ge \mathcal{I}_-$ because the product table gives zero. For the colour classes, $\mathcal{I}(S) = \sum_a S_{aa}/\mathsf{p}_a - 1 \le \mathsf{r} - 1$, and $S_{aa} \ge (2\mathsf{p}_a - 1)_+$ by the margins; the two diagonal bounds are attained by the tables $\operatorname{diag}\mathsf{p}$ and by the table of the pair $(U, U + \frac{1}{2})$ for U uniform on a circle divided into arcs of lengths p_a, which give $\mathcal{I}_+ = \mathsf{r} - 1$ and $\mathcal{I}_- = -\min\{1, (1 - \mathsf{p}_{\max})/\mathsf{p}_{\max}\}$. □

Proof of Theorem 43. For a deterministic schedule the eigenvalues at level j of the surrogate and of the shuffle are

$$h_\lambda \frac{\mathbb{E}[(\mathsf{U})_j]}{(n)_j} = h_\lambda \prod_s \frac{(J_s)_j}{(n)_j} \qquad \text{and} \qquad h_\lambda^t \prod_s \frac{(J_s)_j}{(n)_j}$$

(Section 65.8); they differ by the factor $h_\lambda^{t-1} = 1 + O_j(t/n)$ at fixed j, and both satisfy the tail bounds of Lemma 46, the surrogate with a single hook factor. For independent permanental steps the factorial moment of the untouched set is m_j^t with $m_j = \mathbb{E}[(J)_j]/(n)_j$; at fixed j both multipliers equal $(h_\lambda m_j + O_j(e^{-cn}))^t$ up to the single-hook factor, since a failure of the branching formula requires a selected block smaller than j, and at high levels $m_j \le (\rho+\epsilon)^j + e^{-c_\epsilon n}$. Inserting these into (74) and using Lemma 49, respectively the pointwise kernel bound, proves both comparisons. □

66 The permanental cofactor ratio in order five

Throughout this section $\Lambda(A) = \lambda_{\max}(C(A))/\operatorname{per} A$ for a correlation matrix A, as in Section 36, and γ_n is its maximum in order n. For correlation matrices $C(A)$ is Hermitian positive semidefinite and $\operatorname{per} A > 0$, by Lemma 69. The ratio is unchanged under simultaneous permutation of rows and columns, under entrywise complex conjugation, and under conjugation $A \mapsto DAD^*$ by a diagonal unitary matrix D.

66.1 The question and the results

Theorem 28 determines γ_4 with its maximizers, and Theorem 12 does the same in every order beyond an absolute threshold, the maximizer being the optimal positive equicorrelation matrix up to diagonal unitary conjugation in both cases. Order five is the next order after Theorem 28; Section 6 records numerical ascent there that ends at the equicorrelation value. Comparing the two theorems leaves the following question.

Question 50. *Is $\gamma_5 = \max_{0<b\le 1} F_5(b)$, where $F_5(b) = \Lambda((1-b)I + bJ)$, with equality exactly on the diagonal unitary orbit of the optimal positive equicorrelation matrix of order five?*

The ratio Λ differs from the entry-weighted matrix $[a_{ij}\operatorname{per} A(i \mid j)]$ of the Bapat–Sunder permax conjecture (Conjecture 3 of [58], as recorded in [38, Section 7]), which compared the largest eigenvalue of that matrix with $\operatorname{per} A$ and was disproved by Drury [59]; see [2] for this and for its relation with the Bapat–Sunder Hadamard-product question, for which Drury gave a counterexample of order seven [29]. The raw ratio in order four is at most 2 by Lavi's isolated-row argument [3, Section 4, (4.1)].

Let $b_5 \in (0.2773470462, 0.2773470463)$ be the unique positive root of

$$R_5(b) = 583b^8 + 968b^7 + 824b^6 + 316b^5 + 170b^4 + 32b^3 - 8b^2 - 4b - 1, \tag{75}$$

and put $A_* = (1-b_5)I + b_5 J$ and $L_* = F_5(b_5)$; in the notation of Section 1, $A_* = B_5$. Lemma 57 shows that $L_* = \max_b F_5(b) = 1.8727665159\ldots$ and that L_* is the root near 1.87 of the irreducible polynomial

$$3125L^8 - 12500L^7 + 25000L^6 - 26500L^5 + 13450L^4 - 27644L^3 + 31904L^2 - 15788L + 2809. \tag{76}$$

Definition 51. A *monomial unitary* matrix is a product of a diagonal unitary matrix and a permutation matrix. A correlation matrix A of order n is *transitive* if the permutations underlying the monomial unitaries U with $UAU^* = A$ act transitively on $\{1, \ldots, n\}$.

Every correlation matrix invariant under a transitive group of simultaneous row and column permutations is transitive, and so is every diagonal unitary conjugate of such a matrix. The next theorem answers Question 50 for transitive matrices, including singular ones; it extends Theorem 27 from order four, and from ordinary permutation symmetry to monomial unitary symmetry. The same conclusion holds when the symmetries $U\overline{A}U^* = A$ are also admitted in Definition 51: the unitary symmetries then form a subgroup of index at most two, so the order of their permutation group is still divisible by five and it contains a five-cycle (Lemma 58).

Theorem 52 (transitive maximum)**.** *Every transitive correlation matrix A of order five satisfies $\Lambda(A) \le L_*$. Equality holds if and only if $A = DA_*D^*$ for a diagonal unitary matrix D.*

The second theorem is the local form of Question 50 over all complex correlation matrices, at an explicit radius. Let $\mathcal{O} = \{DA_*D^* : D \text{ diagonal unitary}\}$ and let $\operatorname{dist}(A, \mathcal{O})$ be the Frobenius distance from A to $\mathcal{O}$.

Theorem 53 (local rigidity in order five)**.** *If A is a correlation matrix of order five with $r = \operatorname{dist}(A, \mathcal{O}) \le 1/100$, then*

$$\Lambda(A) \le L_* - \tfrac{1}{25} r^2.$$

The second variation of Λ at A_* vanishes exactly on the tangent space of $\mathcal{O}$, and on its orthogonal complement its eigenvalues lie in $[-1.0771, -0.3404]$.

In particular every maximizer of Λ in order five lies in $\mathcal{O}$ or at Frobenius distance greater than $1/100$ from $\mathcal{O}$, and a correlation matrix with $\operatorname{dist}(A, \mathcal{O}) \le 1/100$ and $\Lambda(A) \ge L_* - \delta$ satisfies $\operatorname{dist}(A, \mathcal{O}) \le 5\sqrt{\delta}$.

The third theorem confines the unrestricted constant and describes every matrix that reaches the equicorrelation value.

Theorem 54 (order-five window)**.** *Every correlation matrix A of order five satisfies*

$$\Lambda(A) < \frac{1 + \sqrt{1 + 4\gamma_4}}{2} < 1.876674.$$

Consequently $L_* \le \gamma_5 < 1.876674$. If $\Lambda(A) \ge L_*$, then every unit vector e with $C(A)e = \lambda_{\max}(C(A))e$ satisfies $0.195795 < |e_i|^2 < 0.201051$ for all i, and every principal submatrix A_i of order four satisfies $\Lambda(A_i) > 1.5917$.

Theorem 54 places γ_5 in an interval of length less than 0.0040, which is 0.21% of L_*; the general upper bound $(n^2 + 2n + 3)/(4n)$ recalled in Section 1 equals $19/10$ in order five. Theorem 54 is the recursion of Lemma 71 with the exact seed γ_4, together with a strictness argument; the same recursion seeded with $\gamma_5 < 1.876674$ gives $\gamma_6 < 2.11117$ and $\gamma_7 < 2.34845$. A direct consequence of Theorem 52 is a sharp derivative bound, the order-five transitive analogue of Section 62.5.

Corollary 55. *If A is a transitive correlation matrix of order five and B is positive semidefinite, then*

$$\frac{d}{dt} \operatorname{per}(A + tB)\Big|_{t=0} \le L_* \, (\operatorname{per} A) \operatorname{tr} B,$$

with equality for $A = A_*$ and $B = J$.

Proof. Expansion of the permanent in each entry gives $\frac{d}{dt} \operatorname{per}(A + tB)|_{t=0} = \sum_{i,j} C(A)_{ij} B_{ij} = \operatorname{tr}(C(A)^{\mathsf{T}} B)$. The matrix $C(A)^{\mathsf{T}} = \overline{C(A)}$ is positive semidefinite with the same spectrum as $C(A)$, so this trace is at most $\lambda_{\max}(C(A)) \operatorname{tr} B \le L_* (\operatorname{per} A) \operatorname{tr} B$ by Theorem 52. For A_* the all-ones vector is a top eigenvector of $C(A_*)$, which gives equality for $B = J$. □

Remark 56 (numerical evidence)*.* We maximized Λ over $A = VV^*$ with unit rows of $V \in \mathbb{C}^{5 \times k}$ by quasi-Newton ascent from 1,896 starts of full rank $k = 5$, complex and real, drawn from Gaussian, near-one-factor, sparse and nearly rank-deficient distributions. Of these, 1,876 ascents end at 1.8727665160, within $2 \cdot 10^{-15}$ of L_*, at a matrix with all off-diagonal moduli 0.2773; the remaining 20 stop at lower values. Over singular matrices, with $k = 4, 3, 2$ and 1,040 starts, the largest values found are 1.7988818, 1.7046085 and 1.5350409. This extends the ascent of Section 6.

The proofs run as follows. Section 66.2 collects the exact scalar data. For Theorem 52, a lifted five-cycle reduces a transitive matrix to a circulant (Lemma 58). A circulant correlation matrix is described by its five spectral weights $x_m \ge 0$, $\sum x_m = 1$; its permanent is an explicit quintic $\mathcal{P}(x)$ with integer coefficients, and the cofactor eigenvalues are the partial derivatives $\partial_m \mathcal{P}/5$ (Lemma 59). Translation of the frequency labels is a diagonal phase conjugation, so it suffices to bound $\partial_0 \mathcal{P}/(5\mathcal{P})$ on the simplex. Near the scalar point the margin $5L_*\mathcal{P} - \partial_0 \mathcal{P}$ is convex in the four transverse weights and stationary on the scalar fibre (Lemma 60); elsewhere a Bernstein subdivision certificate in exact integer arithmetic proves $\partial_0 \mathcal{P}/(5\mathcal{P}) < 1.8727665$ (Lemma 61). For Theorem 53 the symmetric group and complex conjugation split the tangent space into five sectors, on each of which the second variation is a scalar computed exactly (Lemmas 62 and 63). Explicit bounds for the Taylor coefficients of the permanent and of the cofactor matrix along lines (Lemma 64) and a Schur-complement form of the top eigenvalue (Lemma 65) control the cubic remainder. Theorem 54 combines the isolated-row identities with Theorem 28 and its equality case (Section 66.5).

66.2 Scalar matrices of order five

Let $P_m(b) = \operatorname{per}((1-b)I_m + bJ_m) = \sum_{j=0}^{m} \binom{m}{j} d_j b^j$, where d_j is the number of derangements of j points; the sum groups the permutations by their moved points. Thus

$$P_2 = 1 + b^2, \quad P_3 = 1 + 3b^2 + 2b^3, \quad P_4 = 1 + 6b^2 + 8b^3 + 9b^4, \quad P_5 = 1 + 10b^2 + 20b^3 + 45b^4 + 44b^5.$$

Lemma 57. *For $0 < b \le 1$ the cofactor matrix of $A_{5,b} = (1-b)I + bJ$ is $C(A_{5,b}) = P_4(b)I + Z_5(b)(J - I)$ with*

$$Z_5(b) = b + 3b^2 + 9b^3 + 11b^4,$$

its largest eigenvalue is $M_5(b) = P_4 + 4Z_5 = 1 + 4b + 18b^2 + 44b^3 + 53b^4$, attained on the constant vector with spectral gap $5Z_5(b)$, and $F_5 = M_5/P_5$. The derivative is $F_5' = -4R_5/P_5^2$ with R_5 as in (75)*. The polynomial R_5 is irreducible and has exactly one positive root $b_5 \in (0.2773470462, 0.2773470463)$; F_5 increases on $(0, b_5]$ and decreases on $[b_5, 1]$, so $L_* = F_5(b_5)$ is the strict maximum of F_5. Moreover $1.8727665 < L_* < 1.8727666$, and L_* is a root of the irreducible polynomial* (76)*.*

Proof. The diagonal cofactors are permanents of $A_{4,b}$. The off-diagonal cofactor $\operatorname{per} A_{5,b}(1 \mid 2)$ is a polynomial in b whose direct expansion gives $Z_5(b)$, and all off-diagonal cofactors agree by permutation invariance. Hence $C(A_{5,b}) = P_4 I + Z_5(J - I)$, with eigenvalue $P_4 + 4Z_5$ on the constant vector and $P_4 - Z_5$ on its orthogonal complement, and $Z_5 > 0$ for $b > 0$. The derivative formula is a polynomial identity. The coefficient sequence of R_5 has one sign change, so Descartes' rule gives exactly one positive root, and $R_5(0.2773470462) < 0 < R_5(0.2773470463)$ locates it. Since $R_5(0) < 0 < R_5(1)$, $F_5' > 0$ before b_5 and $F_5' < 0$ after it. The lower bound for L_* is the value $F_5(0.2773470462) > 1.8727665$, and the upper bound follows from $L_* \le M_5(0.2773470463)/P_5(0.2773470462) < 1.8727666$, since M_5 and P_5 increase in b. The resultant of $R_5(b)$ and $M_5(b) - LP_5(b)$ with respect to b is a nonzero rational multiple of (76), which therefore vanishes at $L = F_5(b_5)$. Irreducibility of R_5 and of (76) over $\mathbb{Q}$ is checked by exact factorization. All statements are verified in exact arithmetic by `verify_transitive_order5.py`. □

The matrix A_* is positive definite, with eigenvalues $1 + 4b_5$ and $1 - b_5$, and $C(A_*)$ has the simple top eigenvalue $M_5(b_5)$ with eigenvector $\mathbf{1}$ and gap $g = 5Z_5(b_5)$. The values agree with Table 1.

66.3 Transitive correlation matrices

Throughout this subsection $\omega = e^{2\pi \mathrm{i}/5}$ and indices are taken modulo 5 in $\{0, \dots, 4\}$. Let

$$f_m = 5^{-1/2}\big(\omega^{mj}\big)_{j=0}^{4} \qquad (m = 0, \dots, 4).$$

Lemma 58. *If A is a transitive correlation matrix of order five, then after a simultaneous permutation of rows and columns and a conjugation by a diagonal unitary matrix, A is a circulant: $A_{jk} = \sum_{m=0}^{4} x_m \omega^{m(j-k)}$ with $x_m \ge 0$ and $\sum_m x_m = 1$.*

Proof. The group $\mathcal{G}$ of underlying permutations is transitive on five points, so 5 divides $|\mathcal{G}|$ by the orbit-stabilizer theorem, and Cauchy's theorem gives an element of order five, which in S_5 is a five-cycle. Relabel so that it is $\sigma(j) = j + 1$, and let $U = \Psi \Pi_\sigma$ be a lift with $UAU^* = A$, where $\Psi = \operatorname{diag}(\psi_j)$ is unitary and $(\Pi_\sigma)_{j+1,j} = 1$. Choose ς with $\varsigma^5 = \prod_j \psi_j$, and define $\xi_0 = 1$ and $\xi_{j+1} = \varsigma\, \xi_j / \psi_{j+1}$; the product condition makes this consistent around the cycle. With $\Xi = \operatorname{diag}(\xi_j)$ one checks entrywise that $\Xi U \Xi^* = \varsigma\, \Pi_\sigma$. Hence $A' = \Xi A \Xi^*$ satisfies $\Pi_\sigma A' \Pi_\sigma^{\mathsf{T}} = A'$, that is $A'_{j+1,k+1} = A'_{jk}$, so A' is a Hermitian circulant. Its eigenvectors are the f_m, and writing its eigenvalues as $5x_m \ge 0$ gives the displayed form; the unit diagonal gives $\sum x_m = 1$. □

For $x \in \mathbb{R}^5$ let $A(x)$ be the circulant of Lemma 58 and $\mathcal{P}(x) = \operatorname{per} A(x)$. Let $\Delta_4 = \{x \in \mathbb{R}^5 : x \geq 0,\ \sum x_m = 1\}$.

Lemma 59. *The permanent of $A(x)$ is the homogeneous quintic*

$$\mathcal{P}(x) = 120 \sum_i x_i^5 + 150 \sum_i \sum_{s=1,2} x_i^3 x_{i+s} x_{i-s} + 100 \sum_i \sum_{s=1,2} x_i x_{i+s}^2 x_{i-s}^2 + 25\, x_0 x_1 x_2 x_3 x_4. \tag{77}$$

It is invariant under the translations $x_m \mapsto x_{m+l}$ and the dilations $x_m \mapsto x_{km}$, $k \in \{1,2,3,4\}$. The eigenvalues of $C(A(x))$ are $\frac{1}{5}\partial_m \mathcal{P}(x)$, $m = 0, \ldots, 4$, so for $x \in \Delta_4$

$$\Lambda(A(x)) = \max_m \frac{\partial_m \mathcal{P}(x)}{5\mathcal{P}(x)},$$

and $\mathcal{P} > 0$ on Δ_4.

Proof. Expanding $\prod_j A(x)_{j,\pi(j)}$ over $\pi \in S_5$ and over label sequences $(m_0, \ldots, m_4)$ gives

$$\mathcal{P}(x) = \sum_{(m_j)} \Big(\prod_j x_{m_j} \Big) \sum_{\pi \in S_5} \omega^{\sum_j m_j (j - \pi(j))}.$$

Fix an exponent vector $\mathbf{a}$ and let $n_\rho(\mathbf{a})$ count the pairs, of a label sequence with monomial $x^{\mathbf{a}}$ and a permutation π, whose exponent sum is ρ modulo 5. Multiplying all positions by a unit k modulo five, $(m_j) \mapsto (m_{k^{-1}j})$ and $\pi \mapsto k\pi k^{-1}$, is a bijection of these pairs that multiplies the exponent by k. Hence $n_1 = n_2 = n_3 = n_4$, and the coefficient of $x^{\mathbf{a}}$ is $n_0(\mathbf{a}) - n_1(\mathbf{a}) \in \mathbb{Z}$ because $\sum_{\rho \neq 0} \omega^\rho = -1$. Enumerating the $5^5 \cdot 120$ pairs gives the 26 nonzero coefficients in (77). A translation of labels is the conjugation by $\operatorname{diag}(\omega^{lj})$, and a dilation is a simultaneous permutation of rows and columns; both preserve the permanent.

Since $\partial A(x)/\partial x_m = 5 f_m f_m^*$ and $\partial \operatorname{per} A / \partial A_{jk} = C(A)_{jk}$, we get $\partial_m \mathcal{P} = 5 \sum_{j,k} C_{jk} (f_m)_j \overline{(f_m)_k} = 5 f_m^{\mathsf{T}} C \overline{f_m}$. The cofactor matrix of a circulant is a circulant, because $C(\Pi A \Pi^{\mathsf{T}}) = \Pi C(A) \Pi^{\mathsf{T}}$ for permutation matrices Π, so $\overline{f_m} = f_{-m}$ is an eigenvector of C and $\frac{1}{5}\partial_m \mathcal{P}$ is the eigenvalue of $C(A(x))$ on f_{-m}. These five values exhaust the spectrum. Finally $\mathcal{P}(x) = \operatorname{per} A(x) > 0$ on Δ_4, as recalled at the start of this section. □

By the translation invariance, $\partial_m \mathcal{P}(x) = \partial_0 \mathcal{P}(\tau_m x)$, where $(\tau_m x)_k = x_{k+m}$. Hence Theorem 52 reduces to the single inequality

$$G_L(x) = 5L\, \mathcal{P}(x) - \Big(\sum_k x_k \Big) \partial_0 \mathcal{P}(x) \geq 0 \qquad (x \in \Delta_4) \tag{78}$$

at $L = L_*$, with its equality case. At the scalar point $x^\circ = (\frac{1+4b_5}{5}, \frac{1-b_5}{5}, \ldots, \frac{1-b_5}{5})$ we have $A(x^\circ) = A_*$.

Lemma 60 (convexity box)**.** *Let $\mathcal{B}$ be the set of $x \in \Delta_4$ with $|x_0 - \frac{211}{500}| \leq \frac{1}{200}$ and $|x_k - \frac{1-x_0}{4}| \leq \frac{1}{200}$ for $k = 1, 2, 3, 4$. Then $G_{L_*} \geq 0$ on $\mathcal{B}$, with equality only at x°.*

Proof. Write $a = x_0$, $x_k = \frac{1-a}{4} + y_k$ for $k = 1, 2, 3$ and $x_4 = \frac{1-a}{4} - y_1 - y_2 - y_3$, and let $Q_L(a, y) = G_L(x)$. For $L \in [1.8727665, 1.8727666]$, $|a - \frac{211}{500}| \leq \frac{1}{200}$ and $|y_k| \leq \frac{1}{200}$, the Hessian H of Q_L in (y_1, y_2, y_3) is positive definite. Indeed, at the centre $(a, y, L) = (\frac{211}{500}, 0, 1.8727665)$ the Hessian H_0 is an explicit rational matrix with smallest eigenvalue about 107.8; expanding each Hessian entry about the centre and summing the absolute values of the monomial contributions at the box radii

bounds every row sum of $|H - H_0|$ by $B_H < 60.24$; and the three leading principal minors of $H_0 - B_H I$ are positive rationals. Thus $H \succeq H_0 - B_H I \succ 0$ on the box.

For fixed a the fibre $\{y : x \in \mathcal{B}\}$ is convex, contains $y = 0$ and lies in the cube $|y_k| \le \frac{1}{200}$, so $Q_L(a, \cdot)$ is strictly convex on it. The dilation $x_m \mapsto x_{2m}$ fixes x_0, permutes (x_1, x_2, x_4, x_3) cyclically and preserves $\mathcal{P}$ and $\partial_0\mathcal{P}$ (Lemma 59). It therefore preserves $Q_L(a, \cdot)$ and acts on y as a linear map ϖ whose only fixed vector is 0, because a fixed vector has $y_1 = y_2 = y_3 = -(y_1 + y_2 + y_3)$. The transpose ϖ^{T} then has no nonzero fixed vector either, and differentiating $Q_L(a, \varpi y) = Q_L(a, y)$ at $y = 0$ shows that the gradient at $y = 0$ is fixed by ϖ^{T}, so it vanishes. Hence $Q_L(a, y) \ge Q_L(a, 0)$, with equality only at $y = 0$.

At $y = 0$ the circulant is $A_{5,\beta}$ with $\beta = (5a - 1)/4 \in (0.27, 0.29)$, its mode-zero eigenvalue is $M_5(\beta)$ (Lemma 57), and so $Q_L(a, 0) = 5P_5(\beta)(L - F_5(\beta))$. At $L = L_*$ this is nonnegative and vanishes only at $\beta = b_5$, that is at x°, by Lemma 57. Since $L_* \in [1.8727665, 1.8727666]$, the convexity applies at $L = L_*$, which proves the lemma. □

Lemma 61 (Bernstein certificate)**.** *Let $L_0 = 1.8727665$. Then $G_{L_0} > 0$ on $\Delta_4 \setminus \mathcal{B}$.*

Proof. Let Σ be a simplex with vertices $v^0, \dots, v^4 \in \Delta_4$ and write $x = \sum_k \theta_k v^k$ with barycentric coordinates $\theta \ge 0$, $\sum \theta_k = 1$. The homogeneous quintic G_{L_0} becomes a homogeneous quintic $\sum_{|\mathbf{a}|=5} g_{\mathbf{a}} \theta^{\mathbf{a}}$. If every $g_{\mathbf{a}} \ge 0$ and every coefficient of a pure power θ_k^5, which equals $G_{L_0}(v^k)$, is positive, then $G_{L_0} > 0$ on Σ, because some $\theta_k \ge \frac{1}{5}$. Starting from Δ_4 itself, where $\theta = x$, we bisect a longest edge $v^i v^j$ whenever a simplex is neither certified in this way nor contained in $\mathcal{B}$. On the child in which v^j is replaced by the midpoint, the substitution is $\theta_i = \theta'_i + \theta'_j/2$, $\theta_j = \theta'_j/2$; multiplying by 2^5 keeps every coefficient an integer. A simplex is contained in the convex set $\mathcal{B}$ when its five vertices are. The script `verify_transitive_order5.py` runs this subdivision in exact integer arithmetic. It terminates with 1,477 certified simplices and 56 simplices inside $\mathcal{B}$, at maximum depth 34. The leaves cover Δ_4, which proves the lemma. □

Proof of Theorem 52. By Lemma 58 we may assume $A = A(x)$ with $x \in \Delta_4$; the relabeling and the diagonal conjugation change neither Λ nor membership in $\mathcal{O}$, because A_* is fixed by every relabeling. By Lemmas 60 and 61, $\partial_0\mathcal{P}(z) \le 5L_*\mathcal{P}(z)$ for every $z \in \Delta_4$, with equality only at $z = x^\circ$; on $\Delta_4 \setminus \mathcal{B}$ the inequality holds even with $L_0 < L_*$. Applying this to $z = \tau_m x$ and using Lemma 59 gives $\Lambda(A(x)) \le L_*$. Equality forces $\tau_m x = x^\circ$ for some m, so $A(x)$ is the conjugate of A_* by a diagonal unitary of the form $\mathrm{diag}(\omega^{lj})$. Conversely every DA_*D^* is transitive and attains L_*. □

66.4 Local rigidity at an explicit radius

Let V be the real vector space of Hermitian matrices of order five with zero diagonal, with the real Frobenius inner product $\langle X, Y\rangle = \mathrm{Re}\,\mathrm{tr}(XY^*)$. Define the subspaces

$$V_c = \mathbb{R}(J - I), \qquad V_s = \{(w_j + w_k)_{j\ne k} : w \in \mathbb{R}^5,\ \textstyle\sum w_j = 0\},$$
$$V_r = \{X \in V \text{ real} : X\mathbf{1} = 0\}, \qquad V_g = \{(\mathrm{i}(\vartheta_j - \vartheta_k))_{j\ne k} : \vartheta \in \mathbb{R}^5\},$$
$$V_i = \{\mathrm{i}Y :\ Y \text{ real antisymmetric},\ Y\mathbf{1} = 0\},$$

of dimensions $1, 4, 5, 4, 6$. The space V_g is the tangent space of $\mathcal{O}$ at A_*, since $\frac{d}{dt} e^{\mathrm{i}t\Theta} A_* e^{-\mathrm{i}t\Theta}|_{t=0} = \mathrm{i}[\Theta, A_*]$ has entries $\mathrm{i}b_5(\vartheta_j - \vartheta_k)$ for $\Theta = \mathrm{diag}(\vartheta)$. The group S_5 acts on V by $X \mapsto \Pi X \Pi^{\mathsf{T}}$, and complex conjugation acts by $X \mapsto \overline{X}$. These sectors are the order-five instances of the four transverse sectors and the phase kernel of Theorem 1.

Lemma 62. *$V = V_c \oplus V_s \oplus V_r \oplus V_g \oplus V_i$ is an orthogonal decomposition, and every real quadratic form Φ on V invariant under S_5 and complex conjugation has the form $\Phi(X) = \sum_\iota \kappa_\iota \|X_\iota\|^2$, where X_ι is the component of X in V_ι, $\iota \in \{c, s, r, g, i\}$.*

Proof. Write $X = R + \mathrm{i}Y$ with R real symmetric and Y real antisymmetric. Invariance under conjugation gives $\Phi(R + \mathrm{i}Y) = \Phi(R - \mathrm{i}Y)$, so the polar form of Φ pairs R with $\mathrm{i}Y$ to zero and $\Phi(X) = \Phi(R) + \Phi(\mathrm{i}Y)$.

On real symmetric matrices, use the ten unordered pairs $\{j, k\}$ as coordinates. An S_5-invariant symmetric bilinear form has a Gram matrix that is constant on the orbits of S_5 on pairs of pairs; these orbits are given by equal pairs, pairs sharing one point, and disjoint pairs. The Gram matrix is therefore $\alpha_0 I + \alpha_1 \mathcal{E} + \alpha_2(\mathbf{J} - I - \mathcal{E})$, where $\mathcal{E}$ is the adjacency matrix of the relation "sharing one point" and $\mathbf{J}$ is the all-ones matrix of order ten. Direct computation gives $\mathcal{E}X = 6X$ on V_c, $\mathcal{E}X = X$ on V_s (because $\sum_{l \neq j,k}(w_j + w_l + w_k + w_l) = w_j + w_k$ when $\sum w = 0$) and $\mathcal{E}X = -2X$ on V_r (because zero row sums give $\sum_{l \neq j,k}(X_{jl} + X_{kl}) = -2X_{jk}$). Also $\mathbf{J}$ vanishes on $V_s \oplus V_r$. The dimensions sum to ten, so these are the eigenspaces of the Gram matrix.

On real antisymmetric Y, an invariant symmetric bilinear form is $\sum \eta(j, k, l, m) Y_{jk} Y'_{lm}$ with η invariant under S_5 and antisymmetric in (j, k) and in (l, m). If $\{j, k\} \cap \{l, m\} = \emptyset$, the transposition of l and m fixes j and k, and antisymmetry forces $\eta = 0$. Since S_5 is transitive on ordered triples of distinct points, the remaining values are fixed by $\eta(j, k, j, k)$ and $\eta(j, k, j, l)$ together with antisymmetry. The form is thus a combination of $\langle Y, Y' \rangle$ and $\langle Y, \mathcal{S}Y' \rangle$, where $(\mathcal{S}Y)_{jk} = \sum_{l \neq j,k}(Y_{jl} + Y_{lk})$. Here $\mathcal{S}Y = 3Y$ on gradients $Y_{jk} = \vartheta_j - \vartheta_k$ and $\mathcal{S}Y = -2Y$ when $Y\mathbf{1} = 0$, and the dimensions $4 + 6$ exhaust the space. Orthogonality of the five subspaces follows from the self-adjointness of $\mathcal{E}$ and $\mathcal{S}$. □

For $X \in V$ and real r put $A(r) = A_* + rX$ and expand

$$\operatorname{per} A(r) = \sum_{k=0}^{5} \operatorname{per}_k(X)\, r^k, \qquad C(A(r)) = \sum_{k=0}^{4} K_k(X)\, r^k,$$

so that $\operatorname{per}_0 = P_5(b_5)$ and $K_0 = C(A_*)$. Let $u = 5^{-1/2}\mathbf{1}$ and

$$c_2(X) = \frac{1}{P_5(b_5)}\Big(u^* K_2 u + \frac{\|(I - uu^*)K_1 u\|^2}{g} - L_* \operatorname{per}_2\Big). \tag{79}$$

Lemma 63. *For every $X \in V$ one has $u^* K_1 u = L_* \operatorname{per}_1$, and $c_2(X) = \sum_\iota \kappa_\iota \|X_\iota\|^2$ with*

$$\kappa_c \in (-0.538516, -0.538515), \qquad \kappa_s \in (-0.294204, -0.294203),$$
$$\kappa_r \in (-0.174449, -0.174448), \qquad \kappa_i \in (-0.17023965, -0.17023964), \qquad \kappa_g = 0.$$

Moreover $(I - uu^*)K_1 u$ vanishes on $V_c \oplus V_r \oplus V_i$ and has norm at most $0.09776\|X\|$ on V_s.

Proof. The map $X \mapsto u^* K_1 u - L_* \operatorname{per}_1$ is a real linear functional on V invariant under S_5 and conjugation, hence a multiple of $X \mapsto \langle X, J - I \rangle$ by Lemma 62. Along $X = J - I$ the matrices $A(r) = A_{5, b_5 + r}$ stay scalar and $\Lambda(A(r)) = F_5(b_5 + r)$ is stationary at $r = 0$, so the functional vanishes.

The form c_2 is a quadratic form invariant under S_5 and conjugation, so Lemma 62 applies, and $\kappa_\iota = c_2(X)/\|X\|^2$ for any nonzero $X \in V_\iota$. We use the representatives $J - I$; the matrix of V_s with $w = (1, -1, 0, 0, 0)$; $X_{12} = X_{34} = 1$, $X_{13} = X_{24} = -1$ (zero row sums); $X_{12} = X_{23} = \mathrm{i}$, $X_{13} = -\mathrm{i}$; and the tangent vector $\mathrm{i}[\Theta, A_*]$ for $\vartheta = (1, 0, 0, 0, 0)$, with the remaining upper entries zero and Hermitian completion. Exact expansion of all permanents to second order in r (script `verify_hessian_order5.py`) gives the following values at $b = b_5$, where $\|X\|^2 = 20, 12, 8, 6, 8b^2$ respectively.

sector	$\|(I-uu^*)K_1u\|^2$	u^*K_2u	per_2
V_c	0	$6(53b^2+22b+3)$	$10(44b^3+27b^2+6b+1)$
V_s	$\frac{18}{5}(17b^3+7b^2+b-1)^2$	$-\frac{24}{5}(9b^2+b-1)$	$-6(11b^3+3b^2-b-1)$
V_r	0	$\frac{4}{5}(15b^2+2b+1)$	$4(4b^3+3b^2-2b+1)$
V_i	0	$-\frac{3}{5}(13b^2+6b-1)$	$-3(4b^3+b^2+2b-1)$
V_g	$4b^2(11b^3+9b^2+3b+1)^2$	$\frac{24}{5}b^2(7b^2+4b+1)$	$4b^2(11b^3+9b^2+3b+1)$

Evaluating (79) with interval arithmetic on $b \in [0.2773470462, 0.2773470463]$ gives the stated enclosures, and an interval of width below $3\cdot 10^{-10}$ containing 0 for V_g. In fact $\kappa_g = 0$. Near A_* the top eigenvalue of the cofactor matrix is simple, so Λ is smooth there, and Λ is constant along each curve $\zeta(t) = e^{\mathrm{i}t\Theta}A_*e^{-\mathrm{i}t\Theta}$, whose velocities $\zeta'(0)$ fill V_g. Since the first variation vanishes on V, the t^2 coefficient of $\Lambda(\zeta(t))$ equals $c_2(\zeta'(0))$, which is therefore zero. The quadratic form $X \mapsto \|(I-uu^*)K_1u\|^2$ is invariant under S_5 and conjugation, so Lemma 62 makes it diagonal in the sectors; by the first column of the table it vanishes on V_c, V_r and V_i, and on V_s its value per unit $\|X\|^2$ is $\frac{3}{10}(17b^3+7b^2+b-1)^2 < 0.09776^2$. □

Lemma 64. *Let $X \in V$ with $\|X\| = 1$. For $1 \le k \le 5$ let N_k be the number of sets of k off-diagonal cells of the 5×5 grid with distinct rows and distinct columns, and for $1 \le k \le 4$ let N'_k be the largest number of such sets inside the grid of one minor $(j \mid l)$. Then $(N_1,\dots,N_5) = (20,130,320,265,44)$, $(N'_1,\dots,N'_4) = (13,48,53,11)$, and*

$$|\operatorname{per}_k(X)| \le \sqrt{N_k/k!}\,P_{5-k}(b_5), \qquad \|K_k(X)\|_{\mathrm{F}} \le (5-k)\sqrt{N'_k/k!}\,P_{4-k}(b_5).$$

Also $|\operatorname{per}_1(X)| \le \sqrt{20}\,Z_5(b_5)$.

Proof. Expanding $\operatorname{per}(A_* + rX)$ by choosing in each row either the entry of A_* or that of rX, and grouping by the set $\mathcal{M}$ of chosen cells of X, gives $\operatorname{per}_k = \sum_{|\mathcal{M}|=k} X^{\mathcal{M}} \operatorname{per} A_*[\mathcal{M}]$, where $X^{\mathcal{M}} = \prod_{c\in\mathcal{M}} X_c$, $\mathcal{M}$ ranges over sets of k off-diagonal cells with distinct rows and columns, and $A_*[\mathcal{M}]$ is the submatrix of A_* on the remaining rows and columns. That submatrix has entries 1 on a partial permutation pattern and b_5 elsewhere; raising further entries from b_5 to 1 completes the pattern to a permutation and does not decrease the permanent of this nonnegative matrix, so $0 \le \operatorname{per} A_*[\mathcal{M}] \le P_{5-k}(b_5)$. Cauchy–Schwarz and $\sum_{\mathcal{M}} |X^{\mathcal{M}}|^2 \le e_k(|X_c|^2) \le \|X\|^{2k}/k!$, where e_k is the elementary symmetric function, give the first bound. For the cofactor entries, $K_k[j,l]$ is the same sum restricted to sets $\mathcal{M}$ inside the minor grid $(j \mid l)$, with complementary permanents at most $P_{4-k}(b_5)$. Applying Cauchy–Schwarz entrywise and summing over (j,l), where each $\mathcal{M}$ lies in exactly $(5-k)^2$ minor grids, gives the bound for $\|K_k\|_{\mathrm{F}}$. Finally $\operatorname{per}_1 = Z_5(b_5)\sum_{j\ne k} X_{jk}$, and $|\sum_{j\ne k} X_{jk}| \le \sqrt{20}\|X\|$. The counts are enumerated in `verify_local_radius_order5.py`. □

Lemma 65. *Let $\mu(r) = \lambda_{\max}(C(A(r)))$, $\varrho(r) = \sum_{k=1}^{4}\|K_k\|r^k$, $E_m(r) = \sum_{k\ge m}\|K_k\|r^k$ and $\beta_1 = (I-uu^*)K_1u$. If $2\varrho(r) < g$, then*

$$\Big|\mu(r) - M_5(b_5) - r\,u^*K_1u - r^2\Big(u^*K_2u + \frac{\|\beta_1\|^2}{g}\Big)\Big| \le E_3 + \frac{2\varrho\,(r\|\beta_1\| + E_2)^2}{g(g-2\varrho)} + \frac{2r\|\beta_1\|E_2 + E_2^2}{g}.$$

Proof. Decompose $K(r) = C(A(r))$ along u and $u^\perp$: $\alpha(r) = u^*K(r)u$, $\beta(r) = (I-uu^*)K(r)u$, and $\Gamma(r)$ the compression of $K(r)$ to $u^\perp$. At $r=0$, $\alpha = M_5(b_5)$, $\beta = 0$ and $\Gamma = (M_5(b_5)-g)I$ by Lemma 57. Weyl's inequality gives $|\mu - M_5(b_5)| \le \varrho$ and $\lambda_{\max}(\Gamma) \le M_5(b_5) - g + \varrho$, so $\mu - \lambda_{\max}(\Gamma) \ge g - 2\varrho > 0$. A top eigenvector therefore has a nonzero u-component, and the eigenvalue equation in block form gives the exact identity $\mu = \alpha + \beta^*(\mu-\Gamma)^{-1}\beta$. Here $\|(\mu-\Gamma)^{-1} - g^{-1}I\| \le 2\varrho/(g(g-2\varrho))$,

because $\|(\mu-\Gamma)-gI\| \le |\mu-M_5(b_5)| + \|\Gamma-\Gamma(0)\| \le 2\varrho$. The remaining terms come from $|\alpha - M_5(b_5) - ru^*K_1u - r^2u^*K_2u| \le E_3$, from $\|\beta - r\beta_1\| \le E_2$, which gives $\|\beta\| \le r\|\beta_1\| + E_2$, and from $\big|\|\beta\|^2 - r^2\|\beta_1\|^2\big| \le 2r\|\beta_1\|E_2 + E_2^2$. $\square$

Proof of Theorem 53. The orbit $\mathcal{O}$ is compact, so there is a diagonal unitary D with $\|D^*AD - A_*\| = r$. For $r>0$ put $X = (D^*AD - A_*)/r \in V$. Differentiating $\Theta \mapsto \|D^*AD - e^{\mathrm{i}\Theta}A_*e^{-\mathrm{i}\Theta}\|^2$ at its minimum gives $X \perp V_g$. Since $\Lambda(A) = \Lambda(A_* + rX)$, it suffices to bound $\Lambda(A(r)) = \mu(r)/\operatorname{per}A(r)$ for unit $X \perp V_g$ and $0 < r \le \frac{1}{100}$.

By Lemma 57, $M_5(b_5) = L_*P_5(b_5)$, and by Lemma 63, $u^*K_1u = L_*\operatorname{per}_1$. Therefore

$$\mu(r) - L_*\operatorname{per}A(r) = r^2P_5(b_5)\,c_2(X) + \varepsilon_\mu(r) - L_*\sum_{k\ge3}\operatorname{per}_k\, r^k,$$

where ε_μ is the error bounded in Lemma 65. Lemma 63 gives $c_2(X) \le \kappa_i < -0.17023964$ on $V_g^\perp$ and $\|\beta_1\| \le 0.09776$. Insert the bounds of Lemma 64 into Lemma 65 and into the tail $\sum_{k\ge3}|\operatorname{per}_k|r^k$. Every error term is then r^3 times a function of r that is nondecreasing on $[0,\frac{1}{100}]$ (a polynomial with nonnegative coefficients, in one term divided by the decreasing positive function $g - 2\varrho(r)$), so the ratio of the total error to r^2 increases with r. At $r = \frac{1}{100}$ these bounds are evaluated in exact rational arithmetic by the script `verify_local_radius_order5.py`. It uses $P_5(b_5) \ge P_5(0.2773470462)$, $P_m(b_5) \le P_m(0.2773470463)$, $g \ge 5Z_5(0.2773470462)$ and $L_* < 1.8727666$, and it gives $2\varrho < g$ and

$$\mu(r) - L_*\operatorname{per}A(r) \le -0.2074\,r^2, \qquad \operatorname{per}A(r) \le 2.5697 \qquad (0 < r \le \tfrac{1}{100}).$$

Since $\operatorname{per}A(r) > 0$, dividing gives $\Lambda(A(r)) - L_* \le -0.0807r^2 \le -r^2/25$.

The second variation of Λ at A_* in direction X is $2c_2(X)$, because the first variation vanishes. By Lemma 63 it is zero on V_g and has eigenvalues $2\kappa_c, 2\kappa_s, 2\kappa_r, 2\kappa_i \in [-1.0771, -0.3404]$ on $V_g^\perp$. $\square$

66.5 The unrestricted window

The next lemma is the eigenvector form of the isolated-row bound used in Sections 36 and 62.1.

Lemma 66 (row projection)**.** *Let A be a correlation matrix of order $n \ge 2$ with $\ell = \Lambda(A) > 1$, and let e be a unit vector with $C(A)e = \ell(\operatorname{per}A)e$. For each i let A_i be the principal submatrix deleting index i. Then*

$$|e_i|^2 \le \frac{\Lambda(A_i)}{\Lambda(A_i) + 4\ell(\ell-1)}.$$

Proof. Relabel so that $i = n$ and write $A = \begin{pmatrix} T & z \\ z^* & 1 \end{pmatrix}$, $\widehat{C} = C(T)^{\mathsf{T}}$, $p = \operatorname{per}T$ and $q = z^*\widehat{C}z$. Lemma 70 gives $\operatorname{per}A = p + q$, $C(A)_{nn} = p$ and $C(A)_{nj} = (\widehat{C}z)_j$ for $j < n$, and Lemma 69 gives $\widehat{C} \succeq 0$ and $p > 0$. Put $\chi = p/(p+q) \in (0,1]$ and $e' = (e_1, \dots, e_{n-1})$. Row n of the eigenvalue equation reads $(\ell(p+q) - p)e_n = \sum_{j<n}(\widehat{C}z)_j e_j$. Cauchy–Schwarz for the positive semidefinite form $\widehat{C}$ bounds the right side by $(z^*\widehat{C}z)^{1/2}(\overline{e'}^{\,*}\widehat{C}\,\overline{e'})^{1/2}$, and $\overline{e'}^{\,*}\widehat{C}\,\overline{e'} = e'^{*}C(T)e' \le \Lambda(T)\,p\,(1 - |e_n|^2)$. Dividing by $p+q$,

$$(\ell - \chi)^2|e_n|^2 \le \Lambda(T)\,\chi(1-\chi)(1-|e_n|^2).$$

Here $\ell \ge \chi$ because $\ell\operatorname{per}A$ is at least the diagonal entry $C(A)_{nn} = p$. For $\ell > 1$ the function $\chi(1-\chi)/(\ell-\chi)^2$ on $[0,1]$ has its maximum $1/(4\ell(\ell-1))$ at $\chi = \ell/(2\ell-1)$. If $|e_n| = 1$ the display forces $\ell = \chi \le 1$, which is excluded; otherwise rearranging gives the claim. $\square$

Proof of Theorem 54. Let A have order five and $\ell = \Lambda(A)$; we may assume $\ell > 1$. Each A_i has order four, so $\Lambda(A_i) \le \gamma_4$ by Theorem 28. Put $h = 4\ell(\ell-1)$. Since $t \mapsto t/(t+h)$ increases, Lemma 66 and $\sum_i |e_i|^2 = 1$ give

$$1 = \sum_{i=1}^{5} |e_i|^2 \le \sum_{i=1}^{5} \frac{\Lambda(A_i)}{\Lambda(A_i)+h} \le \frac{5\gamma_4}{\gamma_4+h}. \tag{80}$$

Hence $h \le 4\gamma_4$, that is $\ell \le U := (1+\sqrt{1+4\gamma_4})/2$. Suppose $\ell = U$. Then $h = 4\gamma_4$, both inequalities in (80) are equalities, and strict monotonicity forces $\Lambda(A_i) = \gamma_4$ for every i. By the equality case of Theorem 28, every A_i is a diagonal unitary conjugate of $(1-b_4)I + b_4 J$, where $b_4 \approx 0.3131$ is the order-four optimizer. Every off-diagonal entry of A lies in some A_i, so all have modulus b_4. Every triangle $\{j,k,l\}$ misses at least one index, so $A_{jk}A_{kl}A_{lj} = b_4^3 > 0$, this cyclic product being invariant under diagonal unitary conjugation. Conjugating by the diagonal unitary that makes the entries A_{1k} positive, the triangle products through index 1 make every entry positive. Thus A is a diagonal unitary conjugate of A_{5,b_4}, and $\Lambda(A) = F_5(b_4) \le L_* < U$, a contradiction. Hence $\Lambda(A) < U$. With $\gamma_4 < 1.64523$ (Section 62.4), $U < 1.876674$, and A_* gives $\gamma_5 \ge L_*$.

If $\ell \ge L_*$, then $h \ge h_* = 4L_*(L_*-1) > 6.53795$, and Lemma 66 gives $|e_i|^2 \le \gamma_4/(\gamma_4+h_*) < 0.201051$ for every i, hence $|e_i|^2 \ge 1 - 4(0.201051) > 0.195795$. Lemma 66 rearranged reads $\Lambda(A_i) \ge h|e_i|^2/(1-|e_i|^2) \ge 6.53795 \cdot 0.195795/0.804205 > 1.5917$. The rational arithmetic is in `verify_window_order5.py`. □

66.6 Cyclic dephasing of circulant correlations

The Fourier permanent (77) also controls how fast the permanent of a circulant correlation matrix relaxes under cyclic dephasing. For a probability vector μ on $\mathbb{Z}/5$, Schur multiplication by the circulant correlation matrix $A(\mu)$ averages the phase conjugations $\operatorname{diag}(\omega^{sj})$ with weights μ_s, and $A(x) \circ A(\mu) = A(\mu * x)$, where $*$ is convolution on $\mathbb{Z}/5$. Accordingly consider

$$\dot{x}(t) = \sum_{s=1}^{4} a_s(t)\,(T_s - I)\,x(t), \qquad (T_s x)_k = x_{k+s}, \tag{81}$$

with locally integrable rates $a_s \ge 0$; then $x(t) = \mu_t * x(0)$ for probability vectors μ_t, and $A(x(t)) = A(x(0)) \circ A(\mu_t)$. Put

$$\mathcal{H}(x) = -\nabla\mathcal{P}(x)\cdot(T_1 - I)x = 5\mathcal{P}(x) - \sum_k x_{k+1}\,\partial_k\mathcal{P}(x), \qquad \gamma_{\mathrm{deph}} = \inf_{x \in \Delta_4 \setminus \{u\}} \frac{\mathcal{H}(x)}{\mathcal{P}(x)-1},$$

where $u = (1/5,\dots,1/5)$ is the point with $A(u) = I$, and the second expression for $\mathcal{H}$ is Euler's identity for the quintic $\mathcal{P}$.

Theorem 67 (sharp cyclic dephasing rate)**.** *The constant γ_{deph} satisfies $1.200046650748820 < \gamma_{\mathrm{deph}} < 1.200046650748821$. It is attained exactly at the five cyclic translates of one point $x^\natural = (a^\natural, 0, 0, b^\natural, 1-a^\natural-b^\natural)$ with $|a^\natural - 0.269289| < 1/4000$ and $|b^\natural - 0.383354| < 1/4000$. For every solution of (81),*

$$\operatorname{per} A(x(t)) - 1 \le \exp\Bigl(-\gamma_{\mathrm{deph}} \int_0^t \sum_{s=1}^{4} a_s(v)\,dv\Bigr)\bigl(\operatorname{per} A(x(0)) - 1\bigr),$$

and no larger constant than γ_{deph} has this property.

Near the identity $\mathcal{P}(x)-1=\frac{25}{2}|x-u|^2+O(|x-u|^3)$, and along a single unit-rate shift the squared norm of the slowest Fourier harmonic of $x-u$ decays at rate $(5-\sqrt{5})/2\approx 1.382$; so the optimal rate is determined away from the identity, at the face point $x^\natural$, where $A(x^\natural)$ has rank three. The permanent of the internal-state Gram matrix of partially distinguishable photons quantifies their full bunching [60]; for circulant five-photon internal-state Gram matrices under cyclic dephasing, Theorem 67 therefore shows that cyclic dephasing drives the excess full-bunching factor at one output, $\operatorname{per} A-1$, to zero at the sharp exponential rate. It is a quantitative form, on this family of circulant multipliers, of the Bapat–Sunder comparison $\operatorname{per}(A\circ B)\le \operatorname{per}(A)\prod_i b_{ii}$, which fails in general by Drury's counterexamples [29, 59].

Lemma 68. *On Δ_4, $\mathcal{P}\ge 1$ with equality only at u. Along* (81), *$\frac{d}{dt}\mathcal{P}(x(t))=-\sum_s a_s(t)\mathcal{H}_s(x(t))$ for almost every t, where $\mathcal{H}_s(x)=-\nabla\mathcal{P}(x)\cdot(T_s-I)x$, and $\mathcal{H}_s(\delta_k x)=\mathcal{H}_{ks}(x)$ for the dilations $(\delta_k x)_m=x_{km}$, $k\in\{1,2,3,4\}$. In particular $\inf_{x\ne u}\mathcal{H}_s(x)/(\mathcal{P}(x)-1)=\gamma_{\mathrm{deph}}$ for every $s\ne 0$.*

Proof. For a correlation matrix, Lieb's inequality $\operatorname{per} A\ge \operatorname{per} A_{11}\operatorname{per} A_{22}$ for complementary principal blocks [1] gives $\operatorname{per} A\ge 1+|A_{jk}|^2$ for every $j\ne k$, so $\mathcal{P}(x)>1$ unless $A(x)=I$, that is $x=u$. The derivative formula is the chain rule for the absolutely continuous solution. Since $\mathcal{P}\circ\delta_k=\mathcal{P}$ (Lemma 59) and δ_k is a permutation matrix, $\nabla\mathcal{P}(\delta_k x)=\delta_k\nabla\mathcal{P}(x)$; also $T_s\delta_k=\delta_k T_{ks}$. Hence $\mathcal{H}_s(\delta_k x)=-\nabla\mathcal{P}(x)\cdot(T_{ks}-I)x=\mathcal{H}_{ks}(x)$, and $\mathcal{P}(\delta_k x)=\mathcal{P}(x)$. □

Proof of Theorem 67. For a level g let $F_g=\mathcal{H}-g(\mathcal{P}-(\sum_k x_k)^5)$, a homogeneous quintic equal to $\mathcal{H}-g(\mathcal{P}-1)$ on Δ_4; it decreases in g because $\mathcal{P}\ge 1$. The script `verify_dephasing_order5.py` proves the following, in exact rational and integer arithmetic.

1. On the box $|x_i-\frac{1}{5}|\le\frac{1}{4000}$, $i\le 3$, F_g and its tangential gradient vanish at u, and its Hessian in $(x_0,\dots,x_3)$ is positive definite for $g\in[\frac{6}{5},1.200047]$ (centre Hessian with smallest eigenvalue 4.549, perturbation bound 3.28); hence $F_g>0$ on the box minus u.

2. A Bernstein subdivision as in Lemma 61 certifies $F_{1.200047}>0$ on every leaf that lies neither in that box nor in one of the five cyclic translates of $\mathcal{C}=\{|x_0-0.269289|\le\frac{1}{4000},\ |x_3-0.383354|\le\frac{1}{4000},\ 0\le x_1,x_2\le\frac{1}{4000}\}$: 13,004 certified leaves, 100 leaves in the box around u and 63 in the translates of $\mathcal{C}$, at depth at most 54.

3. On $\mathcal{C}$, with $x=(a,x_1,x_2,b,1-a-b-x_1-x_2)$, the partial derivatives of F_g in x_1 and x_2 are positive for $g\in[\frac{6}{5},1.200047]$, and on the face $x_1=x_2=0$ the Hessian of F_g in (a,b) is at least $80I$.

4. At the rational point $(a,b)=(0.269289117288171,\ 0.383353707281529)$ on the face, $F_{g_-}-|\nabla F_{g_-}|^2/160>0$ for $g_-=1.200046650748820$, and $F_{g_+}<0$ for $g_+=1.200046650748821$.

By (4), $\gamma_{\mathrm{deph}}<g_+$. For $g=g_-$: outside the boxes $F_{g_-}\ge F_{1.200047}>0$ by (2); near u, $F_{g_-}>0$ by (1); and on $\mathcal{C}$, decreasing x_1,x_2 to zero lowers F_{g_-} by (3), while on the face the strong convexity in (3) gives $F_{g_-}\ge F_{g_-}(\text{trial})-|\nabla F_{g_-}(\text{trial})|^2/160>0$ by (4). By translation invariance of $\mathcal{P}$ and $\mathcal{H}$ the same holds on every translate, so $F_{g_-}>0$ on $\Delta_4\setminus\{u\}$ and $\gamma_{\mathrm{deph}}>g_-$. On the compact set outside the box around u the ratio $\mathcal{H}/(\mathcal{P}-1)$ is continuous and exceeds 1.200047 near that box, so the infimum is attained; at $g=\gamma_{\mathrm{deph}}$ the same three steps show that $F_{\gamma_{\mathrm{deph}}}$ vanishes only on the faces of the translates of $\mathcal{C}$, where strict convexity leaves one zero in each. This identifies the minimizers. Strong convexity also gives $|(x_0^\natural,x_3^\natural)-(a,b)_{\mathrm{trial}}|\le|\nabla F_{\gamma_{\mathrm{deph}}}(\text{trial})|/40<10^{-6}$, which gives the strict location bounds.

For the evolution, Lemma 68 gives $\frac{d}{dt}(\mathcal{P}-1)\le-\gamma_{\mathrm{deph}}\sum_s a_s(\mathcal{P}-1)$ almost everywhere, and Grönwall's inequality gives the bound. For optimality take $a_1=1$, $a_s=0$ otherwise, and $x(0)=x^\natural$:

then the right derivative of $\mathcal{P}(x(t)) - 1$ at $t = 0$ is $-\gamma_{\mathrm{deph}}(\mathcal{P}(x^{\natural}) - 1)$, so the bound fails for small $t > 0$ with any larger constant. □

66.7 Verification scripts

The following scripts accompany this edition. Each proof script prints its checks and ends with `ALL CHECKS PASSED`.

- `verify_transitive_order5.py`: the integer coefficients (77), the scalar data of Lemma 57, the convexity box of Lemma 60 and the Bernstein certificate of Lemma 61, in exact arithmetic.
- `verify_hessian_order5.py`: the exact second-order expansions and the interval values of Lemma 63.
- `verify_local_radius_order5.py`: the counts and bounds of Lemma 64 and the final rational inequality in the proof of Theorem 53.
- `verify_window_order5.py`: the rational bounds in the proof of Theorem 54.
- `verify_dephasing_order5.py`: the four certificates in the proof of Theorem 67.
- `verify_order_three.py`: the identities of Lemmas 72 and 74 and the scalar facts in Theorem 73.
- `verify_color_tables.py`: the layer and content checks of Lemmas 45 and 47, the identities of Lemma 48, and the finite-n values of Remark 44.
- `search_order5.py` and `summarize_search.py`: the numerical ascent of Remark 56, which plays no role in the proofs.

67 Bordering identities and order three

This section proves, inside the article, the elementary facts on which Sections 36–41, 56–62 and 66 rely: positivity of the cofactor matrix, the bordering identities, the recursive upper bound, the general lower bound, the unrestricted order-three maximum and the rank-two roots-of-unity identities. Throughout, $H(A) = C(A)^{\mathsf{T}}/\operatorname{per} A$ and $\Lambda(A) = \lambda_{\max}(H(A))$ as in Section 36.

Lemma 69 (Gram positivity)**.** *Let A be a correlation matrix of order n. Then $\operatorname{per} A > 0$, and $C(A)$ and $C(A)^{\mathsf{T}}$ are positive semidefinite.*

Proof. Let g be a centered complex Gaussian vector with $\mathbb{E}[g_k \overline{g_l}] = A_{kl}$ and $\mathbb{E}[g_k g_l] = 0$. Wick's formula for complex Gaussian moments [9] gives, for index sets S, T of equal size, $\mathbb{E}[\prod_{k\in S} g_k \prod_{l\in T} \overline{g_l}] = \operatorname{per}[A_{kl}]_{k\in S, l\in T}$. In particular $\operatorname{per} A = \mathbb{E}\prod_k |g_k|^2$, which is positive because each g_k is a nondegenerate complex Gaussian ($\mathbb{E}|g_k|^2 = 1$) and hence nonzero almost surely, and $C(A)_{ij} = \mathbb{E}[\prod_{k\neq i} g_k \prod_{l\neq j} \overline{g_l}]$. For $c \in \mathbb{C}^n$ this gives $\sum_{i,j} \overline{c_i} C(A)_{ij} c_j = \mathbb{E}|\sum_i \overline{c_i} \prod_{k\neq i} g_k|^2 \geq 0$. The transpose of a Hermitian positive semidefinite matrix is its entrywise conjugate, which is again positive semidefinite. □

Lemma 70 (bordering identities)**.** *Let $A = \begin{pmatrix} T & z \\ z^* & 1 \end{pmatrix}$ be a correlation matrix of order $n \geq 2$, with $p = \operatorname{per} T$ and $q = z^* C(T)^{\mathsf{T}} z$. Then*

$$\operatorname{per} A = p + q, \qquad C(A)_{nn} = p, \qquad C(A)_{nj} = (C(T)^{\mathsf{T}} z)_j \quad (j < n).$$

Consequently, with $x = H(A)_{nn}$ *and* $r = (H(A)_{jn})_{j<n}$,

$$x = \frac{p}{p+q} \in (0,1], \qquad r = xH(T)z, \qquad z^*H(T)z = \frac{1-x}{x}, \qquad \|r\|^2 \le \Lambda(T)\, x(1-x).$$

Proof. Expand per A along its last row: the diagonal entry contributes p, and the entry $\overline{z_l}$ in column $l < n$ multiplies $\operatorname{per} A(n \mid l)$. Expanding that minor along its last column, which is z without its n-th entry, gives $\operatorname{per} A(n \mid l) = \sum_{k<n} z_k \operatorname{per} T(k \mid l)$. Hence $\operatorname{per} A = p + \sum_{k,l<n} z_k \overline{z_l}\, C(T)_{kl} = p + q$. The same expansion of the minor $A(n \mid j)$ along its last column gives $C(A)_{nj} = \sum_{k<n} z_k C(T)_{kj} = (C(T)^{\mathsf{T}} z)_j$, and $C(A)_{nn} = \operatorname{per} T$. Dividing by $\operatorname{per} A = p + q$, and using $p > 0$ and $q \ge 0$ from Lemma 69, gives the normalized identities. Finally $\|H(T)z\|^2 \le \lambda_{\max}(H(T))\, z^*H(T)z$ because $H(T) \succeq 0$, so $\|r\|^2 \le x^2\Lambda(T)(1-x)/x$. □

Lemma 71 (recursive upper bound and the general lower bound)**.** *Let* $c_1 = 1$ *and* $c_n = \big(1 + \sqrt{1 + (n-1)c_{n-1}}\big)/2$ *for* $n \ge 2$. *Then for every* $n \ge 2$

$$\frac{(n+1)^2}{4n} < \gamma_n \le c_n \le \frac{n^2 + 2n + 3}{4n},$$

so that $c_n = n/4 + 1/2 + O(1/n)$.

Proof. Clearly $\gamma_1 = 1 = c_1$, and $c_n \ge 1$. Let A have order $n \ge 2$ with $\ell = \Lambda(A) > 1$ (if $\ell \le 1$ there is nothing to prove), and suppose $\gamma_{n-1} \le c_{n-1}$. Lemma 66, whose proof uses only Lemmas 69 and 70, gives for a unit top eigenvector e

$$1 = \sum_{i=1}^{n} |e_i|^2 \le \sum_{i=1}^{n} \frac{\Lambda(A_i)}{\Lambda(A_i) + 4\ell(\ell-1)} \le \frac{n c_{n-1}}{c_{n-1} + 4\ell(\ell-1)},$$

so $4\ell(\ell-1) \le (n-1)c_{n-1}$ and $\ell \le c_n$. For the comparison with $u_n = (n^2 + 2n + 3)/(4n)$, induct: $c_1 = 1 < u_1$, and if $c_{n-1} \le u_{n-1}$ then $1 + (n-1)u_{n-1} = (n^2+6)/4 \le (n + 3/n)^2/4$, so $c_n \le (1 + (n + 3/n)/2)/2 = u_n$. For the lower bound take $b = 2/(n+1)$ in (1), so that $t = (1-b)/b = (n-1)/2$; the Poisson term is positive, hence $\Lambda(B(b)) > (1+t)(1-t/n) = (n+1)^2/(4n)$. □

The next two statements give the unrestricted maximum in order three. For an order-three correlation matrix with off-diagonal entries A_{12}, A_{23}, A_{31}, write x, y, z for their squared moduli, $s = x + y + z$, and $\tau = \operatorname{Re}(A_{12}A_{23}A_{31})$ for the real triangle product.

Lemma 72 (order-three invariants and equalization)**.** *For an order-three correlation matrix* A,

$$\operatorname{per} A = 1 + s + 2\tau, \qquad \operatorname{tr} C(A) = 3 + s, \qquad e_2(C(A)) = 3 + s - 6\tau,$$
$$\det C(A) = (1 - 2\tau)^2 - (s-2)^2 + 4(1-x)(1-y)(1-z).$$

Let A' *have all three off-diagonal moduli equal to* $\sqrt{s/3}$ *and real triangle product* τ. *Then* A' *is a correlation matrix with* $\operatorname{per} A' = \operatorname{per} A$ *and the same eigenvalues as* A, *and* $\Lambda(A') \ge \Lambda(A)$, *with equality only if* $x = y = z$.

Proof. After a diagonal unitary conjugation $A_{12} = \alpha \ge 0$, $A_{23} = \beta \ge 0$ and $A_{31} = w \in \mathbb{C}$, so that $\tau = \alpha\beta \operatorname{Re} w$. The cofactors are $C_{11} = 1 + \beta^2$, $C_{22} = 1 + |w|^2$, $C_{33} = 1 + \alpha^2$, $C_{12} = \alpha + \beta w$, $C_{13} = \alpha\beta + w$ and $C_{23} = \beta + \alpha w$, with $C_{ji} = \overline{C_{ij}}$. Direct expansion, checked symbolically in `verify_order_three.py`, gives the four displayed formulas; in the last one $4(1-x)(1-y)(1-z)$

collects the terms $4(xy+yz+zx-xyz)-4s+4$. The characteristic polynomial $\det(\lambda-A)=(\lambda-1)^3-s(\lambda-1)-2\tau$ depends only on s and τ.

The moduli of A satisfy $|\tau|\le\sqrt{xyz}\le(s/3)^{3/2}$ by the arithmetic–geometric mean inequality, so a phase with $(s/3)^{3/2}\cos\phi=\tau$ exists, and the matrix A' with entries of modulus $\sqrt{s/3}$ and triangle argument ϕ has the same characteristic polynomial as A; in particular it is positive semidefinite, and $\operatorname{per}A'=\operatorname{per}A$. The matrices $C(A)$ and $C(A')$ have the same trace and the same e_2, and $\det C(A')-\det C(A)=4[(1-s/3)^3-(1-x)(1-y)(1-z)]\ge 0$ by the arithmetic–geometric mean inequality for the nonnegative numbers $1-x,1-y,1-z$, with equality only if $x=y=z$. Thus $\det(\lambda-C(A'))=\det(\lambda-C(A))-\Delta$ with $\Delta\ge 0$; at the largest root of the second polynomial the first is $-\Delta$, so $\lambda_{\max}(C(A'))\ge\lambda_{\max}(C(A))$, strictly if $\Delta>0$. □

For an equal-modulus matrix choose a diagonal unitary gauge in which $A_{12}=A_{23}=A_{31}=be^{\mathrm{i}\theta}$ with $b\ge 0$ and $|\theta|\le\pi/3$, and put

$$a=b\cos\theta,\qquad \zeta=b^2\sin^2\theta,\qquad N(a)=1+2a+3a^2,\qquad D(a)=1+3a^2+2a^3,\qquad F_3=N/D.$$

Then $s=3(a^2+\zeta)$ and $\tau=a^3-3a\zeta$.

Theorem 73 (order three)**.** *For an equal-modulus order-three correlation matrix in this gauge,*

$$\Lambda(A)=\frac{N(a)-\zeta}{D(a)+3(1-2a)\zeta}.$$

If $a<4^{-1/3}$, then $\Lambda(A)\le F_3(a)$, with equality only if $\zeta=0$; if $a\ge 4^{-1/3}$, then $\Lambda(A)\le 4/3$. Consequently $\gamma_3=F_3(r_*)$, where $r_*\in(0.35771,0.35773)$ is the unique positive root of $3r^4+4r^3+6r^2-1$, and every order-three correlation matrix with $\Lambda(A)=\gamma_3$ is a diagonal unitary conjugate of $(1-r_*)I+r_*J$.

Proof. In this gauge A is the Hermitian circulant with $A_{j,j+1}=be^{\mathrm{i}\theta}$, so $C(A)$ is a circulant. Its eigenvalues are $1-b^2+2bc_k+4b^2c_k^2$ with $c_k=\cos(\theta+2\pi k/3)$, $k=0,1,2$; this follows by applying $C(A)$ to the Fourier vectors and is checked in `verify_order_three.py`. The quadratic $c\mapsto 1-b^2+2bc+4b^2c^2$ is convex and symmetric about $-1/(4b)$. For $|\theta|\le\pi/3$ the largest of the c_k is $\cos\theta$, and the largest and smallest sum to $\cos(|\theta|+\pi/3)\ge -1/2\ge -1/(2b)$ since $b\le 1$; hence the top eigenvalue is attained at $c=\cos\theta$ and equals $1-b^2+2a+4a^2=N(a)-\zeta$. With $\operatorname{per}A=1+s+2\tau=1+8a^3+3(1-2a)b^2=D(a)+3(1-2a)\zeta$ this gives the formula.

Write $G(a,d)=(1-d+2a+4a^2)/(1+8a^3+3(1-2a)d)$ with $d=b^2=a^2+\zeta$. Its d-derivative has the sign of $-(1+8a^3)-3(1-2a)(1+2a+4a^2)=16a^3-4$. If $a<4^{-1/3}$, G decreases in d, so $\Lambda(A)\le G(a,a^2)=F_3(a)$, with equality only at $d=a^2$. If $a\ge 4^{-1/3}$, G increases in d, and positive semidefiniteness, $\det A=1-3d+8a^3-6ad\ge 0$, gives $d\le(1-2a+4a^2)/3$; at this endpoint $G=(1+2a)^2/[3(1-2a+4a^2)]\le 4/3$, because $4(1-2a+4a^2)-(1+2a)^2=3(1-2a)^2\ge 0$.

Next, $F_3'=2(1-6a^2-4a^3-3a^4)/D^2$. The numerator decreases on $a>0$, is positive at zero and negative at one, so F_3 increases up to the unique positive root r_* and decreases afterwards; exact evaluation locates $r_*\in(0.35771,0.35773)$, and $F_3(r_*)>F_3(5/14)>4/3$ while $r_*<4^{-1/3}$. For an arbitrary order-three correlation matrix, Lemma 72 and the above give $\Lambda(A)\le\max\{\sup_a F_3(a),4/3\}=F_3(r_*)$, and $A=(1-r_*)I+r_*J$ attains it. If $\Lambda(A)=F_3(r_*)$, equality in Lemma 72 forces equal moduli, the bound $4/3$ excludes the high branch, equality in the first bound forces $\zeta=0$, and the uniqueness of the maximizer of F_3 forces $a=r_*$. Thus all moduli equal r_* and the triangle product is positive; the diagonal unitary that makes A_{12} and A_{13} positive then makes A_{23} positive too. □

Lemma 74 (roots of unity in rank two)**.** *Let $n \geq 3$, let $u = (\omega^k)_{k=0}^{n-1}$ with $\omega = e^{2\pi \mathrm{i}/n}$, and for $0 < t < 1$ let $A_t = (1-t)J + tuu^*$. Then A_t is a correlation matrix and*

$$u^* A_t u = n^2 t, \qquad u^* C(A_t)^{\mathsf{T}} u = n\, n!\, t^{n-1}, \qquad \operatorname{per} A_t = n!\big[(1-t)^n + t^n\big].$$

Proof. A_t is positive semidefinite with unit diagonal, and $u^*\mathbf{1} = 0$ gives the first identity. With independent standard complex Gaussians Z_1, Z_2, the vector $g_k = \sqrt{1-t}\, Z_1 + \sqrt{\tau}\, \omega^k Z_2$ has covariance $(1-t)J + \tau uu^*$ for $\tau \geq 0$, and $\prod_k (X + \omega^k Y) = X^n - (-Y)^n$. Hence, by the Wick formula used in the proof of Lemma 69, which holds for every positive semidefinite covariance, $\operatorname{per}((1-t)J + \tau uu^*) = \mathbb{E}|(\sqrt{1-t}Z_1)^n - (-\sqrt{\tau}Z_2)^n|^2 = n![(1-t)^n + \tau^n]$, the cross term having mean zero. At $\tau = t$ this is the third identity. Differentiating in τ at $\tau = t$ gives $\sum_{j,k} C(A_t)_{jk} u_j \overline{u_k} = u^* C(A_t)^{\mathsf{T}} u = n\, n!\, t^{n-1}$. □

Statements and Declarations

Funding. This research was self-funded.

Competing interests. The author owns Luminary Oddities and has no other relevant competing interests.

Data and code availability. The verification scripts cited in the text are included with this article.

AI assistance. OpenAI Codex assisted with derivations, proof development, source checks and independent internal review of each part through Section 65. OpenAI ChatGPT supplied working notes and a manuscript for Section 66 (the circulant reduction, the Bernstein certificate strategy, the order-five second variation, the local radius and the cyclic dephasing rate), and OpenAI Codex supplied a manuscript for Sections 65.16–65.17; these were independently re-derived and re-verified. Anthropic Claude edited, consolidated and cross-referenced the parts, checked computations, and assisted with derivations, proof development, computations, source checks and independent internal review for Section 66. Scott Schubr directed the research and is responsible for the content.

References

[1] E. H. Lieb, *Proofs of some conjectures on permanents*, J. Math. Mech. **16** (1966), 127–134. doi:10.1512/iumj.1967.16.16008.

[2] L. Pioge, K. K. Pietrasz, B. Seron, L. Novo, and N. J. Cerf, *A logical implication between two conjectures on matrix permanents*, Linear Algebra Appl. **725** (2025), 309–318. doi:10.1016/j.laa.2025.07.011; arXiv:2508.00111.

[3] Y. Lavi, *A conjecture on the permanent of positive semidefinite matrices*, 2026. Preprint, version 1; also doi:10.5281/zenodo.23001182.

[4] M. Marcus and R. Merris, *A relation between the permanental and determinantal adjoints*, J. Austral. Math. Soc. **15** (1973), 270–271. doi:10.1017/S1446788700013173.

[5] R. Grone, C. R. Johnson, E. M. Sá, and H. Wolkowicz, *A note on maximizing the permanent of a positive definite Hermitian matrix, given the eigenvalues*, Linear Multilinear Algebra **19** (1986), 389–393. Author copy.

[6] R. Bhatia, P. Grover, and T. Jain, *Derivatives of tensor powers and their norms*, Electron. J. Linear Algebra **26** (2013), 604–619. doi:10.13001/1081-3810.1675.

[7] P. Feinsilver and J. McSorley. Zeons, permanents, the Johnson scheme, and generalized derangements. *International Journal of Combinatorics*, 2011, Article 539030. doi:10.1155/2011/539030.

[8] S. Sra. A Koteljanskii inequality for permanents. arXiv:2609.39979, 2026.

[9] C. Fassino, G. Pistone, and M. P. Rogantin, *Computing the moments of the complex Gaussian: Full and sparse covariance matrix*, Mathematics **7** (2019), 263. doi:10.3390/math7030263.

[10] A. Barvinok. Two algorithmic results for the traveling salesman problem. *Mathematics of Operations Research*, 21(1):65–84, 1996. doi:10.1287/moor.21.1.65.

[11] A. Barvinok. A remark on approximating permanents of positive definite matrices. *Linear Algebra and its Applications*, 608:399–406, 2021. arXiv:2005.06344.

[12] P. McCullagh, *An asymptotic approximation for the permanent of a doubly stochastic matrix*, Author manuscript.

[13] A. Hyvärinen. Estimation of non-normalized statistical models by score matching. *Journal of Machine Learning Research*, 6:695–709, 2005. Journal article.

[14] L. Devroye, A. Mehrabian, and T. Reddad. The total variation distance between high-dimensional Gaussians with the same mean. arXiv:1810.08693.

[15] R. Adamczak and P. Wolff. Concentration inequalities for non-Lipschitz functions with bounded derivatives of higher order. arXiv:1304.1826.

[16] F. Voigtlaender. A general version of Price's theorem. *Journal of Theoretical Probability*, 34:1474–1485, 2021. doi:10.1007/s10959-020-01017-w.

[17] S. Caracciolo, A. D. Sokal, and A. Sportiello. Algebraic/combinatorial proofs of Cayley-type identities for derivatives of determinants and pfaffians. *Advances in Applied Mathematics*, 50(4):474–594, 2013. Published version.

[18] A. S. Bandeira, N. Boumal, and A. Singer. Tightness of the maximum likelihood semidefinite relaxation for angular synchronization. *Mathematical Programming*, 163:145–167, 2017. arXiv:1411.3272.

[19] H. Hatami. Refutation of the Exponential Time Hypothesis, lecture notes, 2026. Lecture notes.

[20] K. Rodtes and M. F. Anwar. Permanents of block matrices. *Linear Algebra and its Applications*, 698:94–101, 2024. doi:10.1016/j.laa.2024.05.024.

[21] A. Ottolini and S. Starr. Pal's permanent conjecture: proof for block uniform matrices. arXiv:2605.25274.

[22] M. A. Aouichaoui and F. Zhang. A solution to a permanent conjecture of Pate. Preprints 202609.2464, 2026.

[23] S. Rahimi-Keshari, A. P. Lund, and T. C. Ralph, *What can quantum optics say about computational complexity theory?*, Phys. Rev. Lett. **114** (2015), 060501. arXiv:1408.3712.

[24] N. M. Temme, *Uniform asymptotic expansions of the incomplete Gamma functions and the incomplete Beta function*, Math. Comp. **29** (1975), 1109–1114. CWI archive.

[25] K. Kirkpatrick and T. Nawaz, *Asymptotics of mean-field $O(N)$ models*, arXiv:1602.03257.

[26] N. Klein, J. Orellana, S. Brincat, E. K. Miller, and R. E. Kass, *Torus graphs for multivariate phase coupling analysis*, arXiv:1910.11044.

[27] S.-G. Hwang and R. Meshulam. An inequality for subpermanents of positive semidefinite Hermitian matrices. *Linear and Multilinear Algebra*, 38:177–180, 1995.

[28] R. B. Bapat and V. S. Sunder, On majorization and Schur products, *Linear Algebra Appl.* **72** (1985), 107–117. DOI.

[29] S. W. Drury, A counterexample to a question of Bapat and Sunder, *Electron. J. Linear Algebra* **31** (2016), 69–70. doi:10.13001/1081-3810.3178.

[30] V. S. Shchesnovich, The permanent-on-top conjecture is false, *Linear Algebra Appl.* **490** (2016), 196–201. DOI.

[31] P. Pant and R. Singh, Structural classes for Chollet's permanent conjecture, arXiv:2604.24192v2, 2026.

[32] S. Sa-nguansin and K. Rodtes, Permanents of correlation matrices and the Chollet permanental conjecture, *Linear Multilinear Algebra* **73** (2025), 4084–4096. DOI.

[33] F. Zhang, Notes on Hadamard products of matrices, *Linear Multilinear Algebra* **25** (1989), 237–242. DOI.

[34] J. Chollet, Is there a permanental analogue to Oppenheim's inequality?, *Amer. Math. Monthly* **89** (1982), 57–58. DOI.

[35] N. Anari, L. Gurvits, S. Oveis Gharan and A. Saberi, Simply exponential approximation of the permanent of positive semidefinite matrices, *Proc. 58th IEEE FOCS* (2017), 914–925. arXiv:1704.03486.

[36] G. James and A. Kerber, *The Representation Theory of the Symmetric Group*, Addison-Wesley, 1981.

[37] A. Bristiel and P. Caputo, *Entropy inequalities for random walks and permutations*, Ann. Inst. H. Poincaré Probab. Statist. 60 (2024). arXiv:2109.06009.

[38] F. Zhang, *An update on a few permanent conjectures*, Spec. Matrices **4** (2016), 305–316. doi:10.1515/spma-2016-0030; arXiv:1608.02844.

[39] G. W. Soules, *An approach to the permanental-dominance conjecture*, Linear Algebra Appl. **201** (1994), 211–229. doi:10.1016/0024-3795(94)90117-1.

[40] R. Merris, *The permanental dominance conjecture*, in *Current Trends in Matrix Theory* (F. Uhlig and R. Grone, eds.), Elsevier, New York, 1987, 213–223.

[41] Y. Wu and P. Yang, *Optimal estimation of Gaussian mixtures via denoised method of moments*, preprint, arXiv:1807.07237; published in Ann. Statist. 48 (2020), 1981–2007. Proposition numbers refer to the arXiv version.

[42] P. Diaconis, *The cutoff phenomenon in finite Markov chains*, Proc. Natl. Acad. Sci. USA **93** (1996), 1659–1664. doi:10.1073/pnas.93.4.1659.

[43] E. Lubetzky and A. Sly, *Cutoff phenomena for random walks on random regular graphs*, Duke Math. J. **153** (2010), 475–510. doi:10.1215/00127094-2010-029.

[44] G.-Y. Chen and L. Saloff-Coste, *The cutoff phenomenon for ergodic Markov processes*, Electron. J. Probab. **13** (2008), 26–78. doi:10.1214/EJP.v13-474.

[45] J. Salez, *Cutoff for non-negatively curved Markov chains*, J. Eur. Math. Soc. **26** (2024), 4375–4392. doi:10.4171/JEMS/1348.

[46] E. Nestoridi and S. Olesker-Taylor, *Limit profiles for reversible Markov chains*, Probab. Theory Related Fields **182** (2022). doi:10.1007/s00440-021-01061-5.

[47] L. Teyssier, *Limit profile for random transpositions*, Ann. Probab. 48 (2020), 2323–2343. arXiv:1905.08514.

[48] L. Teyssier, *Cutoff profiles for conjugacy invariant random walks on symmetric groups*, arXiv:2605.28770, 2026.

[49] V. Féray and P. Śniady, *Asymptotics of characters of symmetric groups related to Stanley character formula*, Ann. of Math. 173 (2011), 887–906. Article.

[50] P. Diaconis and M. Shahshahani, *Time to reach stationarity in the Bernoulli–Laplace diffusion model*, SIAM J. Math. Anal. **18** (1987), 208–218. doi:10.1137/0518016.

[51] J. Shen, *The cutoff profile for random transpositions on repeated cards in the full range of parameters*, preprint, 2026. arXiv:2604.23890.

[52] S. Olesker-Taylor and D. Schmid, *Limit profile for the Bernoulli–Laplace urn*, preprint, 2024. arXiv:2409.07900.

[53] R. Goenka and K.-I. Prodromidis, *Cutoff profile for the balanced generalized Bernoulli–Laplace model*, preprint, 2026. arXiv:2610.28550.

[54] E. Nestoridi and S. Olesker-Taylor, *Limit profiles for projections of random walks on groups*, Electron. J. Probab. **29** (2024), 1–22. doi:10.1214/24-EJP1207.

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

[56] Y. Filmus, R. O'Donnell and X. Wu, *Log-Sobolev inequality for the multislice, with applications*, Electron. J. Probab. **27** (2022), 1–30. doi:10.1214/22-EJP749; theorem numbers refer to arXiv:1809.03546v1.

[57] P. Diaconis and L. Saloff-Coste, *Logarithmic Sobolev inequalities for finite Markov chains*, Ann. Appl. Probab. **6** (1996), 695–750. doi:10.1214/aoap/1034968224.

[58] R. B. Bapat and V. S. Sunder, *An extremal property of the permanent and the determinant*, Linear Algebra Appl. **76** (1986), 153–163. doi:10.1016/0024-3795(86)90220-X.

[59] S. W. Drury, *A counterexample to a question of Bapat & Sunder*, Math. Inequal. Appl. **21** (2018), 517–520. doi:10.7153/mia-2018-21-37.

[60] M. C. Tichy, *Sampling of partially distinguishable bosons and the relation to the multidimensional permanent*, Phys. Rev. A **91** (2015), 022316. doi:10.1103/PhysRevA.91.022316.